Ashley J. Welch
Martin J.C. van Gemert *Editors*

Optical-Thermal Response of Laser-Irradiated Tissue

Second Edition

Springer

Optical-Thermal Response of Laser-Irradiated Tissue

Ashley J. Welch · Martin J.C. van Gemert
Editors

Optical-Thermal Response of Laser-Irradiated Tissue

Second Edition

Springer

Editors
Dr. Ashley J. Welch
Dept. Biomedical Engineering
The University of Texas at Austin
1 University Station, C0800
Austin, TX 78712
USA
welch@mail.utexas.edu

Dr. Martin J.C. van Gemert
Dept. of Biomedical Engineering
and Physics
Academic Medical Center
Meibergdreef 9
1105 AZ Amsterdam
The Netherlands
m.j.vangemert@amc.uva.nl

ISBN 978-90-481-8830-7 e-ISBN 978-90-481-8831-4
DOI 10.1007/978-90-481-8831-4
Springer Dordrecht Heidelberg London New York

Preface

This book presents theory and applications of laser light interaction with tissue. In recent years, propagation of light in tissue and the resulting photochemical or photothermal response has received considerable attention owing to the increased number of diagnostic and therapeutic applications of lasers in medicine.

To address these issues, his book is divided into three parts: (1) Tissue Optics; (2) Thermal Interactions; and (3) Medical Applications. In the first part, the basis for light propagation is presented in terms of energy transport. Both steady state and time of flight light propagation are discussed, and governing equations are formulated that incorporate light scattering in tissue. In this second edition, chapters have been added that describe light propagation based upon Maxwell's Equations. Emphasis is placed upon polarization and how these formulations can be incorporated into Monte Carlo simulations.

The second part of the book describes the thermal response of tissue to light. It begins with fundamental concepts in heat conduction including analytic and numerical solutions which provide estimates of the transient temperature field produced by the absorption of laser light. Next, the effect of elevated temperature denaturing tissue components is described in terms of first order rate reductions. Part II includes chapters concerning temperature measurement and tissue thermal properties. The last part of the book (Part III) is devoted to applications of lasers in medicine which can be analyzed or described using the material of Parts I and II. Part III covers optics of fibers, fluorescence, ablation, and Port Wine Stains.

Each chapter is written by a leading authority in the field. All authors have used a common notation, and the material is arranged so there is a progression of material from elementary concepts to our current understanding of laser-tissue interaction. Throughout the book, basic concepts are emphasized; this is not a book of medical applications of lasers. We believe the contents are suitable for understanding the fundamental interactions of laser light and tissue. It is our hope that this material will permit an analysis of proposed medical applications of lasers and provide the basis for evaluating diagnostic systems and establishing dosimetry for therapeutic applications of laser light.

Austin, Texas — Ashley J. Welch
Amsterdam, The Netherlands — Martin J.C. van Gemert

Acknowledgement

The second edition would not have been possible without the contributions of our colleagues, many of whom have been our students. The integration and uniformity of this material is largely due to the efforts of Chris Humphrey. She is the true "Editor" of this book. Chris has been the interface between authors and the retired editors relaxing at home, and she is the final judge of the material (text to figures) comprising this book. Chris performed the same demanding task for the first edition and yet she still agreed to come out of her own retirement to oversee the completion of this book. For one of us (A.J.), it is especially fitting that this be the final accomplishment of our work. We thank all contributors and our readers for sharing the vision that the material of the Second Edition will benefit students, faculty and industry.

Austin, Texas — Ashley J. Welch
Amsterdam, The Netherlands — Martin J.C. van Gemert

Contents

Contributors

R. Rox Anderson Wellman Center for Photomedicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02114, USA, RRAnderson@partners.org

Jennifer K. Barton Department of Biomedical Engineering, University of Arizona, Tucson, AZ, USA, barton@u.arizona.edu

Jonathan M. Cayce Departments of Biomedical Engineering and Neurosurgery, Vanderbilt University, Nashville, TN 37235, USA, jonathan.cayce@vanderbilt.edu

Lee C.L. Chin Medical Physics Department, Odette Cancer Centre, Toronto, ON, Canada, Lee.Chin@sunnybrook.ca

Christopher H. Contag Departments of Pediatrics, Radiology, and Microbiology and Immunology, Molecular Imaging Program, Stanford University, Palo Alto, CA, USA, ccontag@stanford.edu

Kenneth R. Diller Department of Biomedical Engineering, The University of Texas, Austin, TX, USA, kdiller@mail.utexas.edu

Stanislav Emelianov Department of Biomedical Engineering, The University of Texas, Austin, TX, USA

Dirk J. Faber Department of Biomedical Engineering and Physics, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands, d.j.faber@amc.uva.nl

Martin Frenz Institute of Applied Physics, Sidlerstr. 5, 3012, Bern, Switzerland, martin.frenz@iap.unibe.ch

Steven C. Gebhart Department of Biomedical Engineering, Vanderbilt University, Nashville, TN, USA, steven.c.gebhart@vanderbilt.edu

Olav Anton Haugen Norwegian University of Science and Technology, Trondheim, Norway, olav.haugen@ntnu.no

Steven L. Jacques Biomedical Engineering Department, Oregon Health and Science University, Portland, OR, USA, jacquess@ohsu.edu

E. Duco Jansen Departments of Biomedical Engineering and Neurosurgery, Vanderbilt University, Nashville, TN 37235, USA, duco.jansen@vanderbilt.edu

Anthony Kim Department of Medical Biophysics, University of Toronto and Ontario Cancer Institute, Toronto, ON M5G 2M9, Canada, anthonyk@uhnres.utoronto.ca

Peter E. Konrad Departments of Biomedical Engineering and Neurosurgery, Vanderbilt University, Nashville, TN 37235, USA, peter.konrad@vanderbilt.edu

Neil E.I. Langlois Forensic Science, University of Adelaide, Adelaide, Australia, Neil.Langlois@sa.gov.au

Mark A. Mackanos Department of Biomedical Engineering, Vanderbilt University, Nashville, TN, USA, mark.mackanos@vanderbilt.edu

Anita Mahadevan-Jansen Departments of Biomedical Engineering and Neurosurgery, Vanderbilt University, Nashville, TN 37235, USA, anita.mahadevan-jansen@vanderbilt.edu

Boris Majaron Jožef Stefan Institute, Ljubljana, Slovenia, boris.majaron@ijs.si

Massoud Motamedi Center for Biomedical Engineering, University of Texas Medical Branch, Galveston, TX, USA, massoud.motamedi@utmb.edu

J. Stuart Nelson Beckman Laser Institute, University of California, Irvine, CA, USA, jsnelson@uci.edu

John A. Pearce Department of Electrical and Computer Engineering, The University of Texas, Austin, TX, USA, jpearce@mail.utexas.edu

Lise Lyngsnes Randeberg Department of Electronics and Telecommunications, Norwegian University of Science and Technology, Trondheim, Norway, Lise.Randeberg@iet.ntnu.no

Jaro Rička Institute of Applied Physics, Sidlerstr. 5, 3012, Bern, Switzerland, jaroslav.ricka@iap.unibe.ch

Bjørn Skallerud Department of Structural Engineering, Norwegian University of Science and Technology, Trondheim, Norway, bjorn.skallerud@ntnu.no

Willem M. Star Center for Optical Diagnostics and Therapy, Erasmus Medical Center, Rotterdam, The Netherlands, wmstar@xs4all.nl

Lars Othar Svaasand Department of Physical Electronics, Norwegian University of Science and Technology, Trondheim, Norway, svaasand@fysel.ntnu.no

Sharon Thomsen 500 Discovery View Drive, Sequim, WA, USA, slthomsenmd@yahoo.com

Jonathan W. Valvano Department of Electrical and Computer Engineering, The University of Texas, Austin, TX, USA, valvano@mail.utexas.edu

Martin J.C. van Gemert Department of Biomedical Engineering and Physics, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands, m.j.vangemert@amc.uva.nl

Ton G. van Leeuwen Department of Biomedical Engineering and Physics, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands, A.G.J.M.vanLeeuwen@tnw.utwente.nl

Vasan Venugopalan Department of Chemical Engineering and Materials Science and Laser Microbeam and Medical Program, Beckman Laser Institute, University of California, Irvine, CA, USA, vvenugop@uci.edu

I. Alex Vitkin Division of Biophysics and Bioimaging, Ontario Cancer Institute, Toronto, ON, Canada, Alex.Vitkin@rmp.uhn.on.ca

Alfred Vogel Institute of Biomedical Optics, University of Lübeck, Lübeck, Germany, vogel@bmo.uni-luebeck.de

Joseph T. Walsh Biomedical Engineering Department, Northwestern University, Evanston, IL, USA, jwalsh@northwestern.edu

Lihong V. Wang Department of Biomedical Engineering, Washington University, St. Louis, MO 63130-4899, USA, lhwang@seas.wustl.edu

Ashley J. Welch Department of Biomedical Engineering, The University of Texas, Austin, TX, USA, welch@mail.utexas.edu

Jonathon D. Wells Departments of Biomedical Engineering and Neurosurgery, Vanderbilt University, Nashville, TN 37235, USA, Jonathon.wells@lmco.com

William M. Whelan Department of Physics, University of Prince Edward Island; Department of Biomedical Sciences, Atlantic Veterinary College, Charlottetown, PEI, Canada, wwhelan@upei.ca

Brian C. Wilson Department of Medical Biophysics, University of Toronto and Ontario Cancer Institute, Toronto, ON M5G 2M9, Canada, wilson@uhnres.utoronto.ca

Part I
Tissue Optics

Chapter 1
Overview of Optical and Thermal Laser-Tissue Interaction and Nomenclature

Ashley J. Welch and Martin J.C. van Gemert

1.1 Introduction

The development of a unified theory for the optical and thermal response of tissue to laser radiation is no longer in its infancy, though it is still not fully developed. This book describes our current understanding of the physical events that can occur when light interacts with tissue, particularly the sequence of formulations that estimate the optical and thermal responses of tissue to laser radiation. This overview is followed by an important chapter that describes the basic interactions of light with tissue. Part I considers basic tissue optics. Tissue is treated as an absorbing and scattering medium and methods are presented for calculating and measuring light propagation, including polarized light. Also, methods for estimating tissue optical properties from measurements of reflection and transmission are discussed. Part II concerns the thermal response of tissue owing to absorbed light, and rate reactions are presented for predicting the extent of laser induced thermal damage. Methods for measuring temperature, thermal properties, rate constants, pulsed ablation and laser tissue interactions are detailed. Part III is devoted to examples that use the theory presented in Parts I and II to analyze various medical applications of lasers. Discussions of Optical Coherence Tomography (OCT), forensic optics, and light stimulation of nerves are also included.

Since the optical and thermal responses of tissue to laser irradiation are highly dependent upon the characteristics of the laser source, we will describe various laser parameters and their influence upon the tissue response. Basic irradiation parameters for any laser therapeutic procedure are (1) power, irradiation time, and spot size for continuous wave (*cw*) lasers, and (2) energy per pulse, irradiation time, spot size, repetition rate and number of pulses for pulse lasers. Perhaps the most important *cw* irradiation parameter is irradiance [W/m^2]. The irradiance is a function of the power delivered and the laser spot size on the tissue. Low irradiances that do not significantly increase tissue temperature are associated with diagnostic applications

A.J. Welch (✉)
Department of Biomedical Engineering, The University of Texas, Austin, TX, USA
e-mail: welch@mail.utexas.edu

A.J. Welch, M.J.C. van Gemert (eds.), *Optical-Thermal Response of Laser-Irradiated Tissue*, 2nd ed., DOI 10.1007/978-90-481-8831-4_1,

such as OCT, photochemical processes and biostimulation, whereas high irradiances can ablate tissue, induce plasma formation, and create mechanical damage in tissue. The depth that laser light penetrates tissue depends upon the optical properties of tissue which vary with wavelength.

The spectrum of commercially available lasers for diagnostic and therapeutic medical applications stretches from 193 nm to 10.6 μm. Both *cw* and pulse lasers are available in the UV, visible and IR. Since the absorption and scattering of any tissue varies with wavelength, there are dramatic differences in the penetration depth of the radiation from the various lasers. Light at either 193 nm or 2.96 μm is totally absorbed in the first few μm of tissue owing to amino acid absorption in the UV and water absorption in the IR. In contrast, collimated light from 600 nm to 1.2 μm can penetrate several millimeters in tissue and the associated scattered light several cm. Within this red and near IR wavelength window there is a lack of strongly absorbing tissue chromophores. As the collimated beam passes through tissue, it is exponentially attenuated by absorption and scattering. The scattered light forms a diffuse volume around the collimated beam. Heat is generated wherever collimated or diffuse light is absorbed.

This book provides a detailed description of the optical and thermal response of tissue to laser radiation. We examine many of the physical events that are involved in the response of tissue to laser irradiation. Particular emphasis is placed on events that provide some insight into the optical response during treatment procedures or diagnostic applications of laser light.

These events are schematically depicted in Fig. 1.1 and they comprise a simplified optical, thermal and damage model for laser-tissue interaction. In this diagram, we introduce the natural sequence of laser induced events and the governing equations that will be used throughout the book. Modifications, boundary conditions, and solutions of these equations are given in the following chapters.

In Fig. 1.1, we assume that a laser beam with irradiance profile $E(x,y)$ strikes the tissue. The collimated beam in the z direction attenuates exponentially with tissue depth; light scattered from the collimated beam becomes the source for the resulting diffuse (scattered) light in the tissue. Propagation of the scattered light is described by the transport equation [1] (see Chapter 3) which examines the change in radiance with distance in direction $\hat{\boldsymbol{s}}$ at a position $\boldsymbol{r} = x, y, z$. Light (radiance) in direction $\hat{\boldsymbol{s}}$ decreases owing to absorption and scattering (first term on right side of transport equation in Fig. 1.1) and increases due to light that is scattered from other directions $\hat{\boldsymbol{s}}'$ into direction $\hat{\boldsymbol{s}}$ (second term on right side). The total fluence rate, $\phi(\boldsymbol{r})$, is the integration of radiance $L(\boldsymbol{r},\ \hat{\boldsymbol{s}})$ over all directions in space (over 4π steradians). In this book we separate radiance into its two components: scattered light $L_s(\boldsymbol{r},\ \hat{\boldsymbol{s}})$ and collimated (primary) light $L_p(\boldsymbol{r},\ \hat{\boldsymbol{s}}_0)$ (see Chapters 3 and 6).

Most of the light (collimated or diffuse) that is absorbed is converted to heat. Exceptions are light that produces photochemical reactions or photons absorbed by fluorophores that produce fluorescence. Typically, the light energy associated with these reactions is very small relative to the energy that produces thermal reactions. The rate of volumetric heat production $S(\boldsymbol{r})$ is the product of the total fluence rate at $\boldsymbol{r}$ (collimated and diffuse parts) and the absorption coefficient at $\boldsymbol{r}$, $\mu_a(\boldsymbol{r})$.

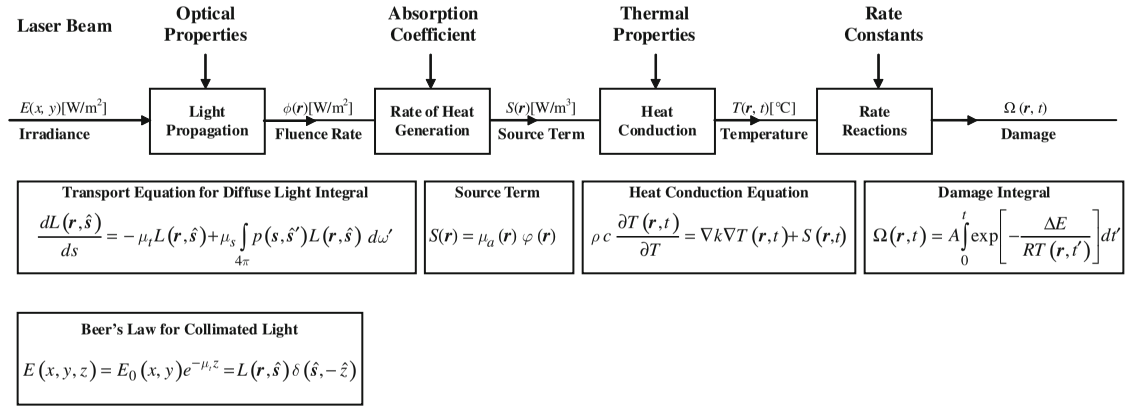


Fig. 1.1 Block diagram of optical and thermal laser tissue interactions. Equations are developed in the various chapters describing tissue optics and thermal interactions

The resulting temperature increase $\Delta T(\boldsymbol{r},\ t)$ produced by the absorbed light and the diffusion of heat to colder regions is given by the heat conduction equation [2] (Chapter 10). Increasing tissue temperature increases reaction rates that can lead to tissue denaturation. These rate processes are represented in the damage integral Ω, based on the first order Arrhenius equation. The block diagram of Fig. 1.1 illustrates that the accuracy of solutions for a particular response is dependent upon the accuracy of the solution of previous events. That is, computation of damage requires accurate predictions of temperature with time which requires knowledge of the rate of heat production which, in turn, is dependent upon an accurate estimate of the fluence rate throughout the tissue. All of these computations require specification of laser parameters at the site of radiation and knowledge of the optical and thermal properties and rate constants of the tissue.

Optical and thermal parameters are not constants, but can dynamically change depending upon the condition of the tissue. Temperature, (de)hydration and thermal damage can alter the absorption and scattering properties of tissue. The most striking example is egg albumin. The clear albumin becomes white when heat coagulates the medium. During coagulation, scattering dramatically increases. Thus the optical properties of a tissue may significantly change during laser irradiation (Chapter 9). Also, the thermal properties of tissue are altered somewhat when heated.

Tissue is a complicated medium, and many of the optical-thermal events produced by laser radiation are interdependent. Yet we have found that many of the assumptions used throughout this book to make formulations tractable, produce solutions that reasonably describe the behavior of laser irradiated tissue. Obviously, the mathematical formulations in Fig. 1.1 are not intended to model tissue at the microscopic level. At best these equations represent the macroscopic response of rather idealized tissue. Nevertheless, the governing equations in Fig. 1.1 provide the basis for successful models used for photodynamic therapy dosimetry, photoacoustic imaging and coagulation of enlarged vessels in the treatment of port wine stains.

1.2 Notation

Often, notation in any field may be ambiguous. When standard notations from two fields such as optics and heat transfer are combined, there are several duplications of symbols. In particular, the field of tissue optics suffers from a proliferation of symbols for the same parameters. Even though symbols are defined in each chapter, the list in Table 1.1 is generally followed throughout the book. Most of the optics notation is taken from the ISO Standard on Quantities and Units of Light and Related Electromagnetic Radiation [3]. When this standard is silent, we have tried to select symbols that have appeared in tissue optics literature and do not conflict with standard optical and thermal notation. In a few chapters it has been necessary to assign more than one meaning to a symbol but we hope the meaning is clear in the context of usage. More complete definitions of optical terms are given in Chapter 3, and thermal terms in Chapter 10. The primary source for our definitions is the special report of the International Non-Ionizing Radiation Committee of the International Radiation Protection Association [4].

Table 1.1 Nomenclature

Symbol	Description	Definition	Units
Basic parameters			
A	Area	–	m^2
c	Speed of light	300,000,000	m/s
c_t	Speed of light in tissue	c/n	m/s
k	Wave number		1/m
$\hat{\boldsymbol{n}}, \hat{\boldsymbol{s}}, \hat{\boldsymbol{u}}, \hat{\boldsymbol{v}}$	Unit vectors		
$\boldsymbol{r}$	Position	(x,y,z)	m
t	Time	–	s
v	Velocity	–	m/s
v_s	Speed of sound	–	m/s
V	Volume	–	m^3
υ	Frequency	cycles per second	1/s
λ	Wavelength	–	nm
ω, Ω	Solid angle		sr
Basic optical parameters			
$C_x,\ C_y,\ C_z$	Directional cosines		
g	Average (expected) cosine of angle of scattering	$g = \frac{\int_{4\pi} p(\hat{\boldsymbol{s}},\hat{\boldsymbol{s}}')(\hat{\boldsymbol{s}} \bullet \hat{\boldsymbol{s}}')\, d\omega'}{\int_{4\pi} p(\hat{\boldsymbol{s}},\hat{\boldsymbol{s}}')\, d\omega'}$	–
n	Refractive index	–	–
$p(\hat{\boldsymbol{s}},\ \hat{\boldsymbol{s}}')$	Phase function of single scattering	$\int_{4\pi} p\left(\hat{\boldsymbol{s}},\ \hat{\boldsymbol{s}}'\right)\, d\omega' = 1$	1/sr
r	Surface reflection	r_{ce} for collimated, external; r_{di} for diffuse, internal, etc.	–
R	Remittance, backscattering	subscripts as necessary	–
T	Transmission	subscripts as necessary	
T_c	Collimated transmission	–	–
T_d	Diffuse transmission	–	–
δ	Penetration depth of collimated light	$\delta = \frac{1}{\mu_t}$ (mean free path for attenuation event)	m
μ_a	Absorption coefficient	probability of absorption per infinitesimal path length Δx is $\mu_a \Delta_x$. ($1/\mu_a$ is mean free path length for absorption event)	1/m
μ_s	Scattering coefficient	probability of scattering per infinitesimal path length Δx is $\mu_a \Delta_x$. ($1/\mu_s$ is mean free path length for absorption event)	1/m

Table 1.1 (continued)

Symbol	Description	Definition	Units
μ_t	Total attenuation coefficient	probability of an interaction per infinitesimal path length Δx is $\mu_t \Delta x,\ \mu_t = \mu_a + \mu_S$	1/m
Dimensionless parameters			
a	Single particle albedo	$a = \frac{\mu_s}{(\mu_a + \mu_s)}$	–
τ	Optical depth	$\tau = (\mu_s + \mu_a)\ d$	–
Reduced optical properties			
a'	Reduced albedo	$a' = \frac{\mu_s'}{(\mu_s' + \mu_a)}$	–
δ_{eff}	Effective penetration depth	$\delta_{eff} = \frac{1}{\mu_{eff}}$	m
μ_{eff}	Effective attenuation coefficient	$\mu_{eff} = \sqrt{3\mu_a(\mu_a + \mu_s')}$	1/m
μ_s'	Reduced scattering coefficient	$\mu_s' = \mu_s(1-g)$	1/m
μ_t'	Reduced total attenuation coefficient	$\mu_t' = \mu_a + \mu_s'$	1/m
μ_{tr}	Transport attenuation coefficient	$\mu_{tr} = \mu_a + \mu_s' = \mu_t'$	1/m
τ'	Reduced optical depth	$\tau = \left(\mu'_s + \mu_a\right) d$	–
Electromagnetic parameters			
$E(\boldsymbol{r}_1,\boldsymbol{r}_2)$	Irradiance at point $\boldsymbol{r}_1, \boldsymbol{r}_2$	$\int L\cos\theta d\omega$	W/m^2
$L\left(\boldsymbol{r}_1, \boldsymbol{r}_2; \hat{\boldsymbol{s}}, \hat{\boldsymbol{s}}'\right)$	Radiance	–	W/m^2/***sr***
$I\left(\hat{\boldsymbol{s}}, \hat{\boldsymbol{s}}'\right)$	Radiant intensity	–	W/***sr***
$U(\boldsymbol{r},t)$	Electric field at the vector position $\boldsymbol{r}$ and time t	–	V/m
$\Gamma\ (r_1, r_2; t_1, t_2)$	Average mutual coherence function between space-time positions ($\boldsymbol{r}_1$,t_1) and ($\boldsymbol{r}_2$,t_2)	$U\ (\boldsymbol{r}_1, t_1)\ U\ (\boldsymbol{r}_2, t_2)$	W/m^2
Light parameters			
$\boldsymbol{E}_0$,$\boldsymbol{E}$	Irradiance at surface	–	W/m^2
$F_+(\boldsymbol{r},\ \hat{\boldsymbol{s}}_0)$	Flux in positive direction $\hat{\boldsymbol{s}}_0$	$F_+(\boldsymbol{r},\ \hat{\boldsymbol{s}}_0) = \int\limits_{(2\pi)+} L(\boldsymbol{r},\ \hat{\boldsymbol{s}})(\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{s}}_0)d\omega$	W/m^2
$F_-(\boldsymbol{r},\ \hat{\boldsymbol{s}}_0)$	Flux in negative direction $\hat{\boldsymbol{s}}_0$	$F_-\left(\boldsymbol{r}, \hat{\boldsymbol{s}}_0\right) = \int\limits_{(2\pi)-} L\left(\boldsymbol{r}, \hat{\boldsymbol{s}}\right)\left(\boldsymbol{s} \cdot \hat{\boldsymbol{s}}_0\right) d\omega$	W/m^2
$\boldsymbol{F}(\boldsymbol{r})$	Flux vector	$\boldsymbol{F}(\boldsymbol{r}) = \int\limits_{4\pi} L\left(\boldsymbol{r},\ \hat{\boldsymbol{s}}\right) \hat{\boldsymbol{s}} d\omega$	W/m^2
H	Radiant exposure	$H = \int E dt$	J/m^2

Table 1.1 (continued)

Symbol	Description	Definition	Units
I	Radiant intensity	$I = \int L dA$	W/sr
$L(\boldsymbol{r}, \hat{\boldsymbol{s}})$	Radiance	$L = \frac{\partial^2 P}{\partial A \partial \omega \cos\theta}$	W/sr·m^2
M	Radiant emittance	$M = \frac{dP}{dA}$	W/m^2
N	Number of photons	–	–
P	Radiant power	$P = \frac{dQ}{dt}$	W
Q	Radiant energy	$Q = h\nu N$	J
W	Radiant energy density	$W = \frac{dQ}{dV}$	J/m^3
ψ	Fluence	$\int \phi dt$	J/m^2
$\phi(\boldsymbol{r}), \phi_s(\boldsymbol{r}), \phi_d(\boldsymbol{r})$	Fluence rate of scattered light	$\phi = \int_{4\pi} L(\boldsymbol{r}, \hat{\mathbf{s}})\, d\omega$	W/m^2
ϕ_t	Total fluence rate		W/m^2
$\phi_p(\boldsymbol{r})$	Fluence rate of primary beam		W/m^2
Thermal and acoustic parameters			
A	Frequency factor	see Chapter 16	1/s
c	Specific heat	see Chapter 10	J/kg·°C
$E_b(T)_{qb}$	Total emissive power of a black body	see Chapters 10, 14	W/m^2
$E_0,\ \Delta E$	Activation energy	see Chapter 16	J/mole
$E(T)\, q_{gray}$	Total emissive power of a gray body	see Chapters 10, 14	W/m^2
h	Heat transfer coefficient	see Chapter 10	W/m^2·°C
k	Thermal conductivity	see Chapter 10	W/m .°C
L_v	Latent heat of vaporization	2.25×10^6 (water at $\rho = 10^3$ kg/m^3; $T =$ 373.1°K $p =$ 1.0133 bars)	J/kg
Q	Thermal energy	–	J
p	Pressure	–	P$_a$, bars, N/m^2
$\dot{q}$	Heat flux	$\dot{q} = -k\frac{\partial T}{\partial n}$	W/ m^2
R	Universal gas constant	$R = 8.32$	J/mole·°C
S	Rate of heat generation	$S = \mu_a \phi$	W/m^3
T	Temperature		°C
w	Perfusion rate		1/s
$W_b(\lambda, T)$	Monochromatic emissive power of a black body		W/m^3
$W(\lambda, T)$	Monochromatic emissive power of a gray body		W/m^3
α	Thermal diffusivity	$\alpha = \frac{k}{\rho c}$	m^2/s
$\in$	Emissivity of a gray body	($\in < 1$)	–
ρ	Density		kg/m^3
Ω	Damage integral	$\Omega(\tau) = A\int_0^\tau \exp\left(-\frac{\Delta E}{RT(\tau)}\right) dt$	–

Table 1.1 (continued)

Symbol	Description	Definition	Units
κ	Isothermal compressibility	($\sim 5 \times 10^{-10}\, Pa^{-1}$) for water or soft tissue	Pa^{-1}
β	Thermal coefficient of volume expansion	($\sim 4 \times 10^{-4} K^{-1}$) for muscle	K^{-1}
Γ	Grueneisen Parameter		–

Unit conversion and useful constants

1 cal $= 4.185\, J = 4.185 \times 10^7$ ergs
$1\, J = 1\ \mathrm{W} \cdot s$
h Planks constant: 6.6256×10^{-34} Js
L_v (latent heat of vaporization of water): 2.25×10^6 J/kg; 2.25×10^3 J/cm^3
ρ(water): 10^3 kg/m^3 $= 1.0$ g/cm^3
ρc(water)=4.27 J/cm^3°C
σ Stefan-Boltzmann constant: 5.6697×10^{-8} W/m$^2 \cdot$ K^4
1 mmHg $= 133.322368$ Pa (N/m^2)

Solid angle vector integrals over spheres and hemispheres

Let $\hat{\boldsymbol{s}}$ be a unit vector and vectors A and B then

1. $\int_{4\pi} \hat{\boldsymbol{s}} d\omega = 0$
2. $\int_{4\pi} \hat{\boldsymbol{s}} \left(\hat{\boldsymbol{s}} \cdot A\right) d\omega = \frac{4\pi}{3} \boldsymbol{A}$
3. $\int_{4\pi} \hat{\boldsymbol{s}} \left(\hat{\boldsymbol{s}} \cdot A\right) \left(\hat{\boldsymbol{s}} \cdot B\right) d\omega = 0$
4. $\int_{4\pi} \left(\hat{\boldsymbol{s}} \cdot B\right) \left(\hat{\boldsymbol{s}} \cdot A\right) d\omega = \frac{4\pi}{3} (A \cdot B)$
5. $\int\limits_{2\pi\ \mu \geq 0} \hat{\boldsymbol{s}}\, d\omega = \pi \hat{\boldsymbol{z}}$
6. $\int\limits_{2\pi\ \mu < 0} \hat{\boldsymbol{s}}\, d\omega = \pi \left(-\hat{\boldsymbol{z}}\right)$
7. $\frac{1}{4\pi} \int\limits_{2\pi\ \mu \geq 0} \left(\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{z}}\right) d\omega = 1/4$
8. $\frac{1}{4\pi} \int\limits_{2\pi\ \mu < 0} \left(-\hat{\boldsymbol{z}} \cdot \hat{\boldsymbol{s}}\right) d\omega = 1/4$
9. $\int\limits_{2\pi\ \mu \geq 0} \left(\hat{\boldsymbol{z}} \cdot \hat{\boldsymbol{s}}\right) \hat{\boldsymbol{s}}\, d\omega = \frac{2\pi}{3} \hat{\boldsymbol{z}}$
10. $\int\limits_{2\pi\ \mu < 0} \left(-\hat{\boldsymbol{z}} \cdot \hat{\boldsymbol{s}}\right) \hat{\boldsymbol{s}}\, d\omega = \frac{2\pi}{3} \left(-\hat{\boldsymbol{z}}\right)$

1.3 Format of Book

In this book, we have assumed that tissue, or at least regions of tissue, can be represented as a homogeneous medium. In Part I (Tissue Optics), we assume this medium consists of bulk absorption and randomly distributed scattering centers that

are sufficiently far apart that interactions at one center are independent of interactions at neighboring centers. With this model, it has been possible to use the transport integro-differential equation to describe light propagation in tissue. This theory has been extended to multiple layers, and by employing numerical Monte Carlo techniques it is possible to predict fluence rates in a volume of tissue that contains other structures (such as blood vessels). It is currently unknown how a 'true' theoretical solution based upon Maxwell's equations would compare with the transport solution. Nevertheless, single scattering, Mie theory and the relation between Electromagnetic Theory and Transport Theory are discussed in Chapters 4 and 7. No matter which model is used, the results are no better than the values of optical properties used to represent tissue. Methods are presented for measuring the absorption and scattering parameters of tissue. Interstitial measurements of fluence rate using fiber optic probes do agree with predictions of fluence rate in the diffusion region.

In Part II (Thermal Interactions), the rate of heat generation and resulting temperature fields are estimated once the light fluence rate is determined. The thermal response of tissue is estimated using various solutions to the heat conduction equation. Methods for temperature measurement provide critical information for experimentally determining temperature fields in laser irradiated tissue. Other chapters in Part II describe thermal denaturation processes, pulsed laser tissue interaction, and pulsed ablation.

Part III of this book (Medical Applications) contains applications of optical and thermal tissue interactions to various medical problems. How the optical properties of tissue impact fluorescence line shapes is one of the examples discussed. Several chapters are used to describe diagnostic and therapeutic applications of lasers, i.e., opto-acoustic and molecular imaging, OCT, and forensic optics, a recent new field in biomedical optics. Therapeutic applications include laser treatment of port wine stains and light stimulation of nerves.

References

1. Ishimaru A. *Wave propagation and scattering in random media. Vol. 1: Single scattering and transport theory*. Academic, New York (1978).
2. Carslaw HS and Jaeger JC. *Conduction of heat in solids*. University Press, Oxford, 2nd edition (1959).
3. Quantities and units of light and related electromagnetic radiations. International Organization for Standards, ISO 31/6-1980 (E).
4. Review of concepts, quantities, units and terminology for non-ionizing radiation protection: A report of the International Non-Ionizing Radiation Committee of the International Radiation Protection Association. *Health Phys*., 49(6):1329–1362 (1985).

Chapter 2
Basic Interactions of Light with Tissue

Joseph T. Walsh

2.1 Introduction

At its simplest, there are but two fates for light that is incident upon a material: either the light is absorbed within the material or it is not.

The light that is not absorbed can either return back toward the source or be transmitted through and out of the material in another direction. While in the material, the light can either travel straight or it can change direction (i.e., it can be scattered). The details of light scattering are covered in detail elsewhere (Chapters 3, 4, 5, and 6).

If the light is absorbed in the material, then a number of vastly different processes can occur: the dominant ones are that the absorbed energy can drive chemical reactions, be re-emitted as light, or be converted into heat. In thermodynamic terms, the optical energy can increase the order in the system (e.g., drive a polymerizing chemical reaction), minimally affect the system, or devolve into a less ordered state (e.g., random vibrations, i.e. heat). Each of these processes is capable of driving further actions.

While the outcomes of a light-tissue interaction are limited (remission, transmission, absorption), one finds that buried in the details of each of these outcomes lies the richness and potential of light-tissue interactions. For example, absorption can cause heating, which can drive ionization, which changes the tissue optical properties, which can increase the absorption of the incident light and thus further increase the heating, leading to a cascade that can induce plasma formation, eventual pressure rises, and explosive expansion of the material.

The purpose of this chapter is to discuss in the most general terms the various light-tissue interactions that are the rich foundation for light-based treatments and diagnostics, and thus the remainder of this book.

But first, to understand the basics of light-materials interactions, a quick review of the basics of optics is appropriate.

J.T. Walsh (✉)
Biomedical Engineering Department, Northwestern University, Evanston, IL, USA
e-mail: jwalsh@northwestern.edu

A.J. Welch, M.J.C. van Gemert (eds.), *Optical-Thermal Response of Laser-Irradiated Tissue*, 2nd ed., DOI 10.1007/978-90-481-8831-4_2,

2.2 Light

Light can be described as either an electro-magnetic wave or a stream of photons. At its simplest, the wave theory of light allows us to describe those aspects of light related to the propagation of light as it is affected by mirrors (reflected), lenses (refracted), or other objects (refracted, diffracted, etc.) and for the destructive and constructive interference effects that we observe. On the other hand, the particle theory of light allows us to understand the quantum nature of light as seen when the absorption of light increases the energy in a system by an amount which is wavelength dependent not radiance-exposure dependent.

The two views are related in that the energy of a photon is proportional to the frequency of the associated wave, namely:

$$E = h\upsilon \tag{2.1}$$

where E is the photon energy, h is Planck's constant ($\sim 6.62 \times 10^{-34}$ Js $= \sim 4.14 \times 10^{-15}$ eVs), and υ is the frequency of the wave (which is related to the wavelength, λ, by the speed of light, c; specifically, $c = \upsilon\lambda$). Beyond the simplest explanations, both the wave and particle view of light come with their own set of insights about the nature of light and the interactions of light with materials. Understanding basic light-tissue interactions requires the use of both views.

One should note that the linkage between wave and particle theory of light leads to a simple observation: as the wavelength of light increases, the photon energy decreases. Thus, 500-nm light contains 2.48 eV photons, whereas 1-μm light contains 1.24 eV photons. One sees that as one traverses the spectrum from the UV-C to the far-IR, the photon energy decreases significantly (see Table 2.1). This decrease is important because high energy photons can drive chemical reactions that lower energy photons cannot. One should note that technically, 'light' refers to that spectral band that can be seen by humans, i.e. the visible, and that optical energy that includes UV or IR is referred to by the more general term, radiation. However, common usage of the term 'light' now often refers to visible radiation as well as radiation in the adjacent UV and IR spectral bands.

If one wishes a more in-depth discussion of the wave or photon nature of light from a mathematical and physics point-of-view, several excellent texts exist [1–4].

Table 2.1 The wavelength and photon energy of each spectral band

Spectral band	Wavelength range	Photon energy range (eV)
UV-C	~100–280 nm	~12.4–4.4
UV-B	~280–320 nm	~4.4–3.9
UV-A	~320–400 nm	~3.9–3.1
Visible	~400–800 nm	~3.1–1.55
Near-IR	~800–1200 nm	~1.55–1.03
Mid-IR	~1.2–7 μm	~1.03–0.18
Far-IR	~7–1000 μm	~0.18–0.0012

2.3 Characteristics of the Incident Light and the Irradiated Material

The interaction of light with any material is dependent upon the properties of the incident light and the optical properties of the material. These properties control the propagation of light within, back from, and through the material as well as any absorption of the light. When using light for treatment or diagnosis, it is generally felt that one can understand but only minimally control the tissue optical properties. On the other hand, it is generally felt that the characteristics of the incident light are selectable and thus are the means by which significant advances in bio-optics have been and will still be made. As will be noted and as is often the case with generalities, there are exceptions to these generalities.

The major optical parameters of the incident light include the following:

- wavelength of the incident light, or equivalently the photon energy
- power or energy of the incident light
- spot size, which can either describe the area irradiated or the diameter (or radius) of the irradiated area
- irradiance or radiant exposure, where the former is power per unit area and the latter is the energy per unit area (naturally, this parameter is derived from the ratio of the previous two parameters)
- duration of the irradiation (often called the pulse duration)
- spatial profile, which is related to the spot size; the spatial profile describes how the irradiance (or radiant exposure) varies across the beam
- temporal profile, which is related to the pulse duration; the temporal profile describes how the irradiance (or radiant exposure) varies with time during the pulse
- spectral profile, which is related to the wavelength; the spectral profile describes the variation in irradiance (or radiant exposure) as a function of wavelength
- polarization state of the light: the Stokes vector, which can be used to describe the polarization of the light, includes I, which quantifies the intensity, Q and U, which quantify the degree of linear polarization, and V, which quantifies the degree of circular polarization (see Chapters 4 and 7).

There are several parameters that are used to describe how a material affects the propagation of light within that material:

- the absorption coefficient describes the absorption of light within the material (Chapter 3)
- the scattering coefficient describes the scatter of the light within the material (Chapter 3)
- the anisotropy describes the variation in direction in which the light propagates following a scattering event (Chapter 3)
- the Mueller matrix describes the transformation of the polarization properties of the incident light (i.e., the incident Stokes vector) into the polarization properties of the transmitted or remitted light (i.e., the resultant Stokes vector) (Chapters 4 and 7).

One should note that each of these material optical properties is generally wavelength dependent. Some of the parameters can be dependent on the irradiance or change over time (e.g., an absorber can bleach at high irradiance). Some parameters are temperature or pressure dependent. And other parameters vary with the polarization of the light. Such dependencies are often the source or bane of diagnostic and therapeutic opportunities.

Each of the above optical and tissue parameters will be explored in more detail below and in subsequent chapters.

2.3.1 Remission

Remission is the propagation of light back from the tissue surface generally toward the source. Remission occurs either because the light is reflected from the front surface interface or the light is scattered within the tissue and then passes through the interface, leaving the tissue and propagating toward the source.

Surface reflections (see Fig. 2.1) are governed by the index of refraction difference between the initial medium and the tissue. In its simplest form, if the initial medium has an index of refraction, $n_i = 1.0$ (e.g. approximately that of air) and the tissue has an index of refraction, $n_t = 1.33$ (e.g. approximately that of water), and the angle of incidence is 0°, then the amplitude of light reflected back toward the light source is a fraction, r, of the incident light's amplitude where

$$r = \frac{(n_t - n_i)}{(n_t + n_i)} = 0.14 \tag{2.2}$$

Given that the power in an optical beam is proportional to the square of the intensity, one sees that the reflectance, R, is r^2 and thus ~2% of the light is reflected at the air-water interface back toward the source. Note that for light that is normally incident on the surface, the propagation of the reflected light is directly back at the source.

If the light hits the interface non-normal, then the light leaves the surface at an angle from the normal identical to the angle at which it arrives (i.e., the angle of incidence equals the angle of reflection) and the amplitude of light reflected is

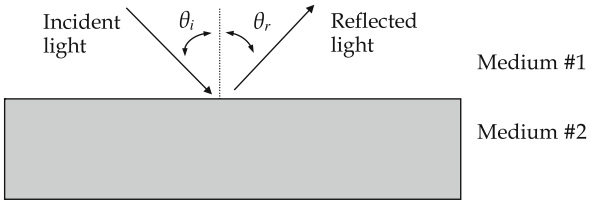


Fig. 2.1 Light incident on an interface between two mediums of differing index of refraction reflects off that interface. Note that the angle of incidence is the same as the angle of reflection ($\theta_i = \theta_r$)

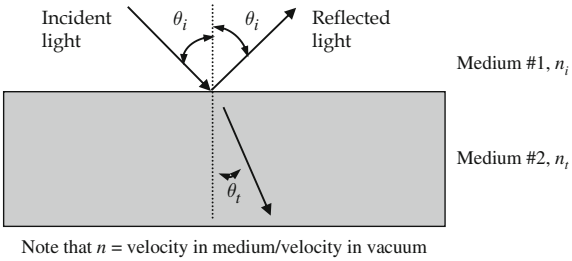


Fig. 2.2 Light incident on an interface between two mediums of differing index of refraction refracts at the interface. Note that the angle of incidence is related to the angle of refraction by the index of refraction ($n_i \sin \theta_i = n_t \sin \theta_t$)

polarization dependent. The fraction of the parallel polarization component, $r_{||}$, of the incident amplitude is

$$r_{\parallel} = \frac{n_t \cos(\theta_i) - n_i \cos(\theta_t)}{n_i \cos(\theta_t) + n_t \cos(\theta_i)} \tag{2.3}$$

and the fraction of the perpendicular polarization component, $r_{\perp}$, of the incident amplitude is

$$r_{\perp} = \frac{n_i \cos(\theta_i) - n_t \cos(\theta_t)}{n_i \cos(\theta_i) + n_t \cos(\theta_t)} \tag{2.4}$$

where n_i is the index of refraction in the incident medium, n_t is the index of refraction of the material into which the light is transmitted, θ_i is the angle at which the incident light hits the surface, and θ_t is the angle at which the light transmitted through the interface propagates.

Note that as light passes from a medium of one index of refraction into another, the direction of propagation changes (see Fig. 2.2). This process is termed refraction and is governed mathematically by Snell's Law:

$$n_i \sin(\theta_i) = n_t \sin(\theta_t) \tag{2.5}$$

One sees that as a light wave propagates from a medium of low refractive index (such as air) to one of higher refractive index (such as tissue), the light bends toward the normal. Also note that if light is propagating from a medium of higher refractive index (e.g. tissue) to one of lower refractive index (e.g. air), then for angles shallower than the critical angle (e.g. $n_i = 1.0$ and $n_t = 1.33$, $\theta_{\text{critical}} \sim 49°$) all of the light reflects back into the medium of higher refractive index (see Fig. 2.3). This critical angle is called Brewster's Angle.

An analysis of Eqs. (2.3) and (2.4), which are a sub-set of the Fresnel Equations, indicates that for a given n_i and n_t there is a θ_i at which $r_{\parallel} = 0$ and there is only transmitted light; this angle is called the polarization angle. Further, consistent with Snell's Law, when $n_i > n_t$ (e.g. when going from tissue to air, as depicted in Fig. 2.3), there is a critical angle, θ_c, independent of polarization; for $\theta_i \geq \theta_c$ all the light is reflected and none of the light is internally transmitted through the

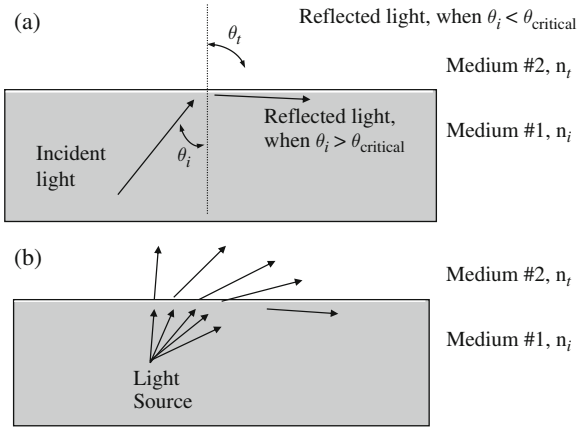


Fig. 2.3 (**a**) Light incident from within medium #1 on an interface between two mediums of differing index of refraction, where $n_i > n_t$, refracts at the interface. Note that the angle of incidence is related to the angle of refraction by Snell's Law ($n_2 \sin\theta_i = n_i \sin\theta_t$). When the angle of incidence is shallower, i.e. further from the normal than the critical angle, then the reflected light stays within the incident medium; this process is called total internal reflection. When $\theta_i = \theta_{\text{critical}}$, then the reflected wave travels along the surface as shown in (**a**). (**b**) When $\theta_i < \theta_c$, not all of the light escapes medium; such light cannot be collected by a detector

interface. This issue is important to remember when collecting light from a tissue; that is, light cannot pass out of the tissue into air and then to a detector if it strikes the tissue-air interface at an angle shallower than the critical angle.

The angle at which the light incident on a tissue surface and reflected at the material's surface leaves that interface is equal to the angle of incidence, as shown in Fig. 2.1. Note that when the surface is optically flat, then the process is called specular reflection: all the reflected light leaves the surface at the same angle and in the same direction; when the surface is rough (e.g., as with the dry surface of skin), then the process is called diffuse reflection: the light leaves the surface in various directions (see Fig. 2.4).

While the above equations are true for all light-tissue interactions, the reality is that most tissue surfaces are not optically flat; thus the angle of incidence varies

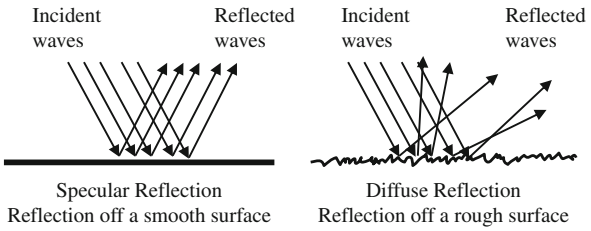


Fig. 2.4 Specular reflection occurs when light reflects off an optically smooth surface; diffuse reflection occurs when light reflects off an optically rough surface

across the tissue surface. As a result, the intensity of the reflected light varies across the surface, and the direction of propagation of the reflected and refracted light varies. Further, the index of refraction within tissue is not a constant – the index of refraction in the cell membrane is different from that in the extracellular fluid. Thus, even within the medium there are reflections and refractions (i.e., the light scatters).

While the light reflected at the air-tissue interface contains information about the index of refraction of the tissue, generally, the more information-rich remitted light is that which has propagated within the tissue and through one or more scattered processes propagates back toward an optical detector placed either in a position to collect remitted or transmitted light. To better understand such remitted light, we need briefly to discuss scattering.

2.3.2 Scattering

Scattering is the process by which light interacts with a material and the direction of propagation changes. Generally, the wavelength of light does not change during a scattering process. Such scattering is termed elastic. That very small fraction that does change wavelength is called inelastic scattering and will be discussed later.

The fundamentals of scattering can be understood in two complementary ways. First, for spatially small materials, one can consider the motions of charged particles; for large scattering centers, one can consider fluctuations in index of refraction.

Charged particles in condensed matter, such as tissue, move in constrained ways. The simplest model of a charged particle is a mass attached to the surrounding material via a spring. Thus, movement of the charged particle is governed by the equations of motion for a mass-spring system (see Fig. 2.5). As with any simple second-order system, this system has a natural frequency. Coupling energy into that system is efficient when the driving force is at the natural, or resonant, frequency. That is, if the frequency of the electromagnetic wave equals that of the resonant frequency, then energy is coupled into the particle – i.e. energy is absorbed. Driving forces off resonance affect the particle, but the transfer of energy into the system

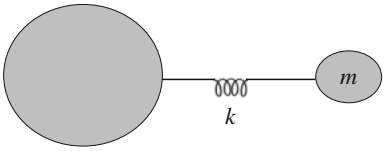


Fig. 2.5 The absorption of an incident light wave by a charged particle can be modeled using a harmonic oscillator model, where the incident electromagnetic wave couples into the mass-spring system. The equation of motion of the system is given by $d^2x/dt^2 = -kx/m$ where x is the length of the spring with spring constant, k, connecting the moving mass, m, and the solid boundary. The mass moves with a natural frequency of $f = (1/2\pi)\sqrt{(k/m)}$

is poor; that is, if the frequency is off resonance, then the amplitude of the forced oscillation of the charged particle is significantly diminished. We note, however, that an oscillating dipole acts like an antenna, which radiates the energy outward – i.e. the incident light wave induces vibrations that cause energy radiation outward in directions not necessarily in the same direction as the incident light. That is, the light changes direction – it is scattered.

In tissue, one can also consider scattering in tissue as occurring where there are fluctuations in the index of refraction within the tissue. These fluctuations can occur on spatial scales much smaller than the wavelength of light, or much larger than the wavelength of light. When the spatial fluctuations are smaller than the wavelength of light (i.e. $r \ll \lambda$; where r is the characteristic size of the particle and λ is the wavelength of the incident light), then the scattering is termed Rayleigh scattering. When the particle size is on the order of, or larger than, the wavelength of light, then the scattering is termed Mie scattering. (Note that Mie scattering theory also accounts for scattering by small particles, since Mie theory reduces to Rayleigh for $r \ll \lambda$).

For a single, spherical, small particle with incident unpolarized light one finds that the fractional intensity, I/I_0 of the light scattered is a function of the scattering angle, θ_s, the wavelength of light, λ, the distance from the particle to the light detector, R, the index of refraction of the particle, n, and the diameter of the particle, d:

$$\frac{I}{I_0} = \frac{1 + \cos^2\theta_s}{2R^2}\left(\frac{2\pi}{\lambda}\right)^4\left(\frac{d}{2}\right)^6\left(\frac{n^2-1}{n^2+2}\right)^2 \tag{2.6}$$

One of the key items to note is that the intensity decreases with the fourth power of the wavelength of light. Thus, blue light is scattered much more efficiently than the longer wavelength red light. Rayleigh scattering by small molecules in the atmosphere is responsible for the blue sky. The molecules in the air (e.g., O_2) scatter the light in a wavelength dependent way. If there were no scattering, then sun light that travels through the atmosphere would not change direction and the sky would look black (that part of the sky away from the sun would not scatter light toward your eye). Such is the case on earth's moon, where there is no atmosphere and thus the sky looks black. Small molecules however do scatter light – blue is better scattered than red – thus, the sky looks blue and, because the light directly from the setting sun travels through more atmosphere than at high noon, the light from the setting sun looks red.

Gustav Mie, and others, derived the solution to Maxwell's equations for the scattering by a particle of any size. The solution is not compact, but does yield two regions of interest: for small particles the solution approximates the Rayleigh solution; for larger particles one finds that the light is significantly forward scattered and there is little wavelength dependence. A good, common demonstration of large particle scattering (sometimes termed Mie scattering) would be clouds and fog. Water molecules in clouds are large and generally the clouds are white in color, thus indicating that all wavelengths are scattering approximately equally.

Further, Mie scattered light is more forward scattered: the scattered light continues to travel in the same general direction and little light at a local location is backward scattered.

In tissue, one encounters particles that range from significantly smaller than the wavelength of light (Rayleigh scattering), to significantly larger than the wavelength of light (Mie scattering). The phenomena that one considers to understand the familiar optics of the atmosphere are quite similar to those used to understand tissue optics.

Consider for a moment a common optical diagnostic technique used in medicine, specifically dermatology. If one has a mole (medically termed a pigmented lesion), then a quick assessment can be made of the potential that it is cancerous: generally, lesions are more suspicious if they are asymmetrical, with blurry borders, contain more than one color, have a diameter larger than ~6 mm, and an elevated surface – one considers the *abcde* nature of the mole. For the purposes of this chapter, the color component is of interest. Pigmented lesions derive their color from melanin – a molecule produced by melanocytes that should reside in the epidermis (the outer most layer of the skin). Melanin looks black because it strongly absorbs light, but quantitatively it more strongly absorbs light in the blue and ultraviolet than in the red. Indeed, melanin's absorption monotonically decreases with increasing wavelength through the visible into the infrared. The selective advantage of having more melanin is that it absorbs incident light (particularly UV light) thus protecting underlying cells from UV-induced, carcinogenic DNA changes; the selective disadvantage of having melanin is that it absorbs the light that drives vitamin D_3 production – although the evolutionary pressures on melanin are likely more nuanced. Thus, while normally melanin and melanin-producing cells reside in the upper skin (namely, the epidermis), in dysplastic and cancerous lesions, the melanocytes have dropped into the dermis. The melanin produced within the dermis still absorbs most of the light incident on the melanin – that is, the melanin looks black – but when the melanin is in the dermis, significant epidermal scattering of the incident light can occur before that light hits the melanin. Given that the scattering of light in the epidermis is typically done by small molecules, the blue light is scattered more significantly in the epidermis and the longer-wavelength light (e.g., red light) is less scattered. Thus, when the skin is illuminated by white light, the blue light is scattered and remitted even before it reaches the absorbing melanin that now lies deeper (i.e., in the dermis); hence, the lesion has a bluish hue. [One can demonstrate the principle by putting a raisin (which is a strong absorber – i.e., black) in a bowl just below the surface of skim milk. Skim milk works best because the large fat globules have been removed and scattering by skim milk is more Rayleigh-like. If the layer of skim milk is the correct thickness, then the raisin appears bluish.]

Thus far, we have limited this general introduction to scattering to issues associated with Mie and Rayleigh scattering. For wavelengths near the size of the characteristic dimensions of the scattering particles, the light is neither isotropically scattered (Rayleigh) nor dominantly forward scattered (Mie). Indeed, there is a transition from isotropic to dominantly forward scattered as the particle size increases. Further, the wavelength dependence disappears as one goes from scattering by small

particles (where short wavelength light is better scattered, Rayleigh) to large particles (where there is little wavelength dependence, Mie). The full mathematical treatise is available and deeper treatments of scattering can be found in Chapters 3, 4, 6 and 7.

The above treatment of scattering assumes an elastic process, i.e. the wavelength of the scattered light is equal to the wavelength of the light incident on the scatterer. In a small fraction of some cases, the wavelength of the scattered light is longer than that of the incident light, i.e. some of the photon energy remains in the molecule following the scattering process and thus the scattered photon has less energy than the incident photon. In such cases of inelastic scattering (which is termed Raman scattering), the energy left in the scatterer is a function of the vibrational states of the scatterer. As such, the wavelength-shifted light is a characteristic of the scatterer, i.e. it is a unique fingerprint of sorts; and thus if the wavelength-dependent spectrum of the scattered light is collected, it can be used to identify the presence of, and quantify the concentration of, scattering molecules. While Raman scattering in bulk solution (e.g. tissue) is generally so weak that practical applications in biophotonics are difficult, Raman scattering can be enhanced (e.g., at the surface of certain metals, particularly gold and silver) to achieve useful light intensities.

2.3.3 Absorption and the Fate of Absorbed Energy

When light is absorbed in tissue, then the absorbing molecule transitions from its ground state to an excited state. The excitation can be either to a higher electronic, vibrational, or rotational state. Such transitions are depicted schematically in Fig. 2.6, a depiction termed a Jablonski diagram. Note that the upward transition is generally instantaneous, femto-seconds or less. Figure 2.6 shows an upward transition between a lower electronic state (labeled S_0) and an upper electronic state (labeled S_2). Within each electronic manifold, for all but the simplest of molecules, there are several vibrational states; and within each vibrational state there are several rotational states. The graphical separation of the states is meant as a depiction of differing energy in each state – higher states have higher energy. Consider as an example, a hemoglobin molecule. Incident light could excite electrons from a ground state to a higher state, that is, one could induce an electronic transition with light. Such light is generally visible or UV light (i.e., light for which the photon energy is relatively high). Alternatively, if the photon energy is lower (e.g., mid-IR) then the absorbed energy might induce a vibrational energy transition. Consider another example, a simple three-atom H_2O molecule. In the vapor phase, absorption might mean the hydrogen-oxygen bonds transition from vibrating asymmetrically to vibrating symmetrically (see Fig. 2.7). If the energy of the absorbed photon is even lower (e.g. far-IR or microwave), then the rotation of the H_2O molecule might transition without changes in either the vibrational or electronic states of the molecule.

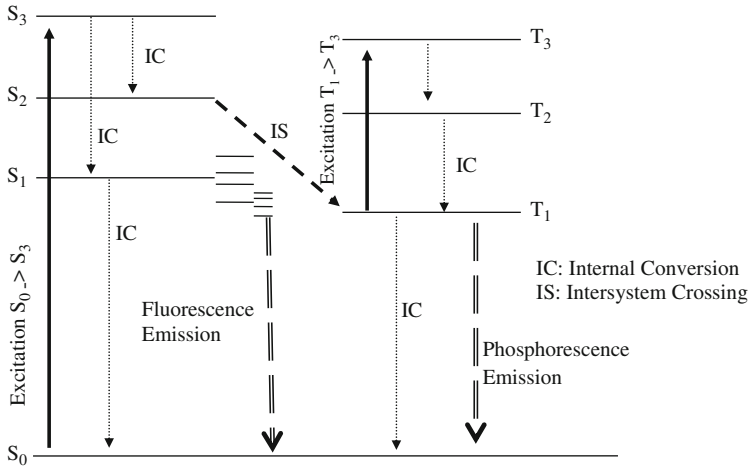


Fig. 2.6 Jablonski diagram showing that absorption of light leads to excitation from S_0 to S_3 after which the absorbed energy can return to the ground state via internal conversions (with or without emission of light) or via an intersystem crossing to the triplet state. Note that an absorption of another photon can lead to excitation in the triplet state. Note also that the fluorescence emission is from the lowest rotational level within the lowest vibrational level of the S_1 state

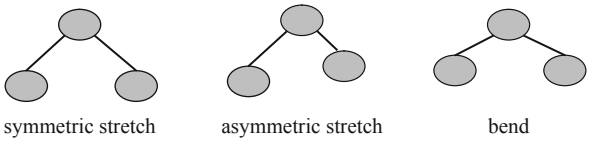


Fig. 2.7 A schematic representation of three vibrational modes for water. Depicted are the symmetric stretch (typically labeled υ_1, which for liquid water is at 3280 cm^{-1} or $\lambda \approx 3.05$ μm), the asymmetric stretch typically labeled υ_3, which for liquid water is at 3490 cm^{-1} or $\lambda \approx 2.87$ μm), and the bending mode (typically labeled υ_2 , which for liquid water is at 1644 cm^{-1} or $\lambda \approx 6.08$ μm)

As noted in Fig. 2.6, an electronic transition might result in a molecule landing in a higher rotational or vibrational state. Transitions within electronic manifolds are generally quite rapid; that is, the rotational and vibrational decays occur quickly and the molecule rapidly transitions to the lowest vibrational and rotational level within an electronic state.

One should note that an increase in the vibrational and rotational level of a molecule manifest itself in an increase in the temperature of absorbing material. It should come as no surprise that a very common result of light absorption is *heating*. The heating of a material with light is thus dependent upon absorption, with the total temperature rise directly proportional to the total energy absorbed. Mathematically, for a short pulse of light, the temperature rise is given by:

$$\Delta T = \frac{\mu_a H}{\rho c} \tag{2.7}$$

where μ_a is the absorption coefficient (typical units, cm^{-1}), H is the radiant exposure (typical units, J cm^{-2}), ρ is the density (gm cm^{-3}), and c is the heat capacity (J $°C^{-1}$ gm^{-1}). For tissue, a reasonable estimate is that $pc \approx 4.2$ J cm^{-3} $°C^{-1}$; thus if a 10 J cm^{-2} pulse of light hits a material with an absorption coefficient of 50 cm^{-1}, then the temperature rise will be ~119°C. This calculation, of course, assumes that the absorbed optical energy does not leave the immediate area of the absorption during the optical pulse (i.e. there is no diffusion during the pulse). Such is the case if the duration of the pulse is shorter than the thermal relaxation time ($\tau_p \approx r^2/4\alpha$, where r is the characteristic dimension of the absorbing target, and α is the thermal diffusivity of the material (for tissue a reasonable estimate is $\alpha \approx 1.3 \times 10^{-3}$ cm^2 s^{-1}). Note that smaller targets cool more quickly than larger ones thus shorter pulses are needed for smaller targets). One sees that to selectively heat a particular absorbing target (e.g. a blood vessel) within a material, one needs to select a wavelength that will be selectively absorbed by the target, a pulse duration that will confine the absorbed energy within the target, and a radiant exposure that will produce a sufficient temperature. As discussed in Chapter 3, light scattering in some tissue such as skin for longer wavelengths (longer than blue) can cause the local fluence [J/cm^2] to be larger than the surface radiant exposure! In such cases, H in Eq. (2.7) must be replaced by fluence ψ.

Figure 2.6 shows that electrons can (typically) exist in one of two states: the singlet state and the triplet state. The spin multiplicity, calculated as $2S+1$, is used to determine the state, where S is the number of electrons multiplied by the spin quantum number. The spin quantum number is either +1/2 (when the spin is up) or −1/2 (when the spin is down). Paired electrons generally prefer to be opposite in spin (i.e. one is up and the other is down); thus the spin multiplicity is $2(+1/2 + -1/2) + 1 = 1$, i.e. a singlet. When the electrons are both up, then the spin multiplicity is $2(+1/2 + +1/2) + 1 = 3$, i.e. a triplet. Generally, molecules do not like to be in the triplet state.

As depicted in Fig. 2.6, an excited state molecule can transition back down toward the ground state via various paths. Internal conversions occur when electrons move from one level to another level within the same state. These conversions can occur with or without the emission of light. When an internal conversion occurs with the emission of light, then the process of emission is termed fluorescence. Note that fluorescence involves emission of photons at lower energy than those that were initially absorbed (with the balance of the absorbed energy generally devolving to vibrations and rotations, i.e. heat, and the ratio of the emitted energy to the absorbed energy termed the quantum yield). Even if the fluorescence occurs from S_2 to the ground state S_0, the excitation would have resulted in not only a transition to the excited electronic state (S_2) but also to an excited vibrational and/or rotational state; but given that upper-level vibrational and rotational states extremely rapidly devolve to the lowest vibrational/rotational levels, the fluorescence process will yield a photon of lower energy than the excitation photon. Certainly, conservation of energy would require the emitted photon to be no greater in energy than the excited photon. This energy shift, which also means that there is a wavelength shift toward the red, is called the Stokes shift.

When a conversion occurs from one state to another with the emission of light, that process is termed phosphorescence. Generally, fluorescence occurs promptly (within less than a few milliseconds) of the absorption, whereas phosphorescence can occur long (seconds to minutes) after the absorption. The long delay is due to the generally long lifetimes of triplet states. Phosphorescence is the basis for many common commercially available glow-in-the-dark products; fluorescence is the basis for the glow one sees when an item is put under a black light. (Black lights emit UV that you cannot see but that excites molecules that emit visible fluorescence). Fluorescence has been shown to have diagnostic potential in several settings as discussed in Chapter 20.

As one can see, another consequence of absorption is *luminescence* either as fluorescence or phosphorescence.

Following absorption, the energy invested in the molecule can drive any of a number of *chemical processes* (see Fig. 2.8). For example, the excited molecule can transfer that energy either within the molecule, in a process termed intramolecular energy transfer, or to another molecule, in a process termed intermolecular energy transfer. These reactions are integral to photodynamic therapy in which light is absorbed by one molecule with the absorbed energy eventually transferred to another molecule that can induce the desired change. For example, a porphyrin could be the absorbing molecule which transfers the energy to oxygen, which is normally in the triplet state but which, upon excitation, becomes singlet. Note that singlet oxygen is a reactive species that can, among other actions, destabilize a cell membrane leading to cell death (e.g., in a tumor). Absorbed energy can also cause simple, but potentially important, isomerizations (e.g., from a molecule arranged as AB to one arranged as BA). For example, there is light-induced isomerization of bilirubin to a form that can be removed by the liver during phototherapy of neonatal bilirubinemia. One of the more dramatic light-induced chemical reactions is dissociation. When high-power, 193-nm, ArF excimer laser radiation strikes a polymer surface (e.g., the collagenous cornea) there is dissociation of the matrix, rapid expansion of the material, and an associated pressure rise that leads to explosive ablation of the material. Finally, an absorbed photon can be sufficiently energetic

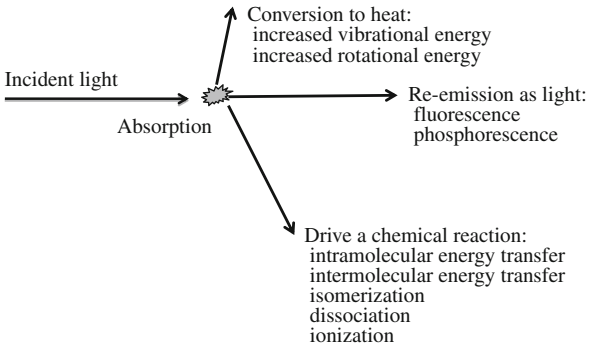


Fig. 2.8 The energy of an absorbed photon can devolve into heat (increased vibrations and rotations in molecules), result in luminescence (fluorescence and phosphorescence), or drive chemical reactions (. . .)

that the electron is excited out of the bond state of the atom, becoming a free electron. Direct ionization is one means of inducing a plasma (which is a sea of free electrons and positively charged ions).

One last point about absorption: if two photons arrive at the same time at a molecule, then there is the possibility that even though neither photon alone could induce a transition from the ground state (S_0) to an excited state (e.g., S_1), the sum of the two photon energies could match the band gap between S_0 and S_1 resulting in an absorption process. Such absorption is called two-photon absorption and is possible only when the incident beam has sufficient energy and is well focused. The advantage of two-photon absorption is that two red photons (e.g. at 800 nm) penetrate deeper in tissue than a blue photon, and when absorbed within tissue at the same location at the same time, will excite the absorber as if it saw a single blue photon (e.g., at 400 nm); then a fluorescent green photon (e.g., at 540 nm) will be emitted. Note that conservation of energy is preserved even though one gets high-energy green fluorescence from excitation from lower-energy red excitation photons.

2.4 Summary

The purpose of this chapter is to introduce the fundamental concepts that form the basis of light-tissue interactions. As was mentioned earlier, it is upon these fundamental concepts that the richness, current utility, and future potential of light-tissue interactions are built. And thus, it is upon these fundamental concepts that the following in-depth chapters of this book rest.

References

1. Born M. *Principles of optics: Electromagnetic theory of propagation, interference and diffraction of light*. University Press, Cambridge, UK, 6th edition (1997).
2. Hecht E. *Optics*. Addison-Wesley, Reading, MA, 4th edition (2002).
3. Sliney D and Wolbarsht M. *Safety with lasers and other optical sources*. Plenum, New York (1980).
4. Yariv A. *Optical electronics*. Holt Rinehart and Winston, New York (1985).

Chapter 3
Definitions and Overview of Tissue Optics

Ashley J. Welch, Martin J.C. van Gemert, and Willem M. Star

3.1 Introduction to Tissue Optics

Optics for laser irradiation of tissue is best described by examining the response of a target within tissue to light. Suppose tissue (e.g. skin) has a chromophore, (e.g. a melanocyte) somewhere inside the tissue at coordinate $\boldsymbol{r}(x, y, z)$ with respect to some frame of reference (Fig. 3.1). What is the rate of heat that is generated in the chromophore when the tissue is irradiated at some wavelength with constant power P over the laser beam radius w_L? *The key question that needs to be answered is: how many photons per second will reach the chromophore and be absorbed.* Tissue optics should provide the answer to this question.

The characteristics of photon propagation, which include scattering and absorption events within tissue, and reflection and transmission at boundaries, govern the number of photons that will reach the melanocyte at coordinate $\boldsymbol{r}$. For example, consider the paths of two photons that reach the target chromophore as illustrated in Fig. 3.2. The changes in direction of propagation shown in the figure suggest that

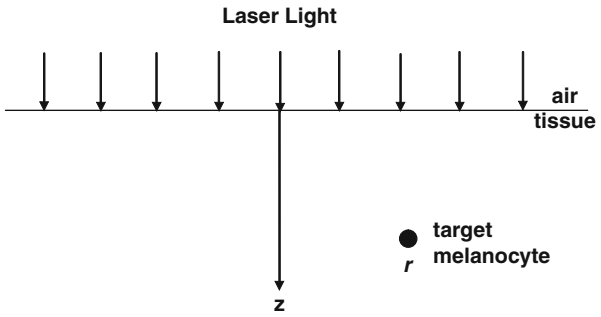


Fig. 3.1 Target melanocyte in skin located at coordinate $\boldsymbol{r}$

A.J. Welch (✉)
Department of Biomedical Engineering, The University of Texas, Austin, TX, USA
e-mail: welch@mail.utexas.edu

A.J. Welch, M.J.C. van Gemert (eds.), *Optical-Thermal Response of Laser-Irradiated Tissue*, 2nd ed., DOI 10.1007/978-90-481-8831-4_3,

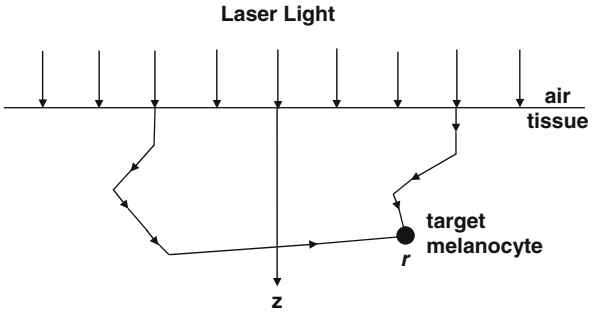


Fig. 3.2 Random paths of two photons from the laser beam that are absorbed by the target melanocyte at ***r***

quite a large number of scattering events take place for each photon. At visible and near IR wavelengths (400–1200 nm) light scattering is an important phenomenon in tissue, and the average distance a photon travels between two scattering events is about 0.01–0.2 mm. As suggested in Fig. 3.2, a scattering event changes the direction of propagation of a photon more often in the "forward" direction than in the "backward" direction, and most photons never "hit" the melanocyte as indicated in Fig. 3.3. Besides scattering, photons can be absorbed by (other) tissue chromophores; they can be totally internally reflected at the tissue-air interface; and/or they can be remitted out of the tissue.

An exact assessment of light propagation in tissue would require a model that characterizes the spatial distribution and the size distribution of tissue structures, their absorbing properties, and their refractive indices. However, for real tissues, such as skin (Fig. 3.4), it is clear that the task of creating a precise representation of either a tissue phantom or as a computer simulation is formidable, if not totally impossible. In this book, therefore, *tissue will be represented as an absorbing bulk material with scatterers randomly distributed over the volume. Within a specified region the material is assumed isotropic and homogeneous, even though this is not a true representation of tissue.*

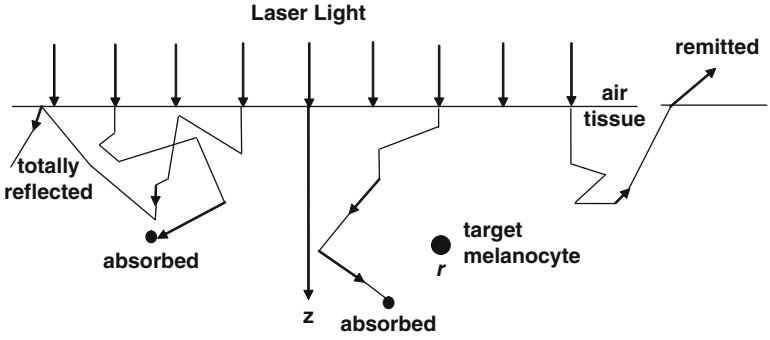


Fig. 3.3 Random paths of most photons miss target melanocyte at ***r***

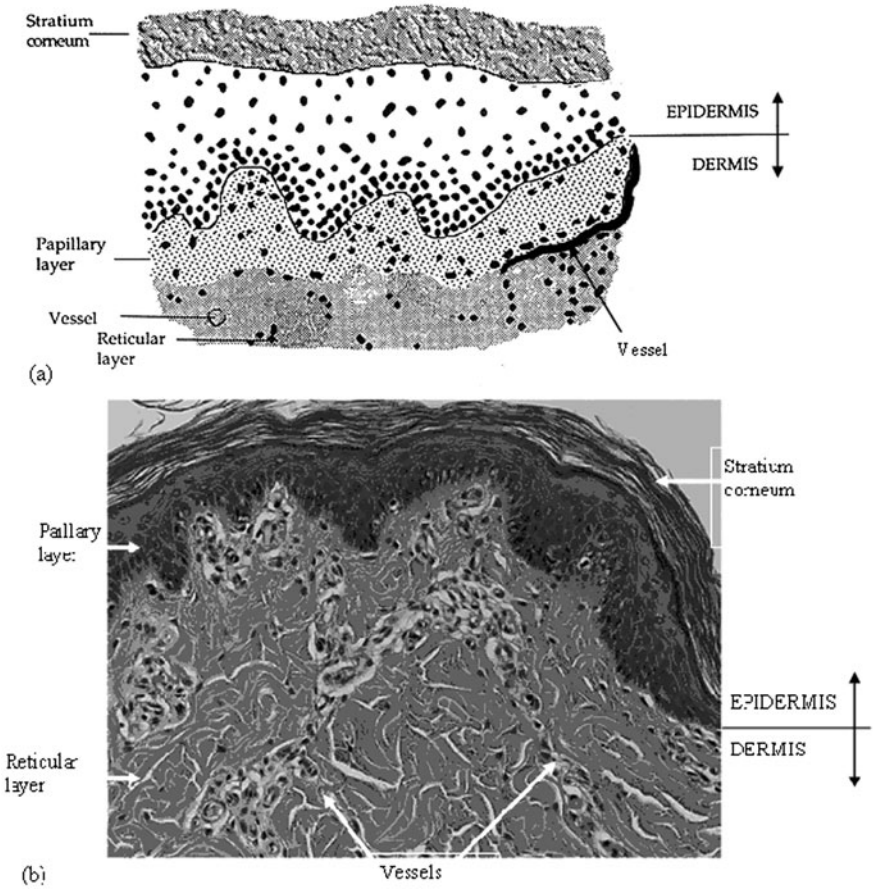


Fig. 3.4 (**a**) Drawing of section through human skin. (**b**) Histological section of pig skin

3.1.1 First Task of Tissue Optics

Finding the light energy that reaches a target chromophore per unit area per unit time at some position $\boldsymbol{r}$ is the first task of tissue optics. This rate of energy delivery per unit area at $\boldsymbol{r}$ is called the fluence rate $\phi(\boldsymbol{r})$, (see the formal definition in Eq. (3.11) below). This task is often simulated by Monte Carlo methods, where a great number of photons are injected into the tissue and their paths are followed as the photons undergo scattering and absorption. The Monte Carlo simulation describes light as particles propagating in a medium containing independent absorption and scattering centers. Monte Carlo results are consistent with transport theory, which is an analytical theory for describing light propagation in turbid media (see Section 3.4.3). The solution of the transport equation provides the spatial distribution of the radiance [W/m^2·sr] for a specified set of tissue scattering and absorption properties.

The fluence rate [W/m^2] is obtained by integrating the radiance over 4π solid angle. Unfortunately, analytical solutions of this integro-differential equation are often not available for situations of interest in laser medicine, like a finite laser beam radial profile and anisotropic scattering of tissue. Under such circumstances approximate methods sometimes provide reasonable answers. The most common method is the diffusion approximation, where light propagation is approximated as a diffusion process.

3.1.2 Second Task of Tissue Optics

The second task of tissue optics, which has so far been the most difficult one, is to develop methods by which the absorbing and scattering properties of tissue can be measured. Such properties are called the optical properties of tissue (see Section 3.2.1), and they are: (i) the absorption coefficient, (ii) the scattering coefficient, (iii) the probability density function that scattering occurs from a certain direction (with unit vector $\hat{\boldsymbol{s}}'$) into another direction (with unit vector $\hat{\boldsymbol{s}}$), (sometimes also called the phase function of single particle scattering), and (iv) the index of refraction of the tissue.

It is obvious that we wish to know these optical properties for all relevant wavelengths, and in all relevant human tissues under in vivo conditions. In other words, the ultimate goal of the second task of tissue optics is to have methods available that can assess all optical properties non-invasively in living tissues.

Unfortunately, tissue optics has not come that far. In vivo optical properties are just being measured, and even optical properties of excised tissues are not known with great accuracy. In many situations absorption and scattering undergo dynamic changes as a result of an increase in temperature, decreasing water concentration or variation in oxygenation. Heat governed denaturation typically increases scattering while a decrease in water concentration decreases scattering.

In this book, Chapters 5 (Monte Carlo modeling of light transport in tissue) and 6 (diffusion approximation) describe methods for accomplishing the first task of evaluating light propagation and fluence rate at any position $\boldsymbol{r}$. Methods for measuring the optical properties of tissue are discussed in Chapter 8.

3.2 Definitions and Geometry

In this book, vectors are denoted by boldface upper or lower case letters such as $\boldsymbol{F}$ (flux). The symbols $\hat{\boldsymbol{n}}, \hat{\boldsymbol{s}}, \hat{\boldsymbol{u}}$, and $\hat{\boldsymbol{v}}$ will be reserved for unit vectors. Integration of functions in vector notation, which occurs at a number of places, is sometimes difficult to visualize. It may then be helpful to integrate over each component separately. For example, when a function of $\hat{\boldsymbol{s}}$ is integrated over all directions in space, it is convenient to express the components of $\hat{\boldsymbol{s}}$ along the x, y, and z axes in spherical polar coordinates: $\hat{s}_x = \sin\theta\cos\varphi$, $\hat{s}_y = \sin\theta\sin\varphi$, $\hat{s}_z = \cos\theta$ (see Fig. 3.5). The element

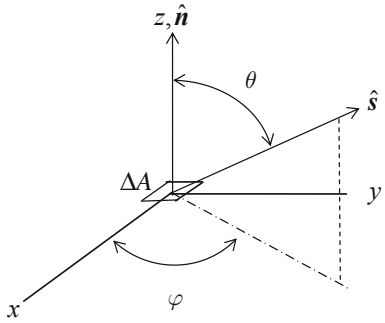


Fig. 3.5 Typical geometry

of solid angle is $d\omega = \sin\theta d\,\theta d\varphi$, where θ is the angle with the z axis and ϕ is the azimuthal angle.

Throughout this book a unit vector $\hat{\boldsymbol{s}}$ at position $\boldsymbol{r}$ is used to represent the location and direction of radiance $L\left(\boldsymbol{r}, \hat{\boldsymbol{s}}\right)$. Quite often a differential area ΔA is placed in the x, y plane so that the normal $\hat{\boldsymbol{n}}$ is aligned with the z axis as illustrated in Fig. 3.5. Associated angles are θ (from z axis to $\hat{\boldsymbol{s}}$) and φ (about the z axis).

3.2.1 Optical Properties

Photons in a turbid medium move in all directions and may be scattered (scattering coefficient $\mu_{\mathrm{s}}[\mathrm{m}^{-1}]$) or absorbed (absorption coefficient $\mu_{\mathrm{a}}[\mathrm{m}^{-1}]$). The absorption and scattering coefficients are defined such that when a photon propagates over an infinitesimal distance ds, the probability for absorption, or scattering, is respectively

probability of absorption in infinitesimal distance ds is $\mu_a ds$
probability of scattering in infinitesimal distance ds is $\mu_s ds$

The mean free path for an absorption event is $1/\mu_a$, and the mean free path for a scattering event is $1/\mu_s$. The sum of μ_a and μ_s is designated as the total attenuation coefficient, μ_t.

$$\mu_t = \mu_a + \mu_s \tag{3.1}$$

Thus the probability that a photon is absorbed in a distance less than s, denoted by the probability distribution function $P_a\,(S < s)$ where S is the random variable representing the distance for the photon to hit a randomly distributed absorption cross section and s is a value is the domain of S which is greater than zero for the above example.

$$P_a\,(S < s) \;=\; F_a\,(s) \;=\; 1 - e^{-\mu_a s} \tag{3.2}$$

The probability distribution function is noted as $F_a(s)$.

Probability distributions are used extensively in Chapter 5 in the Monte Carlo modeling of photon travel in a scattering medium. In probability theory, a distinction is made between probability distribution functions and probability density functions. The probability, for example, that a photon is absorbed in a distance less than s, described by a *probability distribution function* F_a (e.g. $F_a(s) = 1 - \exp[-\mu_a s]$), and the probability that a photon is absorbed between s and $s+ds$, divided by ds, is called the *probability density function* f_a. The integral of any density function f is equal to the distribution function F.

$$F(s) = \int_{-\infty}^{s} f(s')\, ds' \tag{3.3}$$

where s' is the dummy variable of integration in Equation (3.3).

Similar distribution functions describe a scattering event ($F_s(s)$) and either an absorption or scattering event $F_t(s)$

$$F_s(s) = 1 - e^{-\mu_s s} \tag{3.4}$$

and

$$F_t(s) = 1 - e^{-\mu_t s} \tag{3.5}$$

Additional coefficients for optical properties are introduced in the section of this chapter that describes single scattering phase functions (Section 3.4.3.2). Values for wavelength dependent optical properties for blood [1], water [2] and tissue are presented throughout this book. Techniques for measurement of these properties is presented in Chapter 8.

3.2.2 Optical Parameters

This section defines and explains some of the important optical parameters that are used to describe the propagation of light in a turbid material. Terms like the radiance, fluence rate, and flux are described [3–5] in terms of the propagation of photons.

3.2.2.1 Photon Density, $N_0(\boldsymbol{r})$

$N(\boldsymbol{r}, \hat{\boldsymbol{s}})$ at a given point $\boldsymbol{r}$ is defined as the number of photons per unit volume moving in the direction of unit vector $\hat{\boldsymbol{s}}$ within solid angle $d\omega$. The units are $\mathrm{m}^{-3}\mathrm{sr}^{-1}$. The photons density $\left[\frac{\text{photons}}{\text{m}^3}\right]$ at $\boldsymbol{r}$ is

$$N_0(\boldsymbol{r}) = \int_{4\pi} N(\boldsymbol{r}, \hat{\boldsymbol{s}})\, d\omega \tag{3.6}$$

The total photon fluence rate [W/m^2] carried by $N_0(\boldsymbol{r})$ is given by

$$\phi(\boldsymbol{r}) = N_0(\boldsymbol{r})\, h\nu c_t \tag{3.7}$$

where ϕ [W/m^2] is the fluence rate (formal definition in Section 3.2.2.3), c_t[m/s] the speed of light in tissue, $h\nu$[J] the energy per photon, h is Plank's constant, and ν is frequency.

$$c_t = \frac{c}{n_t} \tag{3.8}$$

where n_t is the index of refraction of the medium and c[m/s] is the speed of light in a vacuum.

3.2.2.2 Radiance $L(\boldsymbol{r}, \hat{\boldsymbol{s}})$

Let us now introduce the radiance $L(\boldsymbol{r}, \hat{\boldsymbol{s}})$ as

$$L(\boldsymbol{r}, \hat{\boldsymbol{s}}) = N(\boldsymbol{r}, \hat{\boldsymbol{s}})\, h\nu c_t \ [\mathrm{W}/\mathrm{m}^2 \cdot \mathrm{sr}] \tag{3.9}$$

Radiance $L(\boldsymbol{r}, \hat{\boldsymbol{s}})$ is the quantity used to describe the propagation of photon energy. It is the radiant power per unit of solid angle about unit vector $\hat{\boldsymbol{s}}$ and per unit area perpendicular to $\hat{\boldsymbol{s}}$. In words, *the radiance* is defined as [3]: "at a point of a surface and in a given direction, the radiant intensity of an element of the surface, divided by the area of the orthogonal projection of this element on a plane perpendicular to the given direction" where the radiant intensity [W/sr] is defined below in Section 3.2.2.8.

Radiance $L(\boldsymbol{r}, \hat{\boldsymbol{s}})$ is related to the power $dP(\boldsymbol{r}, \hat{\boldsymbol{s}})$ [W] flowing through infinitesimal area dA, located at $\boldsymbol{r}$, in the direction of unit vector $\hat{\boldsymbol{s}}$, within infinitesimal solid angle $d\omega$.

$$dP(\boldsymbol{r}, \hat{\boldsymbol{s}}) = L(\boldsymbol{r}, \hat{\boldsymbol{s}}) dA d\omega \tag{3.10a}$$

When direction $\hat{\boldsymbol{s}}$ is not perpendicular to area dA, the projection of dA orthogonal to $\hat{\boldsymbol{s}}$, namely $dA \cos\theta$ has to be used in Eq. (3.10a), where θ is the angle between $\hat{\boldsymbol{s}}$ and the normal $\hat{\boldsymbol{n}}$ of dA. The power flowing through dA at $\boldsymbol{r}$ in direction $\hat{\boldsymbol{s}}$, then is (Fig. 3.6)

$$dP(\boldsymbol{r}, \hat{\boldsymbol{s}}) = L(\boldsymbol{r}, \hat{\boldsymbol{s}}) dA \cos\theta\, d\omega \tag{3.10b}$$

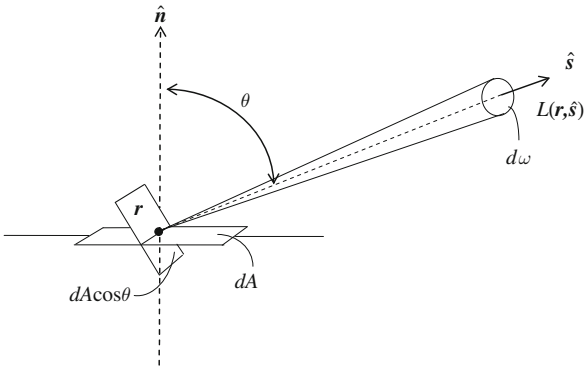


Fig. 3.6 When unit vector $\hat{\boldsymbol{s}}$ is not perpendicular to surface element dA, the projection of dA on a plane perpendicular to $\hat{\boldsymbol{s}}$ must be used, which is $dA\cos\theta$

From Eq. (3.10a), radiance $L(\boldsymbol{r},\hat{\boldsymbol{s}})$ can formally be expressed as

$$\frac{dP(\boldsymbol{r},\hat{\boldsymbol{s}})}{dA\cos\theta d\omega} = L(\boldsymbol{r},\hat{\boldsymbol{s}}) \tag{3.10c}$$

Because of Eq. (3.10c), radiance $L(\boldsymbol{r},\hat{\boldsymbol{s}})$ is sometimes defined as the second derivative of power $P(\boldsymbol{r},\hat{\boldsymbol{s}})$ with respect to $d\omega$ and $dA\cos\theta$ [5].

$$L(\boldsymbol{r},\hat{\boldsymbol{s}}) = \frac{\partial^2 P(\boldsymbol{r},\hat{\boldsymbol{s}})}{\partial\omega\ \partial A\cos\theta} \tag{3.10d}$$

It should be noted, however, that this definition can only have a symbolic meaning, because there is no functional relationship known between $P(\boldsymbol{r},\hat{\boldsymbol{s}})$ and ω and A that will yield an expression for $L(\boldsymbol{r},\hat{\boldsymbol{s}})$ by taking the partial derivative according to Eq. (3.10d).

3.2.2.3 Fluence Rate $\phi(\boldsymbol{r})$

Absorption of photons is the key to clinical medicine. Because an absorbing chromophore at location $\boldsymbol{r}$ inside the tissue can absorb photons irrespective of their direction of propagation, the integral of the radiance over all directions, called fluence rate $\phi(\boldsymbol{r})$ [W/m^2], has more practical significance than the radiance itself. By definition,

$$\phi(\boldsymbol{r}) = \int_{4\pi} L(\boldsymbol{r},\hat{\boldsymbol{s}})d\omega \tag{3.11}$$

The *radiant energy fluence rate* is formally defined as [3]: "at a given point in space, the radiant power incident on a small sphere, divided by the cross-sectional area of that sphere."

3.2.2.4 Net Flux Vector $F(r)$

Flux is a description of photon energy transport per unit area. The net flux vector $\boldsymbol{F}(\boldsymbol{r})$ [W/m^2] is the vector sum of elemental flux vectors $L(\boldsymbol{r},\hat{\boldsymbol{s}})\hat{\boldsymbol{s}}\ d\omega$; that is

$$\boldsymbol{F}(\boldsymbol{r}) = \int_{4\pi} L\left(\boldsymbol{r}, \hat{\boldsymbol{s}}\right)\hat{\boldsymbol{s}}d\omega \tag{3.12a}$$

where integration is taken over the full 4π steradians. Of physical importance is the net energy flux per unit area through an area dA, which is $\boldsymbol{F}(\boldsymbol{r}) \cdot \hat{\boldsymbol{n}}$, where $\hat{\boldsymbol{n}}$ is perpendicular to dA.

Hemispherical fluxes can be defined as the energy flux through dA in either forward direction $\hat{\boldsymbol{n}}$, or backward direction $-\hat{\boldsymbol{n}}$. The hemispherical flux $F_{n+}(\boldsymbol{r})$ is an integral over the solid angle of 2π $\left(0 < \theta \leq \pi/2,\ 0 < \varphi < 2\pi\right)$.

$$F_{n+}(\boldsymbol{r}) = \int_{2\pi} L(\boldsymbol{r},\hat{\boldsymbol{s}})\ (\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{n}})d\omega \tag{3.12b}$$

Similarly, the hemispherical flux passing through dA in the backward direction $-n(\pi/2 < \theta \leq \pi)$ denoted by $F_{n-}(\boldsymbol{r})$ is given as

$$\begin{aligned} F_{n-}(\boldsymbol{r}) &= \int_{2\pi} L(\boldsymbol{r},\hat{\boldsymbol{s}})\ (\hat{\boldsymbol{s}} \cdot -\hat{\boldsymbol{n}})d\omega \\ &= -\int_{2\pi} L(\boldsymbol{r},\hat{\boldsymbol{s}})\ (\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{n}})d\omega \end{aligned} \tag{3.12c}$$

From these forward and backward hemispherical fluxes, Eqs. (3.12b and c), and the basic definition of the net flux vector, Eq. (3.12a), the net flux vector is related to the hemispherical fluxes as

$$\boldsymbol{F}(\boldsymbol{r}) \cdot \hat{\boldsymbol{n}} = F_{n+}(\boldsymbol{r}) - F_{n-}(\boldsymbol{r}) \tag{3.12d}$$

3.2.2.5 Radiant Power (Radiant Energy Flux) and Radiant Energy

The delivery of energy is described in terms of *radiant power (radiant energy flux), P* [W], which is "power emitted, transferred, or received as radiation" [3] or *radiant energy , Q* [J], which is "energy emitted, transferred or received as radiation" [3].

3.2.2.6 Radiant Exitance (Emittance)

Light that is backscattered from tissue can be described in terms of *radiant exitance (emittance) M* [W/m^2] which is the light flux F_{n+} or F_{n-} leaving the tissue, depending upon the direction of the normal. The ISO definition is: "at a point of

a surface, the radiant energy flux leaving an element of the surface divided by the area of that element" [3].

3.2.2.7 Radiant Energy Density

The *radiant energy density, W* [J/m^3], is the "radiant energy in an element of volume, divided by that element" [3].

3.2.2.8 Radiant Intensity

Radiant intensity, I [W/sr], which is used by some authors, is defined by ISO [3] as: "in a given direction from a source, the radiant energy flux (or power) leaving the source, or an element of the source, in an element of solid angle containing the given direction, divided by that element of solid angle."

3.2.2.9 Irradiance

Most diagnostic and medical applications of lasers involve incident laser light on tissue. For cw radiation, the rate of energy delivery per second per unit area of the tissue surface is termed *irradiance* E_0 [W/m^2]. Irradiance is defined for any radiation in [3] as: "at a point of a surface, the radiant energy flux (or power) incident on an element of the surface, divided by the area of the surface." As a laser beam in direction $\hat{\boldsymbol{s}}_0$ propagates into the tissue, it is attenuated by absorption and scattering. We have not found a standard notation for the remnant of collimated light that has traveled some distance in the scattering medium. Ishimaru [4] uses the term primary fluence rate, ϕ_p [W/m^2].

Another formulation incorporates a delta function and employs the well known sampling theorem for integrals which states that for any function $f(\hat{\boldsymbol{s}})$ which is continuous in $\hat{\boldsymbol{s}}_0$

$$\int_{4\pi} f\left(\hat{\boldsymbol{s}}\right) \frac{\delta\left(1-\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{s}}_0\right)}{2\pi} d\omega = f\left(\hat{\boldsymbol{s}}_0\right) \tag{3.13a}$$

Thus, as described in Chapter 6, the representation for power per unit area of a collimated beam, incident at $\boldsymbol{r}$, in a certain direction $\hat{\boldsymbol{s}}_0$ is $E(\boldsymbol{r},\hat{\boldsymbol{s}})\,\delta(1-\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{s}}_0)/(2\pi)$, since

$$\int_{4\pi} E(\boldsymbol{r},\hat{\boldsymbol{s}}) \frac{\delta(1-\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{s}}_0)}{2\pi} d\omega = E(\boldsymbol{r},\hat{\boldsymbol{s}}_0) = \phi_p(\boldsymbol{r}) \tag{3.13b}$$

Suppose a semi-infinite medium is irradiated by uniform diffuse light. Assuming a constant radiance, L_0, and that the unit directional vector $+\hat{\boldsymbol{n}}$ is directed into the medium, the resulting flux, F_{n+} (which is the irradiance [W/m^2]), is

$$F_{n+} = \int_{2\pi} L_0 \cos\theta d\omega = L_0 \int_0^{2\pi} d\phi \int_0^{\pi/2} \cos\theta \sin\theta d\theta = \pi L_0 \tag{3.13c}$$

which would be a diffuse irradiance, E_d.

3.2.2.10 Radiant Exposure

Radiation parameters for pulsed laser systems are usually described in terms of energy delivered per pulse. The delivered energy per unit area is termed the *radiant exposure, H_0* [J/m^2]. Also, this term is used to describe the dose of light for photochemical reactions. Similar as discussed above in Eqs. (3.13a) and (3.13b) for power, the energy of the collimated beam incident per unit area at some position $\boldsymbol{r}$ in the tissue, $H(\boldsymbol{r})$ [J/m^2], is the fluence at $\boldsymbol{r}$ due to the collimated (primary) laser pulse.

3.2.2.11 Source Term $S(\boldsymbol{r})$

The symbol $S(\boldsymbol{r})$ is used to represent the "source term" in both the transport equation for light propagation and in the heat conduction equation. When used in the heat conduction equation, the amount of photon energy per second absorbed locally in tissue, in an infinitesimally small volume at $\boldsymbol{r}$, is equal to

$$S(\boldsymbol{r}) = \mu_a(\boldsymbol{r})\phi(\boldsymbol{r}) \quad [\mathrm{W/m^3}] \tag{3.14}$$

where $\mu_a(\boldsymbol{r})$ is the local absorption coefficient [m^{-1}] of the tissue or a chromophore added to the tissue at $\boldsymbol{r}$. To obtain fluence rate we either need an equation for $L(\boldsymbol{r}, \hat{\boldsymbol{s}})$, that is the transport equation (Section 3.4.3), or a theory to calculate $\phi(\boldsymbol{r})$ directly. For the latter, the diffusion equation can be used, which can be derived from general principles or as an approximation of the transport equation (Chapter 6). In Section 3.4.4 we show the amount of volumetric photon power absorbed at $\boldsymbol{r}$ is equal to the negative divergence of the flux vector $\boldsymbol{F}(\boldsymbol{r})$ (Eq. (3.55); also see Chapter 6, Eq. (6.3a) or Eq. (6.33)).

3.3 Reflection and Transmission at a Surface

The interaction of laser light with tissue is highly dependent upon the method for delivery of the light energy to a selected target. For example, if the laser beam is delivered through air to tissue as depicted in Fig. 3.7, a portion of the beam is reflected at the surface due to the difference in the index of refraction, $n_1 = 1$, of the air, and the index of refraction, $n_2 \approx 1.4$ of the tissue, according to the laws of Fresnel. The fraction of light that is directly reflected from the surface depends upon the angle of incidence of the laser beam and the two indices of refraction. Even when the laser beam is perpendicular to the tissue, a small fraction is reflected

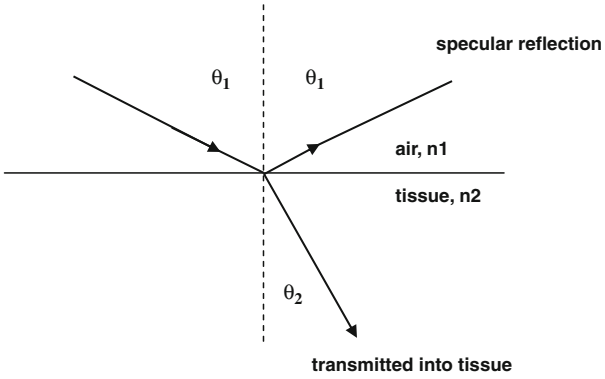


Fig. 3.7 Cartoon of a ray of light incident on tissue at angle θ_1 with respect to the normal of an air-tissue boundary. A portion of the ray is specularly reflected with angle θ_1 and the remainder is transmitted at angle θ_2

(see Eq. (3.15b)). Neglecting polarization, the angular dependence of specular reflection (R_s) at an air-tissue interface is shown in Fig. 3.8 for n_1(air) = 1.0 and n_2(tissue) which can vary from 1.33 to about 1.5. These curves are calculated using Snell's law Eq. (3.15c), and the well-known Fresnel relation for specular reflection R_s, Eq. (3.15a), [6, 7].

$$R_s\ (\theta_1) = \frac{1}{2}\left[\frac{\tan^2(\theta_1 - \theta_2)}{\tan^2(\theta_1 + \theta_2)} + \frac{\sin^2(\theta_1 - \theta_2)}{\sin^2(\theta_1 + \theta_2)}\right],\ \theta_1 \neq 0 \qquad (3.15a)$$

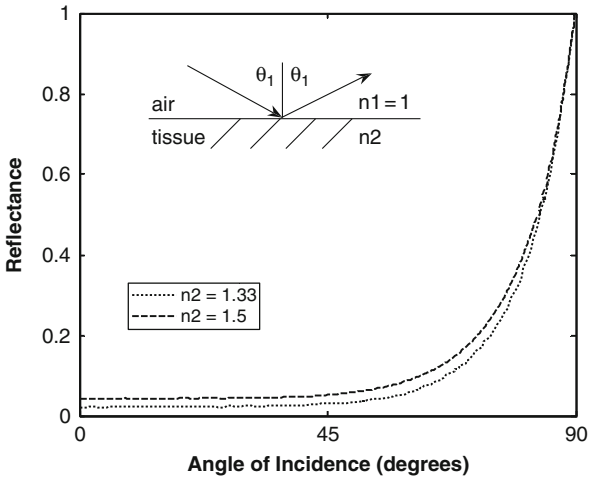


Fig. 3.8 Specular reflection of a ray of light using the Fresnel equation for non-polarized light

where θ_1 and θ_2 are measured with respect to the surface normal as illustrated in Fig. 3.7. When the beam is normal to the surface, Eq. (3.15a) reduces to

$$R_s = \frac{(n_1 - n_2)^2}{(n_1 + n_2)^2} \text{ for } \theta_1 = 0 \tag{3.15b}$$

This derivation uses Snell's law, Eq. (3.15c) below, and taking the limits as θ_1 and θ_2 approach zero. For a smooth flat surface, the reflected beam has the same angle with the surface normal as the incident beam. The angle of the transmitted beam is given by Snell's law as

$$n_1 \sin\theta_1 = n_2 \sin\theta_2 \tag{3.15c}$$

For a unit irradiance the light transmitted into the tissue is

$$T = 1 - R_s \tag{3.16}$$

Equations (3.15) and (3.16) govern the behavior of internal reflections of light in tissue and describe reflection and transmission of a laser beam at the surface of tissue as well. A detailed discussion of these boundary conditions for light propagation in tissue is given in Chapter 6. If light within the tissue strikes the surface with an angle θ_2 greater than the critical angle θ_c, the light is totally internally reflected. The critical angle θ_c is given by

$$\sin\theta_c = \frac{n_1}{n_2}\sin 90^\circ = \frac{n_1}{n_2} \quad \text{for } n_1 < n_2 \tag{3.17a}$$

or

$$\theta_c = \arcsin\left(\frac{n_1}{n_2}\right) \tag{3.17b}$$

3.3.1 Example: Reflection in Fiber Optics

Fresnel's equation governs the ray optics of light transmitted via delivery fibers. Fiber optics provide the basic element for either the delivery of a laser beam or the detection of light. A complete description of the optics of fibers is presented in Chapter 16. The purpose of this short note is simply to introduce the concept of ray entrance into a fiber, numerical aperture, and acceptance angle for a simple core-cladding fiber. Propagation of a ray through a fiber requires the cladding index of refraction, n_2, to be smaller than the core index of refraction, n_1. The limiting acceptance angle of a fiber is determined with the help of Fig. 3.9.

For light to propagate through the fiber, ray angles with respect to the normal to the core-cladding interface must be greater or equal to the critical angle α_c where

$$\alpha_c = \arcsin\left(\frac{n_2}{n_1}\right) \tag{3.18}$$

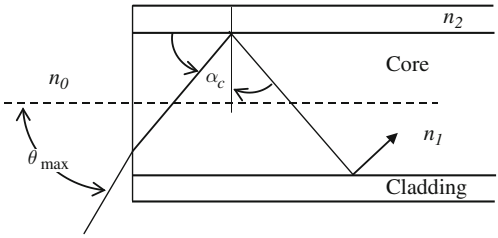


Fig. 3.9 Diagram for maximum acceptance angle θ_{max}. Index of refraction of the environment, core, and cladding are at n_0, n_1, n_2 respectively

The maximum acceptance angle with respect to the core center line, θ_{max}, requires consideration of the refraction of the ray by the environment-core interface. Using Snell's law and an environmental index of refraction n_0

$$n_0 \sin\theta_{max} = n_1 \sin(\pi/2 - \alpha_c) \tag{3.19}$$

The $\sin(\pi/2 - \alpha_c)$ can be represented as

$$\sin(\pi/2 - \alpha_c) = \frac{\sqrt{n_1^2 - n_2^2}}{n_1} \tag{3.20}$$

so

$$n_0 \sin\theta_{max} = \sqrt{n_1^2 - n_2^2} = NA \tag{3.21}$$

where *NA* is called the Numerical Aperture. Although the *NA* is a property of a fiber, the maximum acceptance angle θ_{max} is a function of the environmental index of refraction n_0.

3.3.2 Backscatter

When a laser beam is incident on tissue, the specular reflection from the surface may be a few percent higher than predicted by Eq. (3.15a) which assumes a smooth surface. This is due to the slight divergence of the laser beam and roughness of the tissue surface. Inside the tissue, scattering and absorption attenuate the collimated beam and further "de-collimate" the incident flux as photons are scattered away from the laser beam. Some of the light scattered from the collimated beam undergoes multiple reflections and propagates in the backward direction. Backscattered light that reaches the tissue surface is either internally reflected or transmitted according to Fresnel's relation given in Eq. (3.15a). Thus any measurement of reflection includes both the specular reflection above the irradiated tissue and the transmitted portion of the backscattered flux. That is

$$R_t = R_s + R_d \tag{3.22}$$

where R_t is the total reflectance; R_s is the specular reflection; and R_d is the remitted diffuse light.

3.3.3 Reflection of Diffuse Light

Diffuse light corresponds to a constant radiance for any direction $\hat{\boldsymbol{s}}$. That is

$$L(\boldsymbol{r}, \hat{\boldsymbol{s}}) = L_0 \tag{3.23}$$

If a surface is irradiated uniformly with diffuse light, the total integrated incident power per unit area, or the flux, was calculated in Eq. (3.13c) as

$$E_d = F_{n+} = \pi L_0 \, [\mathrm{W/m^2}] \tag{3.24}$$

for $\hat{\boldsymbol{n}}$ into the medium.

The reflected flux for rays of incident light between θ_1 and $\theta_1 + d\theta_1$ is

$$R_s(\theta_1) L_0 \cos\theta_1 2\pi \sin\theta_1 d\theta_1 \tag{3.25}$$

where $R_s(\theta_1)$ is the Fresnel reflection coefficient at angle θ_1. The total reflected flux [W/m^2] is obtained by integrating equation (3.25) over θ_1 from 0 to $\pi/2$.

$$2\pi L_0 \int_0^{\pi/2} R_s(\theta_1) \cos\theta_1 \sin\theta_1 d\theta_1 = \pi L_0 \int_0^{\pi/2} R_s(\theta_1) \sin(2\theta_1) d\theta_1 \tag{3.26}$$

The integrated specular reflection coefficient r_{sd} for diffuse irradiance of tissue from air is

$$r_{sd} = \frac{\text{reflected flux}}{\text{incident flux}}. \tag{3.27}$$

For uniform diffuse irradiance, Eqs. (3.24) and (3.26) yield

$$r_{sd} = \int_0^{\pi/2} R_s(\theta_1) \sin(2\theta_1) d\theta_1 \tag{3.28}$$

The above equation can be modified to compute internal reflection of light at tissue-air interfaces by including the critical angle for total internal reflection, θ_c. For light traveling to the surface at an angle $\theta_2 > \theta_c$, the Fresnel reflection

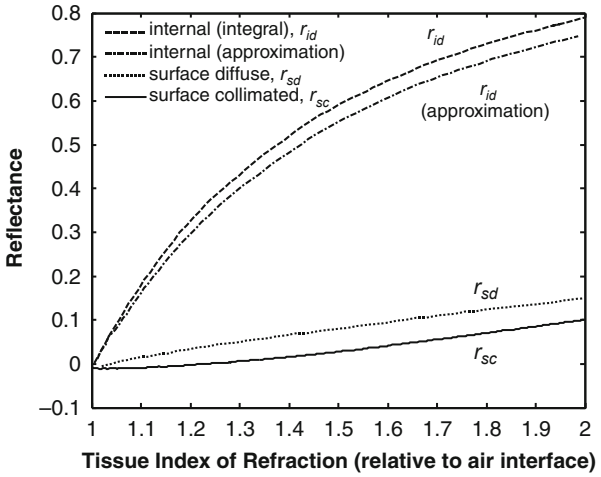


Fig. 3.10 Surface reflectance for (i) a collimated beam normal to the surface, r_{sc}, and (ii) uniform diffuse irradiance of surface, r_{sd}. (iii) Internal reflectance r_{id} of a uniform diffuse light that strikes the interface from inside the medium. r_{id} (approximation) assumes $R_s(\theta_2) = 0$ in Eq. (3.29)

coefficient equals 1. The integrated internal reflection coefficient r_{id}, assuming L is constant, is

$$\begin{aligned} r_{id} &= \frac{\text{reflected flux}}{\text{incident flux}} \\ &= \frac{2\pi \int_0^{\theta_c} R_s(\theta_2) L_0 \cos\theta_2 \sin\theta_2 d\theta_2}{\pi L_0} + \frac{2\pi \int_{\theta_c}^{\pi/2} L_0 \cos\theta_2 \sin\theta_2 d\theta_2}{\pi L_0} \\ &= \int_0^{\theta_c} R_s(\theta_2) \sin(2\theta_2) d\theta_2 + \frac{1}{2}[\cos(2\theta_c) + 1] \end{aligned} \tag{3.29}$$

(See further development of boundary conditions in Chapter 6.)

The effect of index of refraction upon the values of r_{sd} and r_{id} is plotted in Fig. 3.10. Also included in the figure is the Fresnel specular reflection coefficient r_{sc} for a collimated ray of light normal to the surface of the tissue ($\theta = 0$).

3.4 Propagation of Light into Tissue

3.4.1 Propagation of Collimated Light

Light entering tissue is subject to scattering and absorption. A collimated laser beam normal to the surface with uniform irradiance, E_0, has a small portion of the light

reflected at the surface, Eq. (3.11b), and the remaining light is attenuated in the tissue by absorption and scattering according to Beers' law

$$E(z) = E_0(1 - r_{sc})e^{-(\mu_a+\mu_s)z} \tag{3.30}$$

where E_0 is the collimated irradiance [W/m^2]; r_{sc} is the Fresnel surface (specular) reflection of collimated light striking the external surface, Eq. (3.15b); μ_a is the absorption coefficient [m^{-1}]; μ_s is the scattering coefficient [m^{-1}].

$E(z)$ is the fluence rate of collimated light at position z in the tissue [W/m^2] sometimes denoted as the primary fluence rate [4].

$$\phi_p(z) = E(z) \tag{3.31}$$

A more precise definition of the relation between primary radiance, L_p, and E is given in Eq. (3.37).

3.4.1.1 Penetration Depth, δ

The penetration depth δ of the collimated beam is defined here as the mean free path for an absorption or scattering event and it is the depth where $E(\delta) = E_0(1-r_{sc})e^{-1}$, that is, the depth where the percentage of collimated light transmitted into the tissue is reduced to 37%. Thus, the collimated penetration depth is defined as the reciprocal of the attenuation coefficient

$$\delta = \frac{1}{\mu_t} \tag{3.32}$$

3.4.1.2 Optical Depth (OD)

Quite often a distinction is made between the physical thickness, d, of tissue and its optical thickness, OD, which is defined as

$$\text{OD} = \mu_t d \tag{3.33}$$

The transmission of a collimated beam through a tissue with a thickness of one OD is 37%; at three OD the collimated beam is reduced to 5%. Penetration limits for optical diagnostics are dictated by the optical depth of a tissue.

3.4.1.3 Albedo, a

The albedo is defined as the ratio of the scattering coefficient to the attenuation coefficient

$$a = \frac{\mu_s}{\mu_t} \tag{3.34}$$

3.4.1.4 Attenuation of Collimated Light

It is important throughout this book to remember that no matter how complicated the relations become for scattered light, the fluence rate distribution of collimated light with tissue depth is governed by Beer's law, (Eq. (3.30)). For tissue samples that do not scatter light, the absorption coefficient can be determined from the slope of a plot of ℓnT_c (where Tc is the collimated transmission) versus sample thickness, z.

$$\ell n\left[\frac{E(z)}{E_0}\right] = \ell nT_c = \ell n(1 - r_{sc}) - \mu_a z \quad \text{for } \mu_s = 0 \tag{3.35}$$

In theory, when scattering is important, the slope of ℓnT_c versus z provides the attenuation coefficient, μ_t. Unfortunately, it is very difficult to measure the transmission of only collimated light. Invariably, measurements of T_c include scattered light.

Tissue components that absorb light are called *chromophores*. Some of the most important chromophores for visible wavelengths are blood and melanin; the latter are small ($\sim$1 μm) pigment granules in the skin and eye. The absorption coefficient of these granules decreases monotonically with increasing wavelength. The absorption of light by blood is highly dependent upon wavelength and oxygenation, as illustrated in Fig. 3.11. The absorption peaks of oxyhemoglobin occur at 418, 542, and 577 nm. Isobestic points are wavelengths where absorption for oxy- and deoxyhemoglobin are equal (i.e. curves in Fig. 3.11 cross). 548, 568, 587, and 805 nm. Oximetry typically measures reflection from or transmission through tissue using the 805 nm isobestic wavelength and a higher and/or lower wavelength to estimate the oxygen concentration of blood [8]. For UV wavelengths, the primary absorbers are protein and amino acids, whereas water is the important absorbing chromophore for IR wavelengths. Water absorption as a function of wavelength is presented in Fig. 3.12; peak values owing to symmetric and asymmetric stretch modes of water occur at 0.97, 1.44, 1.94, 2.94, 4.68 and 6.1 μm. Selected values about these peaks are listed in Table 3.1. A complete set of values of the water absorption coefficient as a function of wavelength can be found at http://omlc.ogi.edu/spectra/water/abs/idex.html. The peaks at 1.96 and 2.94 μm provide sufficient absorption for ablation of soft tissue and bone. Of considerable interest in the IR are the vibrational modes of peptide O=C–N–H at 6.1 μm (Amide I) and 6.45 (Amide II).

Above a wavelength of 1.4 μm, light scattering in tissue becomes insignificant compared to absorption. The primary chromophore is water, and the absorption coefficient of tissue can be estimated by the percent concentration of water. For example, at 2.0 μm the mass fractions of water in the epidermis and dermis are approximately 0.3 and 0.8 respectively. Thus, for a water absorption coefficient of 69.1 cm^{-1}, μ_a values for the epidermis and dermis of 21 cm^{-1} and 55 cm^{-1}, respectively, are not unreasonable values.

However, near the absorption peaks of water, μ_a is a function of local temperature owing to the symmetric stretch (v_1), symmetric bend (v_2) and asymmetric stretch

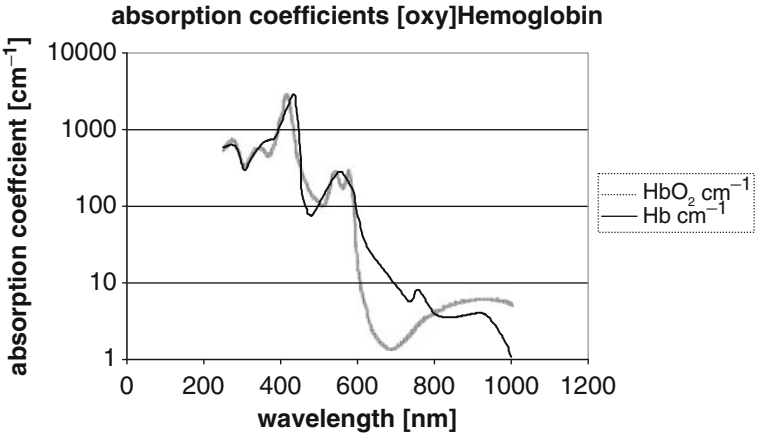


Fig. 3.11 Absorption coefficient of oxy- and deoxyhemoglobin as a function of wavelength. Absorption peaks of Hb and HbO_2 are located at 418, 542, and 577 nm. Wavelengths where the absorbency of HbO_2 are equal (isobestic points) are 548, 568, 587, and 805 nm [1]

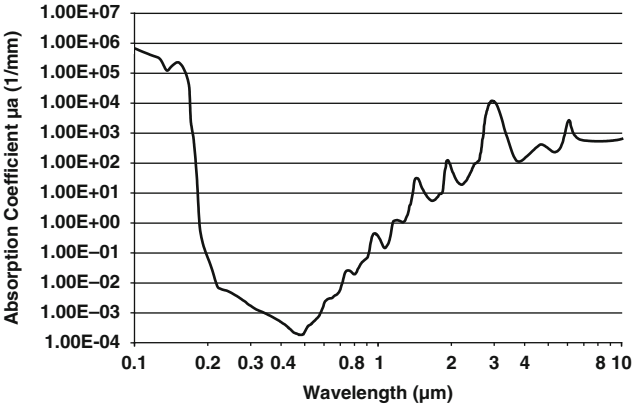


Fig. 3.12 Absorption coefficient of water at room temperature

Table 3.1 Absorption peaks of water in nm and absorption coefficients about these peaks

973 nm		1440 nm		1940 nm		2940 nm		6100 nm	
λ [nm]	μ_a [cm^{-1}]	λ [nm]	μ_a [cm^{-1}]	λ [nm]	μ_a [cm^{-1}]	λ [nm]	μ_a [cm^{-1}]	λ [nm]	μ_a [cm^{-1}]
960	0.42	1400	12.39	1900	66.14	2800	5161	5900	1325
970	0.45	1420	22.12	1920	114.54	2850	8157	6000	2241
975	0.45	1440	28.80	1940	119.83	2900	11,613	6100	2699
980	0.43	1460	28.40	1960	105.79	2950	12,694	6200	1784
990	0.41	1480	21.23	1980	92.03	3100	7783	6300	1127

modes of water [9]. Details of the effect of dynamic changes in optical absorption around water absorption peaks are discussed in Chapter 9. An example of the change in absorption at the 3 μm peak is given in Chapter 15. Basically, the absorption peak shifts to a lower wavelength as water temperature increases. Thus, during an Er:YAG (2.94 μm) pulse irradiation of water, the absorption dramatically decreases whereas absorption increases during an Er,Cr:&SGG (2.79 μm) pulse.

3.4.2 Propagation of Scattered Light

If tissues were just light absorbing media, that is non-scattering, the resulting spatial light distribution following laser irradiation could easily be described by a simple Beer's law of (exponential) attenuation, Eq. (3.30). Only an absorption coefficient would be required to compute the distribution of light in tissue. We all know, however, that tissues are somewhat whitish materials rather than black, demonstrating that they scatter visible light. Scattering is usually caused by random spatial variations in tissue density, refractive index and dielectric constant, and actual light distributions can be substantially different from distributions estimated using Beer's law. For example, scattering can cause the light fluence rate ϕ [W/m^2] just below the surface of the tissue to be larger than the incident irradiance E_0 [W/m^2]. At wavelengths where tissue absorption is low (600–1100 nm) and boundary conditions are matched, the increase in fluence rate relative to irradiance may be a factor of two to three, with a maximum of 5 (Chapter 6, Section 6.4.1.3). Even larger ratios (4–7) occur for hollow organs such as the bladder and blood vessels that act as integrating sphere and cylinder, respectively. Also, scattering extends the light fluence rate beyond the lateral dimensions of the incident irradiance. The magnitude of these effects and their importance depends rather strongly on (1) the scattering and absorbing properties of the tissue, and (2) the diameter of the laser spot relative to the penetration depth of the light.

Most important, the scattering coefficient and absorption coefficient are wavelength dependent. The absorption coefficient of tissue can vary strongly over the wavelength range of clinical laser medicine (from 193 to 10,600 nm), whereas the scattering coefficient typically decreases monotonically with wavelength. However, blood is an exception [10, 11]. The ratio of these coefficients can also vary substantially with (laser) wavelength.

Light scattered from the collimated beam undergoes multiple scattering events as it propagates through tissue. A rigorous description of this propagation in terms of Maxwell's wave equations is not possible at this time. An approach that has proven effective is the transport equation that describes the transfer of energy through a turbid medium [4, 12]. The theory is heuristic and is based on a statistical approximation of photon particle transport in a multiple scattering medium. Ishimaru points out that although "transport theory was developed on the addition of powers, it contains information about the correlation of the fields [4]." Although polarization can be included in transport theory (see Chapter 7), polarization of laser sources is typically neglected.

3.4.3 Transport Equation

The transport equation relates the gradient of radiance L at position $\boldsymbol{r}$ in direction $\hat{\boldsymbol{s}}$ to losses owing to absorption and scattering and to a gain owing to light scattered from all other directions $\hat{\boldsymbol{s}}'$ into direction $\hat{\boldsymbol{s}}$. The equation has the form [4]

$$\frac{dL(\boldsymbol{r},\hat{\boldsymbol{s}})}{ds} = \mu_a L(\boldsymbol{r},\hat{\boldsymbol{s}}) - \mu_s L(\boldsymbol{r},\hat{\boldsymbol{s}}) + \int_{4\pi} p(\hat{\boldsymbol{s}},\hat{\boldsymbol{s}}')L(\boldsymbol{r},\hat{\boldsymbol{s}}')d\omega' + S(\boldsymbol{r},\hat{\boldsymbol{s}}) \tag{3.36}$$

where L is the radiance [W/m^2 · sr], μ_a is the absorption coefficient [1/m], μ_s is the scattering coefficient [1/m], p is the phase (scattering) function [1/sr], and S is the optical source of power generated at $\boldsymbol{r}$ in direction $\hat{\boldsymbol{s}}$ [W/m^3 · sr].

In Chapter 6, the radiance is separated into two components: scattered light L_s and primary light L_p. The primary light represents non-scattered light from external sources. For example, an irradiance $E(\boldsymbol{r},\boldsymbol{s}_0)$ [W/m^2] can be represented as

$$L_p(\boldsymbol{r},\hat{\boldsymbol{s}}) = E(\boldsymbol{r},\hat{\boldsymbol{s}}_0)\frac{\delta(1-\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{s}}_0)}{2\pi} \tag{3.37}$$

where δ is the Dirac delta function and

$$\frac{\delta(1-\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{s}}_0)}{2\pi} = \delta(\varphi-\varphi_0)\delta(\cos\theta-\cos\theta_0) \tag{3.38}$$

A simple proof or demonstration of Eq. (3.38) has been offered in Chapter 6, Appendix 1. Let the integration variable be $\hat{\boldsymbol{s}}$. Multiply both sides of Eq. (3.38) by a function $f(\hat{\boldsymbol{s}}) = f(\varphi,\mu)$ and integrate over $d\omega$. Because of the definition of a delta function, this yields $f(\hat{\boldsymbol{s}}_0)$ on the right hand side. For the left side we choose the coordinates such that $\hat{\boldsymbol{s}}_0$ is in the z direction so that $\mu_0 = 1$, where $\mu = \cos\theta$. The integration of φ gives 2π; and since now $\delta(1-\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{s}}_0) = \delta(1-\mu)$, the angle between $\hat{\boldsymbol{s}}$ and $\hat{\boldsymbol{s}}_0$, the integration over μ yields $f(\mu_0) = f(\hat{\boldsymbol{s}}_0)$.

The optical source term for the scattered light includes the first scattering of primary light, light generated within the tissue, or internal light sources.

The phase function $p(\hat{\boldsymbol{s}},\hat{\boldsymbol{s}}')$ is the normalized probability density function which describes single scattering from unit direction $\hat{\boldsymbol{s}}'$ to unit direction $\hat{\boldsymbol{s}}$ at position $\boldsymbol{r}$; $d\omega'$ is the infinitesimal solid angle about $\hat{\boldsymbol{s}}'$.

The gradient of radiance in the $\hat{\boldsymbol{s}}$ direction has three equivalent forms:

$$\frac{dL(\boldsymbol{r},\hat{\boldsymbol{s}})}{ds} = \hat{\boldsymbol{s}}\cdot\nabla L(\boldsymbol{r},\hat{\boldsymbol{s}}) = \nabla\cdot[L(\boldsymbol{r},\hat{\boldsymbol{s}})\hat{\boldsymbol{s}}] \tag{3.39}$$

(See Chapter 6.)

Equation (3.36) originates from an energy balance of the radiance, incident upon a cylindrical volume with infinitesimal cross section dA [m^2] and infinitesimal

length ds. The decrease in the radiance while transversing infinitesimal distance ds, due to absorption and scattering, is

$$dL(\boldsymbol{r},\hat{\boldsymbol{s}}) = -ds(\mu_a + \mu_s)L(\boldsymbol{r},\hat{\boldsymbol{s}}) \tag{3.40}$$

whereas the increase in $L(\boldsymbol{r},\hat{\boldsymbol{s}})$ due to scattering from the radiance from all other directions $\hat{\boldsymbol{s}}'$ into $\hat{\boldsymbol{s}}$ at position $\boldsymbol{r}$ is

$$dL(\boldsymbol{r},\hat{\boldsymbol{s}}) = ds \int_{4\pi} \mu_s p(\hat{\boldsymbol{s}},\hat{\boldsymbol{s}}')\, L(\boldsymbol{r},\hat{\boldsymbol{s}}')d\omega' \tag{3.41}$$

Adding the contributions of equations (3.39) and (3.40) and dividing by ds yields the net rate of change of radiance $L(\boldsymbol{r},\hat{\boldsymbol{s}})$, over length ds as given in Eq. (3.36). More details are found in Ref. [4] and Chapter 6.

3.4.3.1 Phase Function of Scattering

Although tissues (e.g. tendon, brain tissue, muscle) are typically anisotropic, that is, the optical properties are a function of direction $\hat{\boldsymbol{s}}$, we assume that scattering depends only on the angle θ between unit vector directions $\hat{\boldsymbol{s}}$ and $\hat{\boldsymbol{s}}'$. Scattering from direction $\hat{\boldsymbol{s}}'$ to $\hat{\boldsymbol{s}}$ is illustrated in Fig. 3.13. Our analysis requires tissue to be isotropic in terms of physical properties (such as refractive index, density, etc.). For completeness, the volumetric, single scattering probability density function $p(\hat{\boldsymbol{s}},\hat{\boldsymbol{s}}')$ can therefore be written as

$$p(\hat{\boldsymbol{s}},\hat{\boldsymbol{s}}') = p(\varphi,\cos\theta) \tag{3.42}$$

when $\hat{\boldsymbol{s}}'$ is aligned along the z axis of Fig. 3.5, θ is the angle between $\hat{\boldsymbol{s}}'$ and $\hat{\boldsymbol{s}}$. φ is the azimuthal angle and because of the assumed isotropy, all φ are equally likely. We assume φ and θ are independent random variables so

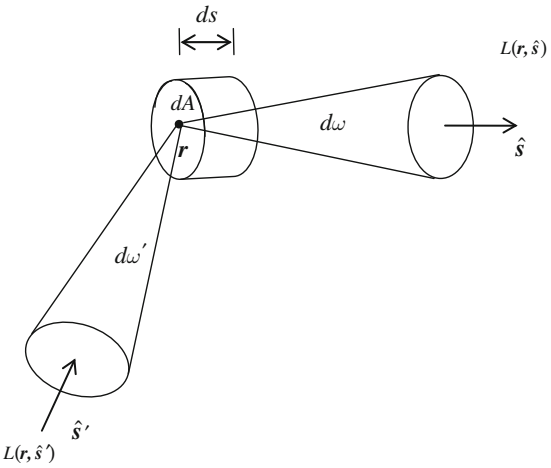


Fig. 3.13 Scattering of radiance incident upon volume $ds\,dA$ at position $\boldsymbol{r}$ from direction $\hat{\boldsymbol{s}}'$ into direction $\hat{\boldsymbol{s}}$

$$p(\hat{\boldsymbol{s}}, \hat{\boldsymbol{s}}') = p(\varphi)p(\cos\theta) = p(\varphi)p(\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{s}}') \tag{3.43}$$

The integral of the density functions over their domains must equal 1.0, that is

$$\int_{4\pi} \boldsymbol{p}(\hat{\boldsymbol{s}}, \hat{\boldsymbol{s}}')d\omega' = \int_{0}^{2\pi} \boldsymbol{p}(\varphi)d\varphi \int_{-1}^{+1} \boldsymbol{p}(\cos\theta)\boldsymbol{d}\cos(\theta) = 1 \tag{3.44}$$

Because of the isotropic scattering assumption for φ, the density function $p(\varphi)$ has a constant value

$$p(\varphi) = \frac{1}{2\pi} \tag{3.45}$$

Including $p(\varphi)$ in the phase function $p(\hat{\boldsymbol{s}}, \hat{\boldsymbol{s}}')$ is necessary for the Monte Carlo simulations described in Chapter 5. If single scattering is isotropic in φ and $\cos\theta$, then $p(\cos\theta) = 1/2$ and the phase function $p(\hat{\boldsymbol{s}}, \hat{\boldsymbol{s}}')$ is equal to a constant

$$p(\hat{\boldsymbol{s}}, \hat{\boldsymbol{s}}') = p(\varphi)p(\mu) = \left(\frac{1}{2\pi}\right)\left(\frac{1}{2}\right) = \frac{1}{4\pi} \tag{3.46}$$

where $\mu = \cos\theta$.

Unfortunately (from a computational point of view), light scattering in tissue is not isotropic even when tissue is isotropic. Single particle light scattering is strongly forward directed. Measurements of angular dispersion are made using thin samples of tissue (thickness much less than one optical depth, $1/\mu_t$, to ensure single scattering). Scattering patterns such as the one in Fig. 3.14 for dermis at 633 nm illustrate the extent of the forward scattering. Several functions such as (1) Henyey-Greenstein, (2) modified Henyey-Greenstein and (3) delta-Eddington have been postulated to represent single scattering phase functions for tissue ([4, 12, 13], see also Chapter 6).

A measure for the degree of anisotropy in scattering is the anisotropy factor g, which represents the expected (average) scattering angle. Total forward scattering means $g = 1$ and isotropic scattering means $g = 0$. Note that although isotropic scattering implies $g = 0$, the inverse is not necessarily true. Thus, $g = 0$ does not necessarily mean isotropic scattering, only that integral over -1 to 1 of $p(\cos(\theta))d\cos(\theta) = 0$. Mathematically, g is defined as the expectation value of the cosine of the scattering angle θ, that is

$$g = \frac{\int_{4\pi} p(\hat{\boldsymbol{s}}, \hat{\boldsymbol{s}}')(\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{s}}')d\omega}{\int_{4\pi} p(\hat{\boldsymbol{s}}, \hat{\boldsymbol{s}}')d\omega} = \int_{4\pi} p(\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{s}}')(\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{s}}')d\omega \tag{3.47}$$

because integration of a density function over its entire domain equals 1.0; that is, $\int_{4\pi} p(\hat{\boldsymbol{s}}, \hat{\boldsymbol{s}}')d\omega = 1$.

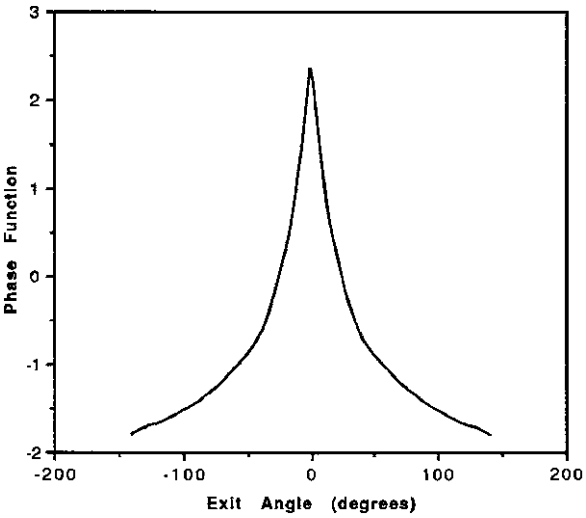


Fig. 3.14 Angular distribution of scattered light at 633 nm for a 0.1-mm-thick sample of human dermis [14]. Magnitude of phase function on $\log_{10}$ scale

The integration is over 4π steradians, since $p(\hat{s} \cdot \hat{s}')$ typically includes the $\frac{1}{2\pi}$ value that represents the isotropic scattering in φ.

For in vitro tissues at the visible and near infrared wavelengths, g turns out to be between about 0.7 and 0.99.

3.4.3.2 Phase Functions

In most literature and all but two chapters in this book, this chapter and Chapter 5, $p(\hat{s}, \hat{s}')$ is set equal to $p(\hat{s} \cdot \hat{s}')$ and the $1/2\pi$ for $p(\varphi)$ is included in $p(\hat{s} \cdot \hat{s}')$. Several single scattering functions have been proposed to represent photon scattering in Monte Carlo simulations, and in the Transport Equation. In Chapter 6, Star shows how these scattering functions are related to a Legendre expansion of the phase function.

Henyey-Greenstein

Perhaps the most used scattering function is the Henyey-Greenstein phase function [12–14] which we will initially write as

$$p^*_{HG}(\mu) = \frac{1}{2}(1 - g^2) \Big/ (1 - 2\,g\mu + g^2)^{3/2} \tag{3.48}$$

where $\mu = \cos\theta$, θ is the angle between $\hat{s}'$ and direction of scattering $\hat{s}$.

The total phase function is

$$p(\hat{s}, \hat{s}') = p(\varphi)p(\mu) = \frac{1}{2\pi}p(\mu) \tag{3.49}$$

which is universally defined as p_{HG}. Further, literature specifies that

$$p(\hat{s},\hat{s}') = p(\hat{s}\cdot\hat{s}') = 1\Big/4\pi\Big[\left(1-g^2\right)\Big/\left(1-2\,g\mu+g^2\right)^{3/2}\Big] = p_{HG} \quad (3.49a)$$

This phase function provides a reasonable estimate of the forward scattering nature of tissue. Although this function may not exactly represent the true phase function of biological tissue, it has been shown experimentally to be a good approximation. It is also quite practical to work with, particularly in model calculations. Furthermore, the detailed behavior of the phase function may not be important for many practical applications, such as calculations of the fluence rate.

It may not be possible to distinguish the "real" phase function from p_{HG} (see also p. 324 of Van de Hulst's book [13]). When $g \to 1, p_{HG}$ approaches the Dirac δ-function, $\delta(1-\hat{s}\cdot\hat{s}')/(2\pi)$.

Modified Henyey-Greenstein

If the anticipated back scatter will be greater than specified by the Henyey-Greenstein function, an isotropic component can be included as

$$p_{HG-m}(\mu) = \left(\frac{1}{2\pi}\right)\frac{1}{2}\left[\beta + (1-\beta)\frac{1-g^2}{(1+g^2-2\,g\mu)^{1.5}}\right] \quad (3.50)$$

where β is the fraction of light isotropically scattered.

Linear Anisotropic Phase Function

The linear anisotropic phase function,

$$p(\hat{s},\hat{s}') = p(\varphi)p(\mu) = \left(\frac{1}{2\pi}\right)\frac{1}{2}(1+3\,g\mu) \quad (3.51)$$

has the distinct disadvantage of having negative values as cos (θ) approaches negative 1. The equation suggests that photons with a negative valued phase function are scattered in the forward direction. Obviously Eq. (3.51) does not qualify as a probability density function, even though the integral of $p(\mu)$ over its domain is equal to one. However, the linear anisotopic phase function is consistent with the truncated phase function in the P_1 or Diffusion Approximation in Chapter 6. Any phase function becomes linear anisotropic when you cut the Legendre series off after the first two terms. Remember, except for Monte Carlo simulation, the $1/2\pi$ factor is included in the $p(\hat{s}\cdot\hat{s}')$ phase function.

A practical disadvantage of the linear anisotopic phase function and thus of Diffusion Theory is the accuracy of solutions when the scattered light is not diffuse: that is, near boundaries and near the collimated beam of light where single scattering events are important. Results are excellent in the diffusion region.

Delta Eddington Approximation

The linear anisotropic scattering function does not mimic microscopic single scattering in tissue. Joseph et al. suggested including a delta function with the linear anisotropic phase function to compensate for the lack of forward scattering [15]. The δ-Eddington phase function has the form:

$$p_{\delta-E}(\mu) = \left(\frac{1}{2\pi}\right)\frac{1}{2}[2f\delta(1-\mu) + (1-f)(1+3g\mu)] \tag{3.52}$$

where $0 \leq f \leq 1$ and f represents the fraction of photons in the $\hat{\boldsymbol{s}}$ direction scattered into $\hat{\boldsymbol{s}}'$. In terms of a primary ray of light with direction $\hat{\boldsymbol{s}}_0$, the first scattering event scatters the fraction f of photons into $\hat{\boldsymbol{s}}_0$. The Delta Eddington phase function reduces the domain of μ that causes $p_{\delta-E}(\mu)$ to be negative. More detail is given in Chapter 6.

3.4.4 Rate of Volumetric Absorption of Photon Energy; Relation Between F(r) and $\mu_a(\boldsymbol{r})\phi(\boldsymbol{r})$

The relation between flux vector $\boldsymbol{F}(\boldsymbol{r})$, Eq. (3.12a), and rate of volumetrically absorbed photon energy at $\boldsymbol{r}$, $\mu_a(\boldsymbol{r})\phi(\boldsymbol{r})$, follows from the transport equation (3.36) with the optical source term set equal to zero, and the notation of Eq. (3.39) (see also Chapter 6). The transport equation with units [W·m^{-3}sr^{-1}] is integrated over all 4π solid angle as follows to form an energy equation [W/m^3].

$$\int_{4\pi} s\cdot\nabla L(\boldsymbol{r},\hat{\boldsymbol{s}})d\omega = -\mu_t \int_{4\pi} L(\boldsymbol{r},\hat{\boldsymbol{s}})d\omega + \mu_s \int_{4\pi}\left[\int_{4\pi} p(\hat{\boldsymbol{s}},\hat{\boldsymbol{s}}')\right]L(\boldsymbol{r},\hat{\boldsymbol{s}})d\omega' d\omega \tag{3.53}$$

Equation (3.39) using the first term can be rewritten as,

$$\int_{4\pi} \hat{\boldsymbol{s}}\cdot\nabla L(\boldsymbol{r},\hat{\boldsymbol{s}})d\omega = \nabla\cdot\int_{4\pi} L(\boldsymbol{r},\hat{\boldsymbol{s}})\hat{\boldsymbol{s}}d\omega \tag{3.54a}$$

by noting that differentiation refers to coordinate $\boldsymbol{r}$ and integration to directions $\hat{\boldsymbol{s}}'$ and reversing the order of integration in the last term of Eq. (3.53). The terms of Eq. (3.53) are then easily evaluated using respectively

$$\int_{4\pi} L(\boldsymbol{r},\hat{\boldsymbol{s}})\boldsymbol{s}d\omega = F(\boldsymbol{r}), \quad \int_{4\pi} L(\boldsymbol{r},\hat{\boldsymbol{s}})d\omega = \phi(\boldsymbol{r}) \quad \int_{4\pi} p(\hat{\boldsymbol{s}},\hat{\boldsymbol{s}}')d\omega = 1; \tag{3.54b}$$

and the following important result is obtained

$$-\nabla\cdot\boldsymbol{F}(\boldsymbol{r}) = \mu_a(\boldsymbol{r})\phi(\boldsymbol{r}) \tag{3.55}$$

Equation (3.55) expresses energy conservation without a topical source term. The negative gradient of flux vector $\boldsymbol{F}(\boldsymbol{r})$ equals the rate of volumetric energy of the photons absorbed at $\boldsymbol{r}$ which is the *heat* source term $\boldsymbol{S}(\boldsymbol{r})$ of the heat conduction equation (see Chapter 10). Chapter 6, Eq. (6.29) illustrates the effect of a non-zero *optical* source term.

3.4.5 Penetration of Light in Tissue

What is the depth of light penetration in tissue when scattering dominates absorption ($\mu_a << \mu_s$)? The e^{-1} penetration depth of the collimated light is $\delta = 1/\mu_t$, but where is the e^{-1} depth of the scattered light? There is no simple answer, which is why several chapters are devoted to estimating the fluence rate of light in tissue. Important factors are (1) μ_a and μ_s, (2) phase function, (3) boundary conditions and diameter of the spot size. For the novice and even for many in the field, it goes against intuition that the fluence rate, ϕ [W/m^2] within tissue can be larger than the irradiance E_0 [W/m^2]. This paradox is illustrated using Monte Carlo simulations of light propagation in bloodless dermis as a function of beam diameter (see Fig. 3.15).

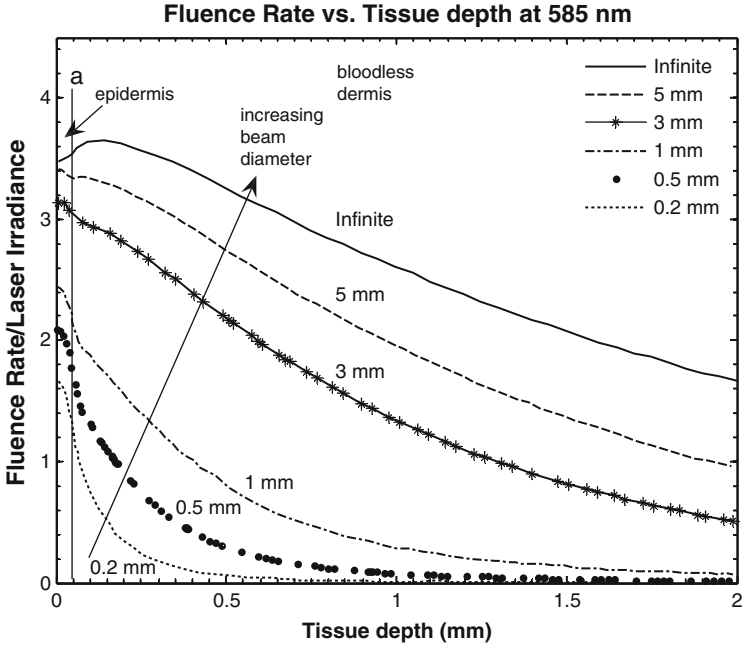


Fig. 3.15 The fluence rate within the skin below the beam center is fluenced (due to scattering) by the beam diameter. For the same incident irradiance (1 W/cm^2), the larger the beam (spot) diameter, the greater is the fluence rate. These Monte Carlo calculations were made for a flat top beam profile at 585 nm using optical properties summarized in Table 3.2

Table 3.2 Optical properties of skin (Van Gemert, MJC, unpublished notes)

	Wavelength (nm)	Absorption coefficient $\mu_a(\text{mm}^{-1})$	Scattering coefficient $\mu_s(\text{mm}^{-1})$	Anisotropy factor g
Epidermis	585	1.8	47	0.79
Dermis	585	0.024	12.9	0.79

The skin consists of two layers, the epidermis and dermis. Optical properties are listed in Table 3.2. The fluence rate 0.15 mm deep for an infinite beam diameter is 3.6 times larger than the irradiance at the air-tissue boundary. Scattering creates more light (fluence rate) just inside the tissue than the irradiance. The e^{-1} depth with respect to the irradiance is approximately 3.0 mm for a 5.0 mm spot diameter. However, the typical measures of collimated and effective penetration depth using dermis optical properties are $\delta = 0.08$ mm and $\delta' = 0.38$ mm respectively. The importance of beam diameter for the penetration depth of the scattered light is clearly illustrated by Fig. 3.15.

An energy balance is obtained by integrating the heat generation $S(\boldsymbol{r}) = \mu_a \phi(\boldsymbol{r})$ throughout the tissue and comparing the answer to the irradiance. The key to the paradox is sufficient backscattering and the small value of μ_a with respect to μ_s and remembering energy loss is associated only with the absorption of photons.

3.5 Photon Sources and Detection

3.5.1 Sources

3.5.1.1 Collimated

The primary source of tissue irradiation is typically a laser beam that is directed to a target through a fiber optic or as a collimated beam. The delivered light is characterized by its spot size, divergence, and radial profile. Quite often the irradiance of a target is computed by dividing the delivered power by the spot size. This creates several uncertainties in the reader's mind. Is there any energy beyond the quoted spot size, and is the irradiance uniform across the laser spot? A uniform irradiance implies a flat top (uniform) beam profile which is unusual for a collimated laser beam, which is typically Gaussian in shape, or perhaps doughnut for a (0,1; 1,0) mode. Some lasers (especially pulsed lasers) can have extremely non-uniform profiles with discernable "hot" spots. For this discussion we will consider two examples: "flat top" and Gaussian profiles, illustrated in Fig. 3.16. These profiles are specified by the following irradiances.

Flat top:

$$\begin{aligned} E(r) &= E_{0f} \text{ for } r \le a \\ E(r) &= 0 \text{ for } r > a \end{aligned} \tag{3.56}$$

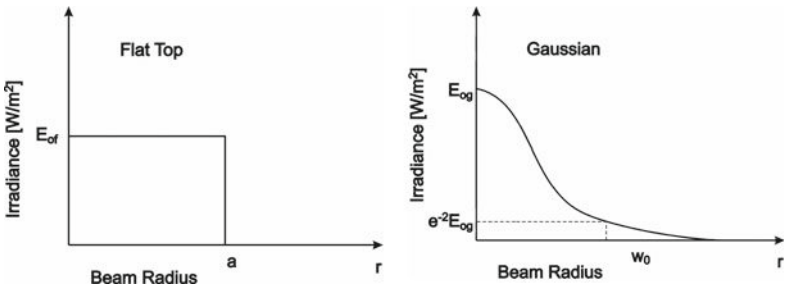


Fig. 3.16 Flat top and Gaussian beam profiles. Eighty five percent of Gaussian beam power is with in the w_0 radius

where a is the radius of the beam.

Gausian:

$$E_{0g} \exp\left(-2r^2 \Big/ w_0^2\right) \tag{3.57}$$

where w_0 is the $1/e^2$ radius of the beam.

The beam power associated with a Gaussian beam is determined by integrating the irradiance profile over r and around its axis.

$$\int_0^{2\pi} \int_0^{\infty} E_{0g} \exp\left(-\frac{2r^2}{w_0^2}\right) r dr d\varphi = \frac{\pi w_0^2 E_{0g}}{2} \tag{3.58}$$

whereas the power associated with a flat top beam with radius a is

$$p = \pi a^2 E_{0f} \tag{3.59}$$

If $a = w_0$ and $E_{0\,g} = 2E_{0f}$, then the same power is associated with each beam. However, 13.5% of the laser power in a Gaussian beam is beyond the $1/e^2$ radius w_0. The half maximum radius for a Gaussian beam is $r = 0.588\, w_0$ which encompasses half of the beam power.

When the laser beam is delivered through a fiber optic, the near field exit beam is a flat top diverging beam; however, the near field only extends to a distance approximately the radius of the fiber core (see Chapter 17). Nevertheless, the flat top profile can be extended by placing a spherical tip or other optics at the exit surface. Far field images become Gaussian with divergences approximately equal to the incoupling angle. Bends in the fiber disrupt modes which increase the exit divergence.

3.5.1.2 Diffuse Sources

A diffuse flux can be created by laser irradiation of a diffusing material or from an output port of an integrating sphere (see Chapter 8). Diffuse light from a surface

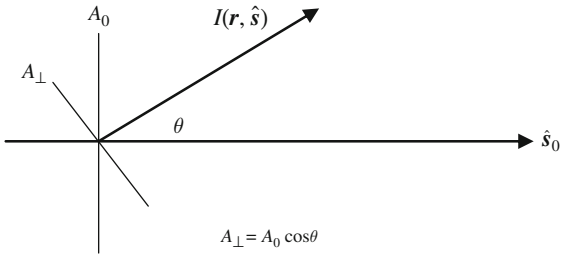


Fig. 3.17 Intensity $I(\boldsymbol{r}, \hat{\boldsymbol{s}})$ from Lambertian source on plane A_0

(source) is characterized by its radiance $L(\boldsymbol{r}, \hat{\boldsymbol{s}})$ [W/m^2 · sr] when the radiance is independent of direction for all $\boldsymbol{r}$ and $\hat{\boldsymbol{s}}$

$$L(\boldsymbol{r}, \hat{\boldsymbol{s}}) = L_0 \tag{3.60}$$

which is designated as a *Lambertian Source*. The intensity $I(\boldsymbol{r}, \boldsymbol{s})[W/sr]$ is determined by integrating L_0 over the surfaces perpendicular to direction $\hat{\boldsymbol{s}}$. By letting θ be the angle between $\hat{\boldsymbol{s}}_0$ and $\hat{\boldsymbol{s}}$ as illustrated in Fig. 3.17

$$I(r, s) = \int_{A_0} L_0 \cos\theta \, dA \tag{3.61}$$

As $\theta \rightarrow \pi/2$, the perceived intensity of a source with finite area decreases to zero because the projected area $A_\perp \rightarrow 0$.

3.5.1.3 Point Source

The isotropic point source can be physically approximated using a highly scattering tip at the end of the fiber optic. In theory, a point source at position $\boldsymbol{r}$ delivering power P creates an intensity $I(\boldsymbol{r})$ [W/sr] of

$$I(\boldsymbol{r}, \hat{\boldsymbol{s}}) = \frac{P}{4\pi} \tag{3.62}$$

in the solid angle $d\Omega$ to an imaginary sphere with radius R

$$dP = \frac{P}{4\pi} d\Omega = \frac{P}{4\pi} \frac{dA}{R^2} \tag{3.63}$$

And the irradiance is

$$E(R) = \frac{P}{4\pi R^2} \tag{3.64}$$

at all points on the sphere.

3.5.2 Detection

The amount of light at the surface of a detector is given by the flux normal to its surface, that is, the irradiance. If the radiance at the surface is $L(\boldsymbol{r},\hat{\boldsymbol{s}}) = L_0$, then the irradiance is πL_0 [W/m^2] (Eq. (3.13c)). Otherwise, it is necessary to integrate $L(\boldsymbol{r},\boldsymbol{s})$ to determine the flux

$$F_n(\boldsymbol{r}) = \int_{2\pi} L(\boldsymbol{r},\boldsymbol{s})\cos\theta d\omega = \int_0^{\pi/2}\int_0^{2\pi} L(\boldsymbol{r},\boldsymbol{s})\cos\theta\sin\theta d\theta d\varphi \tag{3.65}$$

The incremental power reaching a target or detector from radiance at point $\boldsymbol{r}$ in direction $\boldsymbol{s}_0$, as illustrated in Fig. 3.18, is given by

$$dP = L(\boldsymbol{r},\hat{\boldsymbol{s}})dA_s d\Omega \tag{3.66}$$

Note, $d\Omega$ can be replaced by $\frac{dA_d}{R^2}$. Integration over the area of the source produces the power reaching the detector. The irradiance on dA_d is

$$E(dA_d) = L(\boldsymbol{r},\hat{\boldsymbol{s}})\frac{dA_s}{R^2} \tag{3.67}$$

When either or both the source and detector are rotated by angles θ_s and θ_d respectively, the projected areas become $A_s\cos\theta_s$ and $A_d\cos\theta_d$.

$$dP_d = \frac{L_0 dA_s\cos\theta_s dA_d\cos\theta_d}{R^2} \tag{3.68}$$

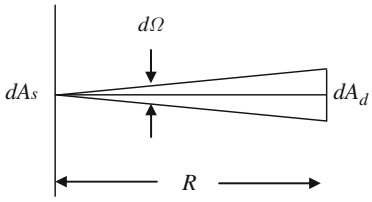


Fig. 3.18 Detector dA_d located at distance R from source

Care must be taken in the placement of detectors to maximize the detected signal. The actual power received at a target can be determined by computing the irradiance reaching the target. Consider a fiber optic detector with a half acceptance angle of 15° and circular source with radius R_s. The distance between source and detector is z_0. (See Fig. 3.19).

By considering a differential source area $dA_s = rdrd\varphi$, the differential power at dA_d is

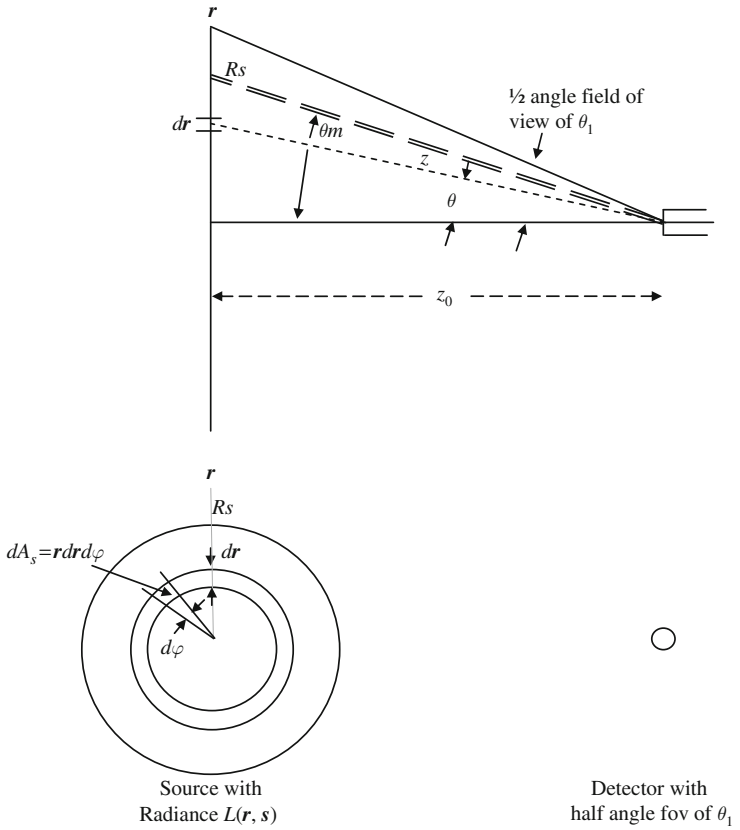


Fig. 3.19 Circular source and detector at distance z_0

$$dP_d = L(\boldsymbol{r},\boldsymbol{s})dA_s \cos\theta \frac{dA_d \cos\theta}{z^2} \tag{3.69}$$

where

$$\begin{aligned} z &= \tfrac{z_0}{\cos\theta} \\ dA_s &= rdrd\varphi \\ r &= z_0 \tan\theta \\ dr &= z_0 \sec^2\theta d\theta \end{aligned}$$

Making these substitutions, the irradiance on dA_d is

$$E(dA_d) = L(\boldsymbol{r},\hat{\boldsymbol{s}}) \cos\theta \sin\theta d\theta d\varphi \tag{3.70}$$

and

$$dP_d = L(\boldsymbol{r},\hat{\boldsymbol{s}})dA_d \cos\theta \sin\theta d\theta d\varphi$$

Irradiance on A_d, a Lambertian surface with radiance L_0, is

$$E(A_d) = L_0 \int_0^{2\pi} \int_0^{\theta} \frac{1}{2} \cos\theta \sin\theta d\theta d\varphi \tag{3.71}$$

where θ_1 is the field of view half angle.

If $\theta_{\max}$ (maximum angle of source to detector) is less than θ_1, then

$$E(A_d) = \pi L_0 \sin^2\theta \Big|_0^{\theta_{\max}} + \pi L_B \frac{\sin^2\theta}{z} \Big|_{\theta_{\max}}^{\theta_1} \tag{3.72}$$

where L_B is the background radiance, which is equal to

$$E(A_d) = \pi L_0 \sin^2\theta_1 \text{ if } \theta_1 < \theta_m \tag{3.73}$$

since $\sin^2\theta_{\max} = \dfrac{r_s^2}{r_s^2 + z_0^2}$

In the near field as $z_0 \to 0$, $E(A_d) = \pi L_0$, while in the far, far field where $z_0 >> r_s$,

$$E(A_d) = \pi L_0 \frac{r_b^2}{z_0^2} \tag{3.74}$$

However, in the not so near field, the radiance entering the fiber is not Lambertian because of the limited field of view of the detector. Also, these equations did not consider the mismatch in index of refraction between the fiber and environment. Exact determination would require a Monte Carlo simulation (see Chapter 5).

3.5.3 Measurement of Fluence Rate

According to the definition of radiant energy fluence rate (Section 3.2.2.3), its measurement requires a probe with isotropic response, which experimentally performs the integration of the radiance as given by Eq. (3.11). Isotropic probes usually consist of a nearly spherical bulb of highly light scattering material, mounted on an optical fiber [16, 17]. Typical dimensions are 1 mm or less for the bulb and 0.2 mm core diameter for the fiber. Both probe and fiber should be as small as possible to minimize disturbance of the light distribution.

The refractive index of the probe material is generally larger than that of the medium in which the measurements are performed (mostly water or mammalian tissue). Part of the scattered light in the probe material is then totally reflected at the boundary with the medium (Chapter 6, Section 6.3.9) causing an increased response relative to a refractive index matched situation with the same fluence rate.

It is therefore important to calibrate a probe in a clear medium with the same refractive index as the scattering medium in which the fluence rate is to be measured, or apply a correction which can be measured and also calculated using diffusion theory [17–19]. Calibration in a clear medium appears to be adequate for quantitative measurements in a turbid medium [20]. Fluence rate measurements in optical phantoms with known optical properties agree reasonably well with theoretical calculations [21, 22].

Tissue response in photodynamic therapy is determined by the number of photons absorbed by a photosensitizer. Clinical photodynamic therapy is often performed in hollow organs, where superficial lesions are to be treated. The contribution of scattered light to the total fluence rate may then be several times larger than the fluence rate of the nonscattered light from the primary light source. To correlate response with "light dose" it is therefore important to measure the total fluence rate using isotropic probes. This has turned out to be quite useful in, e.g., the urinary bladder [23] (see also Chapter 6, Exercise 6.7), oral cavity [24], esophagus [25] and nasopharynx [26].

3.6 Summary

Transport theory is a heuristic model which lacks the physical rigor of multiple scattering EM theory: for example, it does not in itself include effects such as diffraction or interference, even though the absorption and scattering properties of the individual constituent particles may do so. It is fundamental to transport theory that there should be no correlation between the radiation fields. Only quantities such as power or intensity are considered, and the method ignores the behavior of the component wave amplitudes and phases.

There has been a limited amount of work on establishing the connection between the transport and EM theories of radiation propagation in multiply-scattering media, which gives some insight into the physical meaning of the empirical transport parameters. It can be shown, for example, that under certain conditions, the radiance at any point, $\boldsymbol{r}$, in the medium according to transport theory, (i.e. the power flow per unit area per unit solid angle) is equivalent to the statistical average of the (random) time-varying Poynting vector $(\boldsymbol{E}(\boldsymbol{r}, t) \times \boldsymbol{H}(\boldsymbol{r}, t))$ of the local EM field [4, 6]. It is also related, through the Fourier Transform, to the so-called mutual coherence function of multiple scattering EM theory, which is a measure of the correlation between the local fields around the point [4].

Despite such connections, the development and application of transport theory for light propagation in tissue has proceeded largely without reference to EM theory, except for some interpretation of the scattering phase functions in terms of Mie scattering from single particles whose size is much greater than the wavelength. For example, when Mie phase functions are fit to experimental data for soft tissues, the resulting equivalent particle sizes are in the order of the cell diameter [27]. However, this does not mean that tissue really cannot be considered properly as simply a dense cell suspension, since there are significant differences in the scattering

properties of different sub-cellular organelles and the intracellular matrix is certainly not an optically-clear fluid. Thus, the validity of transport theory rests primarily on its utility.

Transport theory has provided a self-consistent framework for studies of light propagation in tissues. For example, if the absorption and scattering properties of a tissue are determined according to the conditions of transport theory (i.e. only radiometric quantities that describe the conservation of power and not wave amplitudes and phases are measured) and, subsequently, these data are used as input to a transport model to calculate, say, the spatial distribution of fluence rate within a tissue volume for given irradiation and boundary conditions, then it is found in general that such distributions agree with experimental values (again measured ignoring light wave properties), although this has never been demonstrated rigorously. Such self-consistency does not imply physical validity, however, and there are some indications that transport theory may break down in the case of highly-structured tissues such as muscle, where the alignment of fibers may cause measurable wave interference effects, and the scatters cannot be considered random [28].

Even accepting that transport theory applies to tissue, there are no exact general solutions to the transport equations in tissue. In other applications of transport theory it has been useful to separate the problems into three regimes: tenuous scattering media where the fraction of the physical volume occupied by the scattering particles (i.e. the volume density) is low, $<10^{-3}$; intermediate scattering media; and dense media (volume density $>10^{-2}$). For tenuous media first-order multiple scattering theory applies, in which the total light fluence rate at any point is given approximately by the attenuated incident light fluence rate and the contribution from photons scattered from other regions of the medium can be ignored.

For the dense medium, which does apply to soft tissues at optical wavelengths, the radiation transport equations can be simplified to yield the Diffusion theory approximately (see Chapter 6). The diffusion equations can be solved either analytically for special cases or, more generally, by numerical techniques. This has been very productive in many applications involving light-tissue interactions. The validity of diffusion theory is limited in tissue to cases where the light has been highly scattered, which depends on the average distance between scattering centers, on the optical albedo, and the source-tissue-detector geometries. However, this includes many cases of interest in laser-tissue interactions in photomedicine at visible and near-infrared wavelengths.

For situations where diffusion theory breaks down, the most useful method has been to apply Monte Carlo modeling to simulate photon transport (see Chapter 5). In this computational technique, the multiple-scattering trajectories of individual photons are traced through the medium, each interaction being governed by the random processes of absorption or scattering. Physical quantities of interest, such as the local fluence, are thus "scored" within the statistical uncertainties of the finite number of photons simulated. The power of the Monte Carlo methods lies in its ability to handle virtually any source, detector and tissue boundary conditions, as well as any combination of tissue optical properties. However, it has the fundamental limitation that the optical parameter space is sampled only one "point" at a time,

so that any single Monte Carlo simulation gives little insight into the functional relationships between measureable quantities and the optical properties.

All the methods based on photon transport have been successful in practice because, to date, the photobiological effects of interest are the result of only the (rate of) local energy absorption in the tissue and are not sensitive to the separate wave amplitudes and phases. Even in this case, however, 3 wave phenomena could alter the local light field intensity and so change the energy absorption rate. Now, some diagnostic applications of tissue optics require EM theory to be applied. Meantime, the essential empiricism of transport theory should be kept in mind, and over-interpretation of the physical meaning of the transport parameters avoided. A more detailed discussion of EM theory applied to tissue optics is contained in Chapters 4 and 7, and Optical Coherence Tomography applications are presented in Chapter 19.

Appendix

Solid Angles

A solid angle is defined as an area on the surface of a sphere with unit radius. The area may have an arbitrary shape and size on the sphere surface. The total area of a sphere surface with radius r is $4\pi r^2$. The solid angle associated with any area on the sphere surface (radius r) is

$$\text{solid angle} = \frac{\text{area on sphere surface}}{r^2} \tag{3.75}$$

with unit of steradians, abbreviated as [sr]. Thus, the maximum possible solid angle is 4π [sr].

Often, an element of solid angle, $d\omega$, is defined about a direction $\hat{\boldsymbol{s}}$. We express the components of $\hat{\boldsymbol{s}}$ in a Cartesian frame of reference as $\hat{\boldsymbol{s}}_x, \hat{\boldsymbol{s}}_y, \hat{\boldsymbol{s}}_z$. In polar coordinates these components are

$$\hat{\boldsymbol{s}}_x = r \sin\theta \cos\varphi \tag{3.76a}$$

$$\hat{\boldsymbol{s}}_y = r \sin\theta \sin\varphi \tag{3.76b}$$

$$\hat{\boldsymbol{s}}_z = r \cos\theta \tag{3.76c}$$

The element of solid angle, $d\omega$, about direction $\hat{\boldsymbol{s}}$, then follows as (Eq. (3.76c)).

$$d\omega = \frac{rd\theta\, r \sin\theta\, d\varphi}{r^2} = \sin\theta\, d\theta\, d\varphi \tag{3.77}$$

References

1. Roggan A, Friebel M, Dőrschel K, Hahn A, and Műller G. Optical properties of circulating human blood in the wavelength range 400–2500 nm. *J. Biomed. Opt.,* 4(1):36–46 (1999).
2. Hale GM and Querry MR. Optical constants of water in the 200-nm to 200-μm wavelength region. *Appl. Opt*., 12:555–563 (1973).
3. Quantities and units of light and related electromagnetic radiations, 2nd edition, *International Standard ISO* 31/6, 1980(E), International Organization for Standardization, Switzerland (1980).
4. Ishimaru A. *Wave propagation and scattering in random media*, Vol. 1. Academic, New York (1978).
5. Chapter 1, *Handbook of optics*, Infrared Target and Background Radiometric Measurements: Concepts, Units and Techniques. Report of the Working Group on Infrared Background (WGIRB) No. 2389-64-T, IRIA, Institute of Science and Technology, University of Michigan, Ann Arbor, Michigan, 1962; also *Infrared Phys.,* 3:139–169 (1963).
6. Hecht E and Zajac A. *Optics*. Addison-Wesley, Reading, MA, 2nd edition (1990).
7. Born M and Wolf E. *Principles of optics*. The Macmillan, New York (1964).
8. Mendelson Y. Blood gas measurement, transcutaneous. In: JG Webster (ed) *Encyclopedia of medical devices and instrumentation*, Vol. 1, pp. 448–460 (1988).
9. Cumming JP and Walsh JT. Erbium laser ablation: The effect of dynamic optical properties, *Appl. Phys. Lett.,* 62:1988–1990 (1993).
10. Faber DJ, Aaiders MC, Mik EG, Hooper BA, van Gemert MJ, and van Leeuwen TG. Oxygen saturation-dependent absorption and scattering of blood. *Phys. Rev. Lett.*, 93(2):028102 (2004).
11. Vodopyanov KL. Saturation studies of H_2O and HDO near 3400 cm-1 using intense picosecond laser pulses. *J. Chem. Phys.,* 94:5389–5393 (1991).
12. Chandrasekhar, S. *Radiative transfer.* University Press, Oxford (1960).
13. van de Hulst HC. *Multiple light scattering: tables, formulas and applications,* Vols. I and II. Academic, New York (1980).
14. Prahl S. *Light transport in tissue*, dissertation, University of Texas, Austin (1988).
15. Joseph HJ, Wiscombe WJ, and Weinman JH. The delta-Eddington approximation for radioactive flux transfer. *J. Atmos. Sci.,* 33:2452–2459 (1976)
16. Marijnissen JPA and Star WM. Quantitative light dosimetry in vitro and in vivo. *Lasers Med. Sci*., 2:235–242 (1987).
17. Van Staveren HJ, Marijnissen HPA, Aalders MCG, and Star WM. Construction, quality assurance and calibration of spherical isotropic light diffusers. *Lasers Med. Sci.,* 10:137–147 (1995).
18. Star WM and Marijnissen JPA. Calculating the response of isotropic light dosimetry probes as a function of the tissue refractive index. *Appl. Opt*., 28:2288–2291 (1989).
19. Marijnissen JPA and Star WM. Calibration of isotropic light dosimetry probes based on scattering bulbs in clear media. *Phys. Med. Biol.,* 41:1191–1208 (1996).
20. Marijnissen JPA and Star WM. Performance of isotropic light dosimetry probes based on scattering bulbs in turbid media. *Phys. Med. Biol*., 47:2049–2058 (2002).
21. Allen V and McKenzie AL. The modified diffusion dipole model. *Phys. Med. Biol.,* 36: 1621–1638 (1991).
22. Rinzema K, Murrer LHP, and Star WM. Direct experimental verification of transport theory in an optical phantom. *J. Opt. Soc. Am*., A15:2078–2088 (1998).
23. D'Hallewin MA, Baert L, Marijnissen JPA, and Star WM. Whole bladder wall photodynamic therapy with in situ light dosimetry for carcinoma in situ of the bladder. *J. Urol.*, 148: 1152–1155 (1992).
24. Tan IB, Oppelaar H, Ruevekamp MC, Veenhuizen RB, Timmers A, and Stewart FA. The importance of in situ light dosimetry for photodynamic therapy of oral cavity tumors. *Head Neck,* 21:434–441 (1999).

25. Van Veen RLP, Aalders MCG, Pasma KL, Siersema PD, Haringsma J, van de Vrie W, Gabeler EEE, Robinson DJ, and Sterenborg HJCM. In situ light dosimetry during photodynamic therapy of Barrett's esophagus with 5-aminolevulinic acid. *Lasers Surg. Med*., 31:299–304 (2002).
26. van Veen RLP, Nyst H, Rai Indrasari S, Adham Yudharto M, Robinson DJ, Tan IB, Meewis C, Peters R, Spaniol SB, Stewart FA, Levendag PC, and Sterenborg HJCM. In vivo fluence rate measurements during Foscan-mediated photodynamic therapy of persistent and recurrent nasopharyngeal carcinomas using a dedicated light applicator. *J. Biomed. Opt*., 11:041107 (2006).
27. Flock ST, Wilson BC, and Patterson MS. Total attenuation coefficients and scattering phase functions of tissues and phantom materials at 633 nm. *Med. Phys.,* 14:835–841 (1987).
28. Twersky V. Absorptions and multiple scattering by biological suspensions. *J. Opt. Soc. Am.,* 60:1084–1093 (1970).

Chapter 4
Polarized Light: Electrodynamic Fundamentals

Jaro Rička and Martin Frenz

4.1 Introduction

In Chapters 4 and 7 we shall combine the photon-particle picture of light propagation and interaction in biological tissues with the electrodynamic model of light. In classical electrodynamics, light is regarded as a rapidly oscillating vector field. The term "light polarization" merely expresses the vector nature of the electromagnetic field. Two field vectors, **E** and **B**, represent the forces experienced by charged matter. Recall the well known expression $\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$ for the Lorentz force acting on a free point-like charged particle. In biological tissue the magnetic forces turn out to be negligible and thus we shall be mainly concerned with the electric field **E**. Because of the rapidity of light oscillations, and because the charges are bound in atoms and molecules, there is no net motion of the charges. The only effect of the imposed electric field is the deformation of the charge system.

Concerning the interaction of laser radiation with the tissue, we can distinguish two extreme regimes: At low laser intensities, corresponding to low field strength E, we are in the linear regime where the charge displacements are proportional to **E**. Of course some of the energy of the laser radiation may be absorbed, causing heating of the tissue, photochemical reactions or re-emission of luminescence. In the low intensity regime, however, the absorption of the laser radiation can be expected to have only a negligible effect on the optical properties of the tissue. This is the regime of optical diagnostics and certain minimally invasive therapeutic methods, such as photodynamic therapy. In this context we are mostly interested in the light propagation through the tissue; here specifically in the effect of the polarization on the light propagation. The other extreme is the highly non-linear regime of laser ablation, but in this regime polarization plays only a minor role. Thus, the contributions in Chapters 4 and 7 are aimed at the diagnostic applications of laser radiation. A topic of growing interest in optical tissue diagnostics, especially in dermatology,

J. Rička (✉)
Institute of Applied Physics, Sidlerstr. 5, 3012 Bern, Switzerland
e-mail: jaroslav.ricka@iap.unibe.ch

A.J. Welch, M.J.C. van Gemert (eds.), *Optical-Thermal Response of Laser-Irradiated Tissue*, 2nd ed., DOI 10.1007/978-90-481-8831-4_4,

is polarization imaging [1–3]. The first step in developing novel diagnostic methods is usually the comparison: one compares the Mueller matrix images of a healthy and diseased tissue, searching for a signature of the disease. In a second step one would like to understand how such images are formed in order to link the signature with changes of the micro-structure of the tissue that are associated with the disease. This is a formidable task, considering the complexity of the biological tissue as an optical medium. In Chapter 4 we compile concepts and tools that we deem to be necessary or convenient for understanding the propagation of polarized light and for polarization analysis. In Chapter 7 these tools will be put into operation for the development of the concepts for realistic optical tissue modeling and polarized Monte Carlo simulations.

4.2 Light as Electromagnetic Wave

4.2.1 Maxwell Equations

The topic of electromagnetism is the subject of numerous standard books (e.g. [4, 5]). We only need to review the matters in a language appropriate for the description of light propagation in biological tissue. (Thereby we shall employ SI-units for electromagnetic quantities, which include the vacuum permittivity $\epsilon_o\,[\mathrm{CV^{-1}m^{-1}}]$ and vacuum permeability $\mu_o\,[\mathrm{VsA^{-1}m^{-1}}]$. Recall that $\epsilon_o\mu_o = 1/c^2$, where c is the speed of light in vacuum.) The core of the electromagnetic theory is the celebrated set of Maxwell Equations (ME) relating the field vectors $\mathbf{E}(\mathbf{r},\ t)\,[\mathrm{Vm^{-1}}]$ and $\mathbf{B}(\mathbf{r}, t)\,[\mathrm{Vsm^{-1}}]$ with their sources, namely with the charge distribution $\rho(\mathbf{r},\ t)\,[\mathrm{C/m^3}]$ and current distribution $\mathbf{j}(\mathbf{r},\ t)\,[\mathrm{A/m^2}]$. Employing the SI units, the Maxwell Equations read:

$$\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0 \tag{4.1}$$

$$\frac{1}{\mu_o}\nabla \times \mathbf{B} - \epsilon_o \frac{\partial \mathbf{E}}{\partial t} = \mathbf{j} \tag{4.2}$$

$$\epsilon_o \nabla \cdot \mathbf{E} = \rho \tag{4.3}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{4.4}$$

Macroscopic currents and free charges play hardly any role in the interaction of the rapidly oscillating light field with condensed matter. Therefore, we only consider the microscopic contributions to ρ and $\mathbf{j}$, which are due to charges bound in atoms or molecules. These contributions originate chiefly from electric dipoles $\mathbf{d}_m$ and magnetic dipoles $\mathbf{m}_m$ that are induced in the molecules by the applied fields. Correspondingly, the charge and current distributions ρ and $\mathbf{j}$ reflect the distributions $\mathbf{P}$ and $\mathbf{M}$ of the molecular electric and magnetic dipoles (traditionally called "dielectric polarization" and "magnetization"). In a simple microscopic model one

would express these distributions in terms of Dirac's δ-functions representing the position of the molecules [6], as for example, $\mathbf{P}(\mathbf{r}) = \sum_m \mathbf{d}_m \delta(\mathbf{r} - \mathbf{r}_m)$. We shall need such a microscopic approach for the development of light scattering in tissues, but for now we neglect the molecular picture and write the bound charges and currents in the standard form [4, 5] as

$$\rho = -\nabla \cdot \mathbf{P} \qquad \mathbf{j} = \frac{\partial \mathbf{P}}{\partial t} + \nabla \times \mathbf{M} \tag{4.5}$$

Note that the equation system (Eqs. (4.1–4.5)) is not yet closed. Since $\mathbf{P}$ and $\mathbf{M}$ have been generated by the fields $\mathbf{E}$ and $\mathbf{B}$, we will have to supply the *constitutive relations* $\mathbf{P} = \mathbf{P}(\mathbf{E}, \mathbf{B})$ and $\mathbf{M} = \mathbf{M}(\mathbf{E}, \mathbf{B})$. To simplify the matter, we restrict ourselves to monochromatic laser radiation. This is legitimate because the output of most lasers is in a good approximation of monochromatic light, and because any time dependence can be synthesized as a superposition of appropriately phased monochromatic waves. Furthermore, we assume that the response of the matter is linear. Thus all field vectors and molecular dipole moments oscillate with the same angular frequency ω. For example

$$\mathbf{P}(\mathbf{r}, t) = \mathbf{P}(\mathbf{r})\, e^{-i\omega t} \quad \text{and} \quad \mathbf{E}(\mathbf{r}, t) = \mathbf{E}(\mathbf{r})\, e^{-i\omega t} \tag{4.6}$$

Note that the vectors are complex quantities, for example $\mathbf{E} = \Re\mathbf{E} + i\Im\mathbf{E}$. The complex notation simplifies the mathematics, because the differential operator $\partial / \partial t$ is replaced with multiplication by $-i\omega$. But this is not the only motivation for the complex notation. We intend to combine the photon picture with the electrodynamic model of light and this requires us to accept at least some of the concepts of the quantum theory. In the quantum world we deal quite naturally with complex wave functions, while the observables are real.

The observable in which we are most interested in the context of light propagation is the photon flux density N [m^{-2}] or the irradiance I [Wm^{-2}] (symbol I is used in the present chapter to avoid confusion with the electric field strength E). In classical electrodynamics, the energy flow is represented by the Poynting vector $\mathbf{S}(t) = [\Re\mathbf{E}(t) \times \Re\mathbf{B}(t)] / \mu_o$. However, since light fields exhibit extremely rapid oscillations, the only meaningful observable is the average $\mathbf{S} = \mathbf{S}(t)$ over at least a couple of oscillation cycles. In the complex notation this quantity is written as[1]:

$$\mathbf{S} = S\hat{\mathbf{s}} = \frac{1}{4}\frac{1}{\mu_o}[\mathbf{E} \times \mathbf{B}^* + \mathbf{E}^* \times \mathbf{B}] = \frac{1}{2}\frac{1}{\mu_o}\Re[\mathbf{E}^* \times \mathbf{B}] \tag{4.7}$$

[1] In most textbooks one finds the Poynting vector defined as $\mathbf{S}' = \frac{1}{2}\Re[\mathbf{E}^* \times \mathbf{H}]$ where $\mathbf{H}$ is the auxiliary fields defined in Eq. (4.9). The discussion whether $\mathbf{S}'$ or $\mathbf{S}$ is the correct expression for the electromagnetic Poynting vector in condensed matter is continuing [7, 8] and is likely to be intensified in the context of the so-called "left-handed metamaterials" [9]. Luckily, magnetization in biological tissues at optical frequencies is negligible, so that $\mathbf{S}' = \mathbf{S}$.

The power dP flowing through an oriented infinitesimal area $d\mathbf{a} = da\ \hat{\mathbf{n}}$ is given by $dP = \mathbf{S}d\mathbf{a} = S(\hat{\mathbf{s}} \cdot \hat{\mathbf{n}})da$, where $\hat{\mathbf{s}}$ is the unit vector pointing in the direction of $\mathbf{S}$, and $\hat{\mathbf{n}}$ is the unit vector specifying the orientation of the surface. To highlight the significance of the Poynting vector, we recall the Poynting's theorem [4, 5] for harmonic fields:

$$\nabla \cdot \mathbf{S} = -\frac{1}{2}\Re\left(\mathbf{j} \cdot \mathbf{E}^*\right) \tag{4.8}$$

Equation (4.8), which follows straightforwardly from the Maxwell Equations, expresses the conservation of electromagnetic energy: divergence of electromagnetic energy flux (left hand side) is due to the rate of exchange of electromagnetic energy with the matter, e.g. heating by absorption (right hand side). One should keep in mind that $\mathbf{E}$ and $\mathbf{B}$ are complicated microscopic fields that consist of the superposition of the field of the driving laser and of the fields whose sources are the oscillating charges in the matter. One can think of the photons as being engaged in nearly endless microscopic interaction loops, which, as we know, slow down their propagation.

4.2.2 Plane Waves in Homogeneous Media

The exact solution of the Maxwell Equations in a randomly heterogeneous medium such as a biological tissue is an impossible task. We are forced to make approximations. A useful zero order approximation is to represent tissue by a homogeneous medium, whose molecular constituents are smoothed by a suitable averaging procedure [6]. Then one regards the dipole distributions $\mathbf{P}$ and $\mathbf{M}$ as smooth vector fields like $\mathbf{E}$ and $\mathbf{B}$, which may be included in two auxiliary fields $\mathbf{D}$ and $\mathbf{H}$:

$$\mathbf{D} = \epsilon_o\mathbf{E} + \mathbf{P} \qquad \mathbf{H} = \frac{1}{\mu_o}\mathbf{B} - \mathbf{M} \tag{4.9}$$

Here $\mathbf{D}$ is the so-called "dielectric displacement," which incorporates the dielectric response of the matter. The magnetic response of the matter is incorporated in $\mathbf{H}$. However, typically, biological tissues do not exhibit a magnetic response at optical frequencies and so we will usually set $\mathbf{M} = 0$. (Magnetic dipoles will return through a back door of chiral, optically active media.)

The smoothed version of the Maxwell Equation is particularly convenient for an approximate yet self-consistent treatment of electromagnetic waves in condensed matter. Recalling from the elementary scalar optics the role of the harmonic plane waves, we make the following *ansatz*: space and time dependence of the field vectors are expressed by common oscillating phase factor $\exp(i\mathbf{k} \cdot \mathbf{r} - i\omega t)$, where ω is the oscillation frequency and $\mathbf{k}$ is the wave vector. In general, $\mathbf{k}$ is a complex vector, whose imaginary part indicates an attenuated or an evanescent wave. In a non-absorbing medium, $\mathbf{k}$ is a real quantity, the unit vector $\hat{\mathbf{k}}$ is perpendicular to the wave front and the magnitude $k = |\mathbf{k}|$ defines the phase velocity of light in

the given direction: $v = \omega/k$. The same phase factor also applies to the material response, as long as we remain in the linear regime. Thus we write, for example, $\mathbf{E}(\mathbf{r}, t) = \mathbf{E}\exp(i\mathbf{k}\cdot\mathbf{r} - i\omega t) = E\hat{\mathbf{e}}\exp(i\mathbf{k}\cdot\mathbf{r} - i\omega t)$, where $\mathbf{E}$ is a constant complex vector, E is the scalar amplitude and $\hat{\mathbf{e}}$ is a complex unit vector such that $\hat{\mathbf{e}}^* \cdot \hat{\mathbf{e}} = 1$. This vector $\hat{\mathbf{e}}$ specifies the polarization state of the wave, as shall be discussed in detail in Section 4.3. For now we only note that $\hat{\mathbf{e}}$ is a real unit vector in the case of a linearly polarized wave. Note that this $\mathbf{E}(\mathbf{r}, t)$ is not yet the solution of the Maxwell Equation (ME); it is only a convenient *ansatz*: in the $\mathbf{k},\omega$-representation the differential operators are transformed into simple vector operations $\partial/\partial t \rightarrow -i\omega$ and $\nabla \rightarrow i\mathbf{k}$, and the common phasor $\exp(i\mathbf{k}\cdot\mathbf{r} - i\omega t)$ can be dropped from all equations. This reduces the set of Maxwell Equations into an exercise in vector algebra:

$$\mathbf{k} \times \mathbf{E} = \omega\mathbf{B} \tag{4.10}$$

$$\mathbf{k} \times \mathbf{H} = \omega\mathbf{D} \tag{4.11}$$

$$\mathbf{k} \cdot \mathbf{D} = 0 \tag{4.12}$$

$$\mathbf{k} \cdot \mathbf{B} = 0 \tag{4.13}$$

We also recall Eq. (4.9), still assuming that the medium is non magnetic, i.e. $\mathbf{M} = 0$. Upon examining this set of equations, one discovers certain geometrical relationships, which are summarized in Fig. 4.1, assuming thereby that $\mathbf{k}$ is real: from Eqs. (4.12) and (4.13), it follows that $\mathbf{k}$ is normal to a plane spanned by $\mathbf{B}$

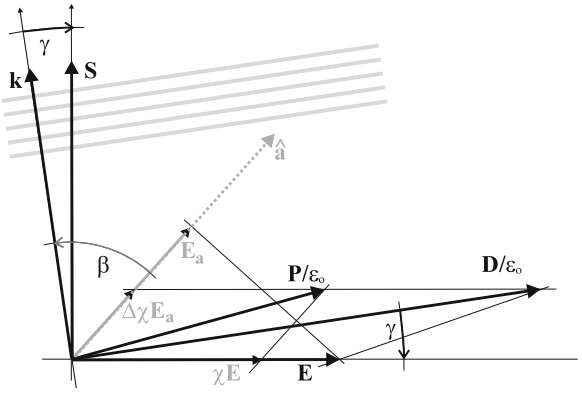


Fig. 4.1 Geometrical relationships in a harmonic plane. Magnetic vector **B** points into the paper. Note that the Poynting vector **S** and the wave vector **k** are not necessarily parallel. The light gray lines represent the wave fronts, which are *per definition* normal to **k**. They are parallel with **D**, but not necessarily with **E**. In an isotropic medium one would have $\mathbf{D} \parallel \mathbf{E}$ and, therefore, $\mathbf{S} \parallel \mathbf{k}$. (The outlined situation illustrates the propagation of an extraordinary wave in an anisotropic medium, $\hat{\mathbf{a}}$ is the anisotropy axis, gray-printed quantities are explained in Section 4.2.2.1.)

and **D**. In addition, according to Eq. (4.11), **D** is normal to a plane spanned by **k** and $\mathbf{H} = \mathbf{B}/\mu_o$ (which points into the paper). Thus, **B**, **D** and **k** form an orthogonal tripod. From Eq. (4.10) we deduce that **B** is normal to the plane **k**, **E**. Thus **B** and **E** are orthogonal and **E** must lay in the plane **k**, **D**. However the angle γ between **E** and $\mathbf{D} = \epsilon_o\mathbf{E} + \mathbf{P}$ is not yet specified, since it depends on the angle between **P** and **E**. Energy propagates in the direction of the Poynting vector (recall Eq. (4.7)), i.e. perpendicularly to **E** and **B**. Note that **S** and **k** are not necessarily parallel but they span the same angle γ as **E** and **D**. In any case, the vectors **E**, **D**, **P**, **k** and **S** lay in a plane to which **B** is normal. The same considerations apply to the real and imaginary parts of the complex field vectors.

After the exercise in geometry, all we have to do is to solve the Maxwell Equations for the unknown magnitude k of the wave vector and for the angle γ. First we eliminate $\mathbf{H} = \mathbf{B}/\mu_o$ from Eq. (4.11), thereby exploiting Eq. (4.10). Then using a well known vector identity $\mathbf{k} \times (\mathbf{k} \times \mathbf{E}) = \mathbf{k}(\mathbf{k} \cdot \mathbf{E}) - (\mathbf{k} \cdot \mathbf{k})\mathbf{E}$, we arrive at a vectorial wave equation in the **k**, ω-representation:

$$(\mathbf{k} \cdot \mathbf{k})\,\mathbf{E} - \mathbf{k}\,(\mathbf{k} \cdot \mathbf{E}) = k_o^2(\mathbf{E} + \mathbf{P}/\epsilon_o) \tag{4.14}$$

Here $k_o = \omega/c$ is the wave number in a vacuum. In order to solve Eq. (4.14), we must provide a constitutive relation $\mathbf{P} = \mathbf{P}(\mathbf{E})$ appropriate for the medium under consideration. The general solution is a little bit tricky and therefore let us first consider the familiar case of isotropic dielectrics:

$$\mathbf{P} = \epsilon_o \chi \mathbf{E} \qquad \mathbf{D} = \epsilon_o\,(1 + \chi)\,\mathbf{E} = \epsilon\epsilon_o\mathbf{E} \tag{4.15}$$

where χ is the *susceptibility* and $\epsilon = 1 + \chi$ is the *relative permittivity*. Combining the isotropic constitutive relation with Eq. (4.12) one finds that $\mathbf{k} \cdot \mathbf{E} = 0$. (In other words, **k** and **E** are perpendicular, which also implies that $\hat{\mathbf{s}} \parallel \mathbf{k}$. In isotropic media energy is transported along **k**.) Thus, Eq. (4.14) reduces to the identity

$$n^2\mathbf{E} = \epsilon\mathbf{E} \quad \text{where} \quad n^2 = \frac{\mathbf{k} \cdot \mathbf{k}}{k_o^2} \tag{4.16}$$

which is fulfilled if $n^2 = (\mathbf{k} \cdot \mathbf{k})/k_o^2 = \epsilon$. The right hand side of Eq. (4.16) is the general definition of *refractive index* n, applicable for real or complex **k**. (Note that if **k** is a complex vector then $\mathbf{k} \cdot \mathbf{k}$ is not the same as $|\mathbf{k}|^2 = \mathbf{k} \cdot \mathbf{k}^*$.) Assuming for simplicity that ϵ and **k** are real (no losses, no evanescence), we obtain the familiar result $\mathbf{k} \cdot \mathbf{k} = k^2 = n^2k_o^2$, where $n = \sqrt{\epsilon}$ is the refractive index. We leave it to the reader to verify that Eq. (4.16) is nothing but a disguise of the familiar wave equation $c_m^2\nabla^2\mathbf{E} = \partial^2\mathbf{E}/\partial t^2$, where $c_m^2 = c^2/\epsilon$ (recall that $\partial/\partial t \to -i\omega$ and $\nabla \to i\mathbf{k}$).

4.2.2.1 Anisotropic Media

Having rehearsed the familiar concepts, we tackle a more general case. Biological tissues are known to exhibit a certain degree of structural anisotropy. They contain elongated fiber-like macromolecules, and those fibers may be organized in oriented structures resulting in structural birefringence. A second, no less important, motivation is that birefringent optical elements are needed for measuring and manipulating the polarization of light. The anisotropy in structure results in the *anisotropy of dielectric response*[2] of the material to the applied field: along certain directions the charges respond to the electric force $\mathbf{E}$ with larger susceptibility than along others. We restrict ourselves to the most common case of uniaxial anisotropy: there is a single special direction, a single anisotropy axis $\hat{\mathbf{a}}$. The uniaxial constitutive relation reads

$$\mathbf{P}/\epsilon_o = \chi\mathbf{E} + \Delta\chi\hat{\mathbf{a}}\left(\hat{\mathbf{a}}\cdot\mathbf{E}\right) \tag{4.17}$$

Here χ is the overall isotropic contribution to the dielectric susceptibility and $\Delta\chi$ is the anisotropic susceptibility excess (positive or negative) in the direction $\hat{\mathbf{a}}$. The isotropic case is recovered by setting $\Delta\chi = 0$. Note that the vector $\mathbf{E_a} = \hat{\mathbf{a}}\left(\hat{\mathbf{a}}\cdot\mathbf{E}\right)$ represents the projection of the $\mathbf{E}$ vector on the anisotropy axis. Often Eq. (4.17) can be found in the short hand tensor notation $P = \epsilon_o\boldsymbol{\chi}\mathbf{E}$, where the components of $\boldsymbol{\chi}$ are $\chi_{kl} = \chi\delta_{kl} + \Delta\chi a_k a_l$. (Adding one more anisotropy axis we would naturally get the biaxial case, but adding still more would not bring anything new, because a 3D tensor has only 3 diagonal components. Note that in the uniaxial case the tensor $a_k a_l$ is a simple projection operator.) We prefer here the vector notation, since it makes it easier to understand the geometrical relationships. Further, we assume for simplicity a lossless medium ($\mathbf{k}$ real) and linear polarization ($\hat{\mathbf{e}}$ real). Inserting the Eq. (4.17) into (4.14) and dividing by the common factor $|\mathbf{k}|\,E$, the wave equations for light propagation in anisotropic media become

$$\hat{\mathbf{e}} - \hat{\mathbf{k}}\left(\hat{\mathbf{k}}\cdot\hat{\mathbf{e}}\right) = \frac{\epsilon}{n^2}\hat{\mathbf{e}} + \frac{\Delta\chi}{n^2}\hat{\mathbf{a}}\left(\hat{\mathbf{a}}\cdot\hat{\mathbf{e}}\right) \tag{4.18}$$

$$\left(\hat{\mathbf{k}}\cdot\hat{\mathbf{e}}\right) = -\frac{\Delta\chi}{\epsilon}\left(\hat{\mathbf{k}}\cdot\hat{\mathbf{a}}\right)\left(\hat{\mathbf{a}}\cdot\hat{\mathbf{e}}\right) \tag{4.19}$$

The second part follows from Eq. (4.12). The task can be formulated as follows: Suppose that in a medium with anisotropy axis $\hat{\mathbf{a}}$ we would like to excite a plane wave whose wave front propagates in the direction $\hat{\mathbf{k}}$, as indicated in Fig. 4.1. We know the inclination cosine $\hat{\mathbf{k}}\cdot\hat{\mathbf{a}} = \cos(\beta)$ and the medium parameters $\epsilon,\ \Delta\chi$. What are the permitted values of $|\mathbf{k}|^2 = n^2k_o^2$ and what polarization directions $\hat{\mathbf{e}}$ are allowed? Note that the questions can be answered quite easily if $\hat{\mathbf{k}}\cdot\hat{\mathbf{a}} = 1$ or if $\hat{\mathbf{k}}\cdot\hat{\mathbf{a}} = 0$. The letter special case, with optical axis $\hat{\mathbf{a}}$ oriented in right angle with $\hat{\mathbf{k}}$, is realized in polarizing optical elements such as retarders or a Glan Thompson prism.

[2] Anisotropy of dielectric response should not be confused with anisotropic scattering from large particles. There is no macroscopic scattering in homogeneous media.

However, in birefringent biological tissues light propagates in a random direction and therefore we need the general results. Upon examining the geometry of the problem, one finds that the sought information on $\hat{\mathbf{e}}$ is contained in two direction cosines $\hat{\mathbf{a}} \cdot \hat{\mathbf{e}}$ and $\hat{\mathbf{k}} \cdot \hat{\mathbf{e}}$. Thus, we convert the vector equation (4.18) into two scalar equations by multiplying both sides with $\hat{\mathbf{a}}$ and $\hat{\mathbf{e}}$, which gives a set of three equations for the three unknowns $\hat{\mathbf{a}} \cdot \hat{\mathbf{e}}$, $\hat{\mathbf{k}} \cdot \hat{\mathbf{e}}$ and n^2:

$$\begin{aligned} \left(\hat{\mathbf{a}} \cdot \hat{\mathbf{e}}\right) - \left(\hat{\mathbf{a}} \cdot \hat{\mathbf{k}}\right)\left(\hat{\mathbf{k}} \cdot \hat{\mathbf{e}}\right) &= \frac{\epsilon}{n^2}\left(\hat{\mathbf{a}} \cdot \hat{\mathbf{e}}\right) + \frac{\Delta\chi}{n^2}\left(\hat{\mathbf{a}} \cdot \hat{\mathbf{e}}\right) \\ 1 - \left(\hat{\mathbf{k}} \cdot \hat{\mathbf{e}}\right)^2 &= \frac{\epsilon}{n^2} + \frac{\Delta\chi}{n^2}\left(\hat{\mathbf{a}} \cdot \hat{\mathbf{e}}\right)^2 \\ \left(\hat{\mathbf{k}} \cdot \hat{\mathbf{e}}\right) &= -\frac{\Delta\chi}{\epsilon}\left(\hat{\mathbf{k}} \cdot \hat{\mathbf{a}}\right)\left(\hat{\mathbf{a}} \cdot \hat{\mathbf{e}}\right) \end{aligned} \tag{4.20}$$

An algebra program (Maple or Mathematica) suggests two solutions, which is not too surprising, since we are dealing with the phenomenon of birefringence. The so-called "ordinary" solution could have been already guessed when looking at Eq. (4.18):

$$\hat{\mathbf{a}} \cdot \hat{\mathbf{e}}_\perp = 0 \qquad \hat{\mathbf{k}} \cdot \hat{\mathbf{e}}_\perp = 0 \qquad n^2 = \epsilon_o \tag{4.21}$$

Apparently, the polarization vector $\hat{\mathbf{e}}_\perp$ is perpendicular to both, $\hat{\mathbf{a}}$ and $\hat{\mathbf{k}}$, and therefore perpendicular to the plane spanned by $\hat{\mathbf{a}}$ and $\hat{\mathbf{k}}$. The electric field oscillates perpendicularly to the propagation direction $\hat{\mathbf{k}}$, which then coincides with the direction $\hat{\mathbf{s}} = \mathbf{S}/|\mathbf{S}| = \hat{\mathbf{e}} \times \hat{\mathbf{b}}$ of energy propagation (recall Eq. (4.7)). Such a wave behaves just as an ordinary plane wave in an isotropic medium; therefore, it is called the "ordinary" wave. Because of $\hat{\mathbf{e}}_\perp \perp \hat{\mathbf{a}}$, the ordinary wave does not feel the anisotropy, i.e., it experiences the refractive index $n = \sqrt{\epsilon}$.

The second solution is the "extraordinary" wave, illustrated in Fig. 4.1:

$$\left(\hat{\mathbf{a}} \cdot \hat{\mathbf{e}}_\parallel\right)^2 = \frac{\sin(\beta)^2}{\sin(\beta)^2 + \cos(\beta)^2\,(\epsilon + \Delta\chi)^2/\epsilon^2} \qquad \left(\hat{\mathbf{k}} \cdot \hat{\mathbf{e}}_\parallel\right) = -\frac{\Delta\chi}{\epsilon}\cos(\beta)\left(\hat{\mathbf{a}} \cdot \hat{\mathbf{e}}_\parallel\right) \tag{4.22}$$

One verifies, that the projection of $\hat{\mathbf{e}}_\parallel$ on the *ak*-plane, namely the vector $\mathbf{e}_{ak} = \hat{\mathbf{k}}(\hat{\mathbf{k}} \cdot \hat{\mathbf{e}}_\parallel) + \hat{\mathbf{a}}(\hat{\mathbf{a}} \cdot \hat{\mathbf{e}}_\parallel)$, is a unit vector. Thus, $\hat{\mathbf{e}}_\parallel = \mathbf{e}_{ak}$ lays in the *ak*-plane. (Note that $\hat{\mathbf{e}}_\parallel$ and $\hat{\mathbf{e}}_\perp$ are orthogonal.) The result for the extraordinary refractive index is:

$$\frac{1}{n^2} = \frac{\cos(\beta)^2}{\epsilon} + \frac{\sin(\beta)^2}{\epsilon + \Delta\chi} \tag{4.23}$$

Note that Eq. (4.23) represents an ellipse. Refractive index experienced by the extraordinary wave depends on the direction $\hat{\mathbf{k}}$ of the wave propagation with respect to the anisotropy axis. In the special case $\hat{\mathbf{k}} \cdot \hat{\mathbf{a}} = \cos(\beta) = 1$ one recovers the ordinary wave with $\hat{\mathbf{a}} \cdot \hat{\mathbf{e}} = 0$ and $n = \sqrt{\epsilon}$. This is because the $\hat{\mathbf{a}}$-component of $\hat{\mathbf{e}}_\parallel$ vanishes and thus the wave does not feel the anisotropy. On the other hand, when

$\hat{\mathbf{k}} \perp \hat{\mathbf{a}}$, then the wave experiences refractive index $\sqrt{\epsilon + \Delta\chi}$. If $\Delta\chi > 0$ (positive birefringence), then wave propagation is slowed with respect to the ordinary wave. Thus, the "slow axis" of a retarder plate made from a positively birefringent material coincides with $\hat{\mathbf{a}}$ (see Section 4.3.2.1). The truly extraordinary feature of the extraordinary wave is the fact that the direction $\hat{\mathbf{s}} = \mathbf{S}/|\mathbf{S}|$ of energy propagation does not coincide with the direction $\hat{\mathbf{k}}$ of wave propagation. Evaluating the extraordinary Poynting vector (combine Eq. (4.7) with Eq. (4.10)), one obtains

$$\mathbf{S}_{ext} = |k| \left[\hat{\mathbf{k}} - \left(\hat{\mathbf{k}} \cdot \hat{\mathbf{e}}_{\parallel} \right) \hat{\mathbf{e}}_{\parallel} \right] \frac{1}{2} \epsilon_o |E|^2 c \tag{4.24}$$

from where it follows

$$\hat{\mathbf{s}} = \left[\hat{\mathbf{k}} - \left(\hat{\mathbf{k}} \cdot \hat{\mathbf{e}}_{\parallel} \right) \hat{\mathbf{e}}_{\parallel} \right] \Big/ \sqrt{1 - \left(\hat{\mathbf{k}} \cdot \hat{\mathbf{e}}_{\parallel} \right)^2} \qquad \hat{\mathbf{k}} \cdot \hat{\mathbf{s}} = \cos(\gamma) = \sqrt{1 - \left(\hat{\mathbf{k}} \cdot \hat{\mathbf{e}}_{\parallel} \right)^2} \tag{4.25}$$

Both the ordinary and the extraordinary waves are simultaneously excited when the anisotropic material is illuminated with a beam whose polarization is neither fully parallel nor fully perpendicular to the plane spanned by $\hat{\mathbf{a}}$ and $\mathbf{k}$. This splits the beam into two beams with orthogonal polarizations $\hat{\mathbf{e}}_{\parallel}$ and $\hat{\mathbf{e}}_{\perp}$: one observes birefringence. Because of the wave splitting, we were allowed to treat the ordinary wave $\mathbf{E}_{\perp}$ and the extraordinary wave $\mathbf{E}_{||}$ separately.

4.2.2.2 Chiral Media

Another special type of material response often found in biological tissues is chiral response, which manifests itself as optical activity, i.e. rotation of the polarization plane of linearly polarized light. A well known example is the optical activity of glucose solution, which suggests that the effect could be a non-invasive measurement of sugar levels in diabetes patients (see, however, [10]). The microscopic origins of optical activity are chiral, i.e. screw-like molecules (such as glucose), whose structure forces the charges to move in a screw-like fashion. Thus, the electric force generates in the medium microscopic coil magnets and we end up with electrically induced magnetic dipole density $\mathbf{M}_e$. Since $\mathbf{E}$ accelerates charges, we expect $\partial \mathbf{M}_e / \partial t \sim \mathbf{E}$, which leads to the chiral wave equation [5]:

$$\left(k^2 - k_o^2 \epsilon \right) \mathbf{E} = -ik\xi \, \hat{\mathbf{k}} \times \mathbf{E} \tag{4.26}$$

where $\epsilon = 1 + \chi$ and ξ are dimensionless material constants. For the solution we use the technique of educated guess. Realizing that the cross-product $\hat{\mathbf{k}} \times \mathbf{E}$ actually means rotation, we guess that circular polarization states could solve the problem. There are two orthogonal circular states, represented by the complex unit vectors $\hat{\mathbf{e}}_L = 1/\sqrt{2} \left(\hat{\mathbf{x}} - i\hat{\mathbf{y}} \right)$ and $\hat{\mathbf{e}}_R = 1/\sqrt{2} \left(\hat{\mathbf{x}} + i\hat{\mathbf{y}} \right)$. The subscripts L and R refer to left- and right-circularly polarized light waves. Inserting these states into Eq. (4.26) we obtain

$$\left(k^2 = k_o^2\epsilon\right)\ \hat{\mathbf{e}}_{L/R} = \pm k\xi\ \hat{\mathbf{e}}_{L/R} \tag{4.27}$$

where + is for $\hat{\mathbf{e}}_L$ and − for $\hat{\mathbf{e}}_R$. This gives two simple quadratic equations for k to be solved. Choosing the solutions such that $k > 0$ (one propagation direction of the wave), one obtains

$$k_{L/R} = \sqrt{k_o^2\epsilon + \xi^2/2} \pm \xi/2 \tag{4.28}$$

where again + is for $\hat{\mathbf{e}}_L$ and − for $\hat{\mathbf{e}}_R$. Thus, we can write the refractive index as $n_{L/R} = n \pm \delta n$, where $n = \sqrt{\epsilon + \xi^2/2k_o^2}$ and $\delta n = \xi/2k_o$. In other words, the two circular polarizations propagate with different phase velocities, undergoing a mutual phase shift during their propagation. One consequence of the refractive index difference is circular birefringence: a light beam passing through a prism made from chiral material would split into two beams with mutually orthogonal circular polarizations. Another consequence of the phase shift is the optical activity: after passing a linearly polarized beam through a chiral slab, one finds the polarization plane rotated in a certain angle with respect to the input.

4.2.3 Beyond Plane Waves

4.2.3.1 Beam-Waves and Their Properties

In the previous sections we pretended to be working with photon states that are plane waves with perfectly defined wave vector **k**, but we could not avoid talking or thinking in terms of "beams" or "rays." This is because there is a problem with the plane waves: they may be legal solutions of Maxwell equations in homogeneous media, but they are only idealized mathematical constructs. They can't exist in practice, because if the amplitude E of a plane wave is finite, then the plane wave transports infinite power. Plane wave photon states represent an extreme of the uncertainty relation: they are perfectly localized in the **k**-space, but they are completely delocalized in the real space. The probability of detecting a photon is the same at any position **r**. Consequently, plane waves are useless for imaging. Because polarization imaging is an increasingly important field of tissue optics, we must find better ways to represent photons. Before the wave nature of light was accepted, classical optics provided a sort of solution: geometrical rays. They are still a useful practical concept, but they are not a good representation of physical photon states: a ray is perfectly localized in transverse direction and it also has a well defined direction of propagation. In other words, a ray is perfectly localized both in **k** and **r**, thus violating the uncertainty relation in a most drastic way. Moreover, it is quite difficult to equip a ray with polarization and to include interference effects. Fortunately, there is a simple way out of the dilemma: there exist photon states which are fairly localized both in **r** and **k** (as needed for imaging), which resemble plane waves in many respects, and which occur quite naturally in every photonic laboratory. Those photon

states are beams. A monochromatic beam as a spatially coherent superposition of plane waves[3]:

$$\begin{aligned}\mathbf{E}(\mathbf{r}) &= \tfrac{E}{2\pi}\int_{-1}^{1}\int_{-1}^{1}\hat{\mathbf{e}}\,(\kappa_x,\,\kappa_y)\,B(\kappa_x,\,\kappa_y)\;e^{i\mathbf{k}\cdot\mathbf{r}}d\kappa_x d\kappa_y \\ \mathbf{B}(\mathbf{r}) &= \tfrac{1}{c}\tfrac{E}{2\pi}\int_{-1}^{1}\int_{-1}^{1}\hat{\mathbf{k}}\left(\kappa_x,\,\kappa_y\right)\;\times\;\hat{\mathbf{e}}\left(\kappa_x,\,\kappa_y\right)\,B\left(\kappa_x,\,\kappa_y\right)\;e^{-i\mathbf{k}\cdot\mathbf{r}}d\kappa_x d\kappa_y\end{aligned} \tag{4.29}$$

Here, $\hat{\mathbf{e}}$ is the polarization vector and $\hat{\mathbf{k}} = \mathbf{k}/k$ is the propagation unit vector with components κ_x, κ_y and κ_z such that $\kappa_x^2 + \kappa_y^2 + \kappa_z^2 = 1$. For simplicity we assume that the beam propagates along the z-axis and that the beam focus is at $\mathbf{r}_f = 0$. The beam profile is determined by the distribution $B(\kappa_x, \kappa_y)$ of the complex amplitudes of the partial waves. $B(\kappa_x, \kappa_y)$ is *quadratically* normalized:

$$\iint \left|B\left(\kappa_x,\,\kappa_y\right)\right|^2 d\kappa_x d\kappa_y \;=\; 1 \tag{4.30}$$

An important special case is a Gaussian beam, a good approximation for the fundamental mode of a laser resonator and the mode emanating from a single mode fiber:

$$B\left(\kappa_x,\,\kappa_y\right) \;=\; \frac{1}{\sqrt{\pi\sigma_k^2}}e^{-\left(\kappa_x^2+\kappa_y^2\right)/2\sigma_k^2} \tag{4.31}$$

The profile of a Gaussian beam is completely characterized by two numbers: the wave number k and the variance of the radial component of $\hat{\mathbf{k}}$, i.e., the average $\sigma_k^2 = \left\langle \kappa_x^2 + \kappa_y^2 \right\rangle = \iint (\kappa_x^2 + \kappa_y^2)\;\left|B\left(\kappa_x,\,\kappa_y\right)\right|^2 d\kappa_x d\kappa_y$. In cylindrical coordinates one has $\sigma_k^2 = \left\langle \sin\left(\beta_k\right)^2 \right\rangle$, where β_k is the angle of the inclination of the partial wave $\hat{\mathbf{k}}$ with respect to the beam axis z. Thus, we identify $\sigma_k = \sqrt{\left\langle \sin\left(\beta_k\right)^2 \right\rangle}$ as the *effective numerical aperture* of a beam (of any beam, not only Gaussian). Note that if $B(\kappa_x, \kappa_y)$ is sufficiently narrow, then $\sigma_k \approx \beta = \sqrt{\left\langle \beta_k^2 \right\rangle}$, where β is the divergence angle of the beam.

Unlike plane waves, beams with finite amplitude E carry finite power P

$$p \;=\; \oint_A \mathbf{S}\cdot d\mathbf{a} \tag{4.32}$$

where A is an arbitrary surface enclosing the source of the beam (i.e. the laser) and $\mathbf{a}$ is the outer normal to this surface. Since lasers usually radiate in one direction, A can be an arbitrary plane intersected by the beam. Upon inserting the Poynting vector that follows from Eq. (4.29) into Eq. (4.32) and using Eq. (4.30), one obtains:

[3] In a medium replace c with $c_m = c/n$ and ϵ_o with $\epsilon\epsilon_o$; k is the wave number in the medium.

$$P = \frac{c\epsilon_o}{2} E^2 \frac{1}{k^2} \langle \kappa_z \rangle \tag{4.33}$$

where $\langle \kappa_z \rangle = \langle \cos(\beta_k) \rangle = \iint |B(\kappa_x, \kappa_y)|^2 \kappa_z d\kappa_x d\kappa_y$ is the mean inclination cosine of the partial waves.

4.2.3.2 The Two Regimes: Quasi-Plane Waves and Quasi-Spherical Waves

A sketch of the wave structure of a typical moderately focused beam is illustrated in Fig. 4.2. One clearly observes two distinct regimes: in the asymptotic regime, far from the focus, the wave fronts are spherical, light propagates as a quasi-spherical wave. The asymptotic version of Eq. (4.29) is ([11] and references therein):

$$\begin{aligned} \mathbf{E}(\mathbf{r})|_{R\to\infty} &= E\,\mathbf{e}(s_x, s_y)\, B(s_x, s_y)\, s_z \frac{ie^{ikR}}{kR} \\ \mathbf{B}(\mathbf{r})|_{R\to\infty} &= \tfrac{1}{c}\hat{\mathbf{s}} \times \mathbf{E}(\mathbf{r})|_{R\to\infty} \end{aligned} \tag{4.34}$$

Here, $R = (x^2 + y^2 + z^2)^{1/2}$ is the distance from the beam focus at $\mathbf{r}_0 = 0$ and $\hat{\mathbf{s}}$ is the unit vector pointing to the direction of observation with the components $s_x = x/R$, $-s_y = y/R$ and $s_z = z/R$. The second part of Eq. (4.34) expresses the transversality of the field in the radiation zone, far from this focus. The asymptotic field of Eq. (4.34) exhibits a remarkable property: the amplitude of the spherical wave $\exp(ikR)/kR$ at a point $\mathbf{r} = \hat{\mathbf{s}}R$ in the far-field is determined by a single partial wave with amplitude $B(\kappa_x, \kappa_y)$ and polarization $\mathbf{e}(\kappa_x, \kappa_y)$, such that $\kappa_x = s_x = x/R$ and $\kappa_y = s_y = y/R$. In fact, each geometrical ray in the far field represents a partial wave. Hence, probing the field at some point far from the focus in the direction of propagation, one picks up a single partial plane wave, weighted with the factor $s_z = z/R$. On the other hand, one can go into the far field against the direction of propagation, before the focus, and do all sorts of beam engineering. For example, by strongly focusing a beam that is radially polarized in the far field, one can even create axial polarization in the focus [12], which may be useful for laser ablation.

In the present context we are more interested in the focal region of the beam, in the so-called Rayleigh range: in that waist-like region close to the focus of the beam,

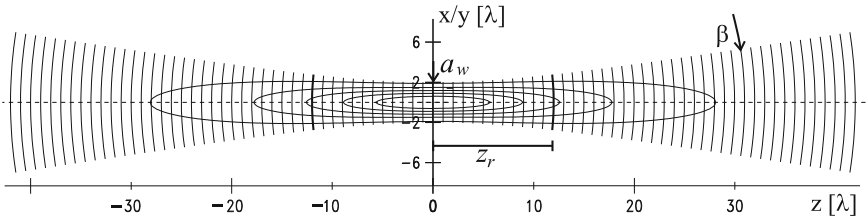


Fig. 4.2 Schematics of the wave structure of a moderately focused Gaussian beam. The contours indicate intensity levels. a_w is the $1/e$ waist radius, $z_r = ka_w^2 = 1/k\sigma_k^2$ is the size of the Rayleigh range. For paraxial beams $\sigma_k \approx \beta$, where β is the $1/e$ divergence angle

the wave fronts appear nearly flat, much like in a plane wave. Strictly speaking, the beam wave in the focal region is *not* a transverse radiation wave such that $\hat{\mathbf{s}} = \hat{\mathbf{e}} \times \hat{\mathbf{b}}$. It can't be, since it is a superposition of waves with varying polarization vectors $\hat{\mathbf{e}}$. However, for the present needs it is sufficient to consider only the so-called paraxial beams such that $\sigma_k^2 \ll 1$. This allows us to neglect the z-components of the partial polarization vectors and to approximate the z-components of the $\hat{\mathbf{k}}$-vectors of the partial waves as $k_z = \sqrt{1 - \left(\kappa_x^2 + \kappa_y^2\right)} \approx 1 - \left(k_x^2 + k_y^2\right)/2$. Furthermore, for the present needs it is sufficient to allow for two restrictions: (i) we send all partial waves through the same polarization filter. Because of the paraxial condition, we always get the same $\hat{\mathbf{e}}$ that does not depend on $\kappa_x,\ \kappa_y$. (ii) We restrict ourselves to the Rayleigh range (viz. Fig. 4.2), whose length is characterized by the so-called Rayleigh-parameter $z_r = 1/k\sigma_k^2$. Within the Rayleigh range, we can approximate the beam wave with a quasi-plane wave:

$$\mathbf{E}(x,\, y,\, z) = \tilde{B}(kx,\, ky) E\, \hat{\mathbf{e}}\, e^{ikz} \tag{4.35}$$

Where $\tilde{B}(kx, ky)$ is a dimensionless spatial profile:

$$\tilde{B}(kx,\, ky) = \frac{1}{2\pi} \iint B\left(\kappa_x,\, \kappa_y\right)\ e^{ik\kappa_x x + ik\kappa_y y} d\kappa_x d\kappa_y \tag{4.36}$$

According to Parselval's theorem, the distribution $\tilde{B}(kx, ky)$ is quadratically normalized such that $\iint \left|\tilde{B}(\kappa_x,\, \kappa_y)\right|^2 k^2\, dxdy = 1$. Since the k-vector distribution $B(\kappa_x,\ \kappa_y)$ and the transverse spatial profile $\tilde{B}(kx, ky)$ are a pair of Fourier transforms, we expect the following general behavior: the larger is the k-vector variance σ_k^2, the smaller is the beam cross-section $a_w^2 \iint \left(x^2 + y^2\right) \left|\tilde{B}(y,\, y)\right|^2 k^2\, dxdy$. For the special case of Gaussian k-distribution from Eq. (4.31) one obtains

$$\tilde{B}(kx,\, ky) = \frac{1}{\sqrt{\pi k^2 a_w^2}} e^{-\left(x^2 + y^2\right)/2a_w^2} \quad \text{where} \quad a_w^2 = \frac{1}{k^2\sigma_k^2} = \frac{z_r}{k} \tag{4.37}$$

In general one must expect $k\, a_w \sigma_k \leq 1$; the Gaussian beam, for which the equality applies, is a minimum uncertainty beam and therefore special.

In subsequent applications we shall find it convenient to express the electric field of the beam wave in terms of a measurable quantity, namely the beam power P [W]. Combining Eq. (4.33) with Eq. (4.35) and setting $\langle \kappa_z \rangle \approx 1 - \sigma_k^2/2 \approx 1$, we write the electric field of the beam wave as

$$\mathbf{E}(x, y, z) = \sqrt{P} \sqrt{\frac{2}{c\epsilon_o}}\, X(x,\, y)\, \hat{\mathbf{e}}\, e^{ikz} \tag{4.38}$$

where $X(x, y) = k\tilde{B}(x, y)$. A convenient model for the profile $X(x, y)$ is the Gaussian

$$X(x,\, y) \,=\, \exp\left(-\left[x^2 \,+\, y^2\right] / 2a_w^2\right) / \sqrt{\pi a_w^2} \tag{4.39}$$

The quantity πa_w^2 defines the effective cross-section of the beam. Thus Eq. (4.38) can be understood as

$$\mathbf{E}(x,\, y,\, z) \,=\, \sqrt{\frac{2}{c\epsilon_o}}\,\sqrt{I(x,\, y)}\;\hat{\mathbf{e}}\, e^{ikz} \,=\, \sqrt{\frac{2}{c\epsilon_o}}\,\sqrt{\bar{I}Y(x,\, y)}\;\hat{\mathbf{e}}\, e^{ikz} \tag{4.40}$$

Here $I(x,\ y) \,=\, P\,|X(x,y|^2$ is the irradiance [W/m^2]. On the right hand side we factored $I(x,y)$ into the mean irradiance $\bar{I} \,=\, P/\pi a_w^2$ and the dimensionless profile $Y(x,\, y) = \exp\left(-\left[x^2 \,+\, y^2\right]/a_w^2\right)$, which characterizes the intensity distribution through the beam cross-section. Recall, however, that sooner or later the quasi plane wave will transform into the quasi spherical wave, Eq. (4.34), which radiates into a solid angle $\Omega_w \,=\, 4\pi/k^2 a_w^2 \,=\, 4\pi\sigma_k^2$. The radiant intensity $\bar{I}_\Omega$ [W/sr] of the beam is $\bar{I}_\Omega \,=\, P/\Omega_w$. Note that for a Gaussian beam $\bar{I}_\Omega \,=\, \bar{I}\, z_r^2/4$ which is yet another manifestation of the uncertainty relation.

It is obvious that the Rayleigh range is a good place in a beam to measure and manipulate the polarization. One only has to make sure that z_r is sufficiently large as compared with the thickness of the polarizing elements. Recall that for a paraxial beam $z_r \,=\, ka_w^2 \,\approx\, 1/k\beta^2$, where β is the $1/e$ divergence angle of the beam. In a typical laser beam with β of a couple of millirads, z_r is many meters long. On the other hand, when the beam is focused with a microscope objective, z_r may shrink to only a fraction of the wavelength. The question arises, what happens with the polarization upon focusing the beam? Linear or circular polarization prepared in the Rayleigh range remains largely unaffected by focusing through a good lens (free of birefringence), up to a numerical aperture close to 1 [11].

4.2.3.3 Light Sources, Dipole Fields

There is still a problem with the beams, as defined mathematically in Eq. (4.29): they do not yet represent physical reality, because they propagate from nowhere to nowhere. They do not have a source, where the energy is pumped into the light field, and they do not have a sink, where the radiation energy is absorbed and re-stored in the constituents of matter. The basic constituents of condensed matter are atoms and molecules whose sizes are much smaller than the light wavelength. The only way the electromagnetic field can interact with molecules is to deform their charge systems into dipoles. Such dipolar interactions are involved in both scattering and absorption of light by molecules, but the effect of absorption on light propagation is rather trivial (a photon path is terminated). We are primarily interested in scattering: the impinging oscillating electric field induces an oscillating dipole in a target molecule which, in turn, re-radiates the electromagnetic energy in form of a spherical wave. As usual, we assume that the field-molecule interaction is linear and that

the impinging field is harmonic; all quantities oscillate with $\exp(-i\omega t)$. We set aside the rare case of chiral molecules and consider only the induced electric dipole:

$$\mathbf{d} = \mathbf{A}_m \cdot \mathbf{E}^{\mathbf{i}} \tag{4.41}$$

This is the molecular version of the constitutive relation in homogeneous continuous media, Eq. (4.17). The polarizability tensor $\mathbf{A}_m$ represents the susceptibility of the molecule to the applied electric field. For convenience we also define a normalized version of $\mathbf{A}_m$, namely $\mathbf{A} = \mathbf{A}_m/a$, where $a = \text{Trace}(\mathbf{A}_m)/3$ is the mean polarizability. With isotropic molecules, the tensor $\mathbf{A}$ reduces to the unit matrix $\mathbf{I}$.

The electromagnetic field generated by a dipole source is given by the Hertz-solution of Maxwell Equations, which represents the fundamental mode in the space of polarized spherical waves. For convenience we reproduce here the electric part of the scattered field. The scattering molecule is positioned in the origin of a coordinate system at $\mathbf{r}_m = 0$. The scattered field can be written as the sum of two terms: $\mathbf{E}^s(\mathbf{r}) = \mathbf{E}^s_{near}(\mathbf{r}) + \mathbf{E}^s_{rad}(\mathbf{r})$. The first term represents the near field:

$$\mathbf{E}^s_{\text{near}}(\mathbf{r}) = \frac{1}{4\pi\epsilon_o} k^3 \frac{e^{ikr}}{kr} \left\{ \frac{1}{k^2 r^2} - i\frac{1}{kr} \right\} \left[3\left(\mathbf{d}\cdot\hat{\mathbf{r}}\right)\hat{\mathbf{r}} - \mathbf{d} \right] \tag{4.42}$$

Note that the near field contributions decay rapidly with $1/(kr)^2$ and $1/(kr)^3$. Near fields do not represent propagating photons. However, they mediate energy exchange at close distances between the molecules and therefore play an important role in optical properties of condensed matter. The second term is the radiation field:

$$\mathbf{E}^s_{\text{rad}}(\mathbf{r}) = \frac{1}{4\pi\epsilon_o} k^3 \frac{e^{ikr}}{kr} \left[\mathbf{d} - \hat{\mathbf{r}}\left(\hat{\mathbf{r}}\cdot\mathbf{d}\right) \right] \tag{4.43}$$

The radiation field $\mathbf{E}^s_{rad}(\mathbf{r}, t)$ is a polarized spherical wave $\exp(-i\omega t + ikr)/kr$ whose polarization is given by the vector $\left[\mathbf{d} - \hat{\mathbf{r}}\left(\hat{\mathbf{r}}\cdot\mathbf{d}\right)\right] = \mathbf{d}_\perp$. Note that this $\mathbf{d}_\perp$ is the component of the induced dipole moment that is perpendicular to the director $\hat{\mathbf{r}}$. Thus, the radiation field is a transverse field, just like a plane wave. Only now we have a spherical wave that propagates radially in the direction $\hat{\mathbf{s}} = \hat{\mathbf{r}}$. The oscillating dipole acts as a source of photons. The radiation power, i.e. electromagnetic energy emitted in unit time, is

$$P_d = \frac{1}{4\pi\epsilon_o} \frac{c}{3} k^4 |d|^2 \tag{4.44}$$

Note here the k^4-dependence, characteristic for dipole radiation. Note also, that with the molecular dipole we introduced a fundamental inhomogeneity into the otherwise homogeneous medium.

4.3 Manipulating Polarization

> The type of optical system with which this section deals is not the more familiar type involving lenses, prisms, etc., but is rather the type which is composed of retardation plates, partial polarizers, and plates possessing the ability to rotate the plane of polarization. We shall therefore be concerned not with the directions of rays of light, but with the state of polarization and the intensity of the light as it passes through the optical system. [13]

4.3.1 Linear and Circular Polarization

In Section 4.2.2 we made an important observation: plane waves and quasi-plane waves are transversal waves: the field vector $\mathbf{E}(\mathbf{r},t) = E\,\hat{\mathbf{e}}\exp(ik\cdot\mathbf{r}-i\omega t)$ oscillates in a plane that is perpendicular to the propagation direction $\hat{\mathbf{s}}$. One can always rotate the coordinate system so that $\hat{\mathbf{z}}$-axis is oriented along the propagation direction $\hat{\mathbf{s}}$ and therefore $\hat{\mathbf{e}} = e_x\hat{\mathbf{x}} + e_y\hat{\mathbf{y}} + 0\cdot\hat{\mathbf{z}} = e_x\hat{\mathbf{x}} + e_y\hat{\mathbf{y}}$. In other words, the polarization vector $\hat{\mathbf{e}}$ can be expressed as a complex linear combination of the two orthogonal unit vectors $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$. These two real unit vectors form the so-called linearly polarized base. Any linear combination with real coefficients, such as for example $\hat{\mathbf{e}} = \cos(\psi)\hat{\mathbf{x}} + \sin(\psi)\hat{\mathbf{y}}$, represents linearly polarized light. The classical observable quantity is $\Re\left[\hat{\mathbf{e}}\exp(ik_sz - i\omega t)\right] = \hat{\mathbf{e}}\cos(k_sz-\omega t) = \left[\cos(\psi)\,\hat{\mathbf{x}} + \sin(\psi)\,\hat{\mathbf{y}}\right]\cos(k_sz-\omega t)$. The real field vector oscillates linearly in a plane rotated by ψ with respect to the xz-plane, as illustrated in Fig. 4.3a.

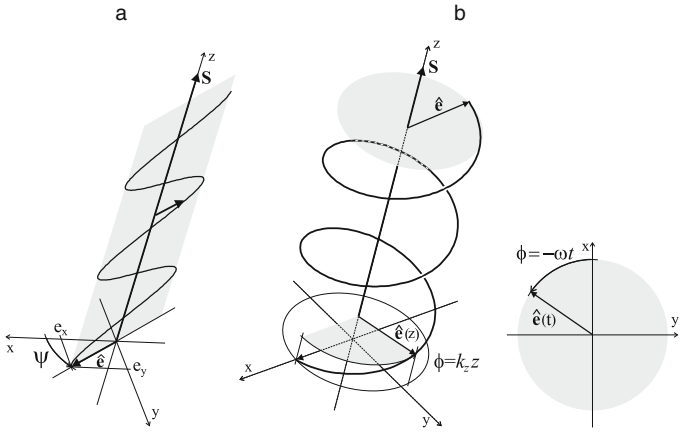


Fig. 4.3 (**a**) Linearly polarized wave. Plotted is the real part of the polarization vector as the function of z at $t = 0$: $\Re\left[\hat{\mathbf{e}}\exp(ik_sz)\right] = \left[\cos(\psi)\,\hat{\mathbf{x}} + \sin(\psi)\,\hat{\mathbf{y}}\right]\cos(-k_sz)$. (**b**) Right circularly polarized wave. The right handed spiral is the real part of the polarization vector: $\Re\left[\hat{\mathbf{e}}_R\exp(ik_sz)\right] = \frac{1}{\sqrt{2}}\cos(-k_sz)\,\hat{\mathbf{x}} - \frac{1}{\sqrt{2}}\sin(-k_sz)\,\hat{\mathbf{y}}$. As the spiral is pushed through the xy-plane (without rotation), the arrow $\Re\left[\hat{\mathbf{e}}_R\exp(-i\omega t)\right] = \frac{1}{\sqrt{2}}\cos(\omega t)\,\hat{\mathbf{x}} - \frac{1}{\sqrt{2}}\sin(\omega t)\,\hat{\mathbf{y}}$ rotates counter-clockwise. Note that the x and y axes appear reversed because we view the rotation in the direction of wave propagation

Combining $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ with complex coefficients, one obtains the general case of elliptically polarized light. Two particularly important combinations are $\hat{\mathbf{e}}_L = \hat{\mathbf{x}}/\sqrt{2} + i\hat{\mathbf{y}}\sqrt{2}$ and $\hat{\mathbf{e}}_R = \hat{\mathbf{x}}/\sqrt{2} - i\hat{\mathbf{y}}\sqrt{2}$. Note that $i = \exp(i\pi/2)$; there is a phase shift $\pm\pi/2$ between the two linearly polarized components:

$$\Re\left[\hat{\mathbf{e}}_{L/R}\exp(ik_s z - i\omega t)\right] = \frac{1}{\sqrt{2}}\cos(\omega t - k_s z)\,\hat{\mathbf{x}} \pm \frac{1}{\sqrt{2}}\sin(\omega t - k_s z)\,\hat{\mathbf{y}} \tag{4.45}$$

where, again, + is for $\hat{\mathbf{e}}_L$ and − for $\hat{\mathbf{e}}_R$. The geometry of the right handed polarization $\hat{\mathbf{e}}_R$ is illustrated in Fig. 4.3b. When frozen in time, the tips of the real part of the polarization vector along a ray form a right handed (continental) screw. The right part of Fig. 4.3b illustrates the time rotation of the tip, as viewed in the direction of propagation: the vector rotates counter-clockwise, which is rather confusing. Circularly polarized states are fundamental in the photon picture of light: a photon in the left-handed state $\hat{\mathbf{e}}_L \equiv |\sigma^+\rangle$ has its spin oriented in the direction of propagation $\hat{\mathbf{s}}$, whereas $\hat{\mathbf{e}}_R \equiv |\sigma^-\rangle$ is a photon state with negative spin. These two complex vectors $\hat{\mathbf{e}}_L$ and $\hat{\mathbf{e}}_R$ are orthogonal. They form the so-called circularly polarized base, but this choice of base is slightly too abstract for the needs of an experimentalist. We shall keep working with the linear base $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$.

4.3.2 Basic Hardware Elements

4.3.2.1 Retarder Plate

The basic optical element for manipulating polarization is a retarder plate, a transparent plan-parallel plate made from birefringent material. The anisotropy axis $\hat{\mathbf{a}}$ is oriented in parallel with the faces of the plate and the plate is oriented normally to the $\hat{\mathbf{k}}$ of the impinging wave so that $\hat{\mathbf{k}} \perp \hat{\mathbf{a}}$. Consequently, the ordinary and the extraordinary waves propagate in the same direction $\hat{\mathbf{s}} \parallel \hat{\mathbf{k}}$ but with different phase velocities c/n_{ord} and c/n_{ext}. Upon traversing the plate the two waves will accumulate a mutual phase shift which may turn linear polarization into elliptical, or vice versa. Figure 4.4 explains the function of the retarder in more detail and defines the coordinate system used in the subsequent formal analysis. In the complex vector notation the action of the retarder can be expressed by the following equation:

$$\hat{\mathbf{e}}(z_{\text{out}}) = e^{ik_x d}\hat{\mathbf{x}}\left[\hat{\mathbf{x}}\cdot\hat{\mathbf{e}}(z_{in})\right] + e^{ik_y d}\hat{\mathbf{y}}\left[\hat{\mathbf{y}}\cdot\hat{\mathbf{e}}(z_{\text{in}})\right] \tag{4.46}$$

Note here the two projection operators $\hat{\mathbf{x}}\hat{\mathbf{x}}\cdot$ and $\hat{\mathbf{y}}\hat{\mathbf{y}}\cdot$ that project the input state $\hat{\mathbf{e}}(z_{\text{in}})$ on the states $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$, respectively. The phasors $\exp(ik_x d)$ and $\exp(ik, d)$ account then for the propagation over a distance $d = z_{\text{out}} - z_{\text{in}}$ with two different phase velocities v_x and v_y. There is no power lost in the plate and so the state $\hat{\mathbf{e}}$ remains a normalized unit vector after propagating through the plate. A particularly

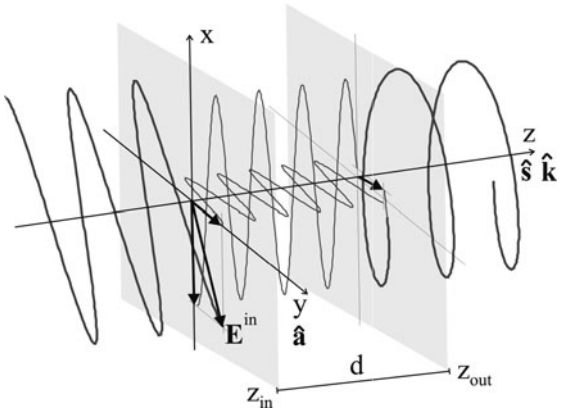


Fig. 4.4 Retarder plate. The anisotropy axis $\hat{\mathbf{a}}$ is oriented in parallel with the faces of the plate, in our case along the $\hat{\mathbf{y}}$-direction. Upon entering the plate, the linearly polarized input wave $\mathbf{E}^{\text{in}}$ splits into the ordinary and extraordinary wave, but both beams propagate in the same direction $\hat{\mathbf{s}}$, because $\hat{\mathbf{k}}$ is normal to the interface and to $\hat{\mathbf{a}}$ (Eqs. (4.22) and (4.25)). The ordinary wave oscillates along $\hat{\mathbf{x}}$ and propagates with phase velocity c/n_x, where $n_x = \sqrt{\epsilon}$. The extraordinary wave oscillates along $\hat{\mathbf{y}}$ and propagates with phase velocity c/n_y, where $n_y = \sqrt{\epsilon + \Delta\chi}$. (In our example, $\hat{\mathbf{a}}$ is a slow axis, such as in a quartz retarder.) Upon passing the thickness d, the two waves accumulated a phase difference $\delta = d(k_y - k_x)$, so that they recombine into an elliptically polarized state, in our example nearly $\hat{\mathbf{e}}_R$. Recall from electrodynamics textbooks that there are certain boundary conditions to be met at each interface. Because $\hat{\mathbf{k}}$ is normal to the interface and to $\hat{\mathbf{a}}$, only two boundary conditions are sufficient in our case: (1) phase matching $\cos(k_{\text{in}} z_i) = \cos(k_x z_i + \phi_x) = \cos(k_y z_i + \phi_y)$; and (2) tangential components of the **E**-vector are continuous: $\mathbf{E}^{\text{in}} = E_x\hat{\mathbf{x}} + E_y\hat{\mathbf{y}}$

important special case of a retarder is the λ/4-plate, whose thickness d is carefully adjusted so that $(k_y - k_x)d = \pm\pi/2$. The special retarder equation reads

$$\hat{\mathbf{e}}\,(z_{\text{out}}) \;=\; e^{ik_x d}\left\{\hat{\mathbf{x}}\left[\hat{\mathbf{x}}\cdot\hat{\mathbf{e}}\,(z_{\text{in}})\right] \;\pm\; i\hat{\mathbf{y}}\left[\hat{\mathbf{y}}\cdot\hat{\mathbf{e}}\,(z_{\text{in}})\right]\right\} \tag{4.47}$$

This looks much like circular polarization. Indeed, upon inserting a linearly polarized input state $\hat{\mathbf{e}}(z_{\text{in}}) = \cos(\phi)\,\hat{\mathbf{x}} + \sin(\phi)\,\hat{\mathbf{y}}$ (and neglecting the common phasor $\exp(ik_x d)$), one obtains $\hat{\mathbf{e}}(z_{\text{out}}) = \cos(\phi)\,\hat{\mathbf{x}} + i\sin(\phi)\,\hat{\mathbf{y}}$. This is elliptical polarization. To obtain a perfectly circular state, one must turn the anisotropy axis of the λ/4-plate in 45° with respect to input polarization, so that $\cos(\phi) = 1/\sqrt{2}$ and $\sin(\phi) = \pm 1/\sqrt{2}$. Turning the plate means a rotation of the coordinate system. We define the rotated coordinate system $\hat{\mathbf{x}}'$, $\hat{\mathbf{y}}'$ through

$$\begin{aligned}\hat{\mathbf{x}} &= \cos(\varphi)\,\hat{\mathbf{x}}' - \sin(\varphi)\,\hat{\mathbf{y}}'\\ \hat{\mathbf{y}} &= \sin(\varphi)\,\hat{\mathbf{x}}' + \cos(\varphi)\,\hat{\mathbf{y}}'\end{aligned} \tag{4.48}$$

The angle ϕ is measured clockwise when looking along the direction of propagation. In the rotated coordinate system the anisotropy axis $\hat{\mathbf{a}}$ is still oriented along $\hat{\mathbf{y}}'$ so that

$$\hat{\mathbf{e}}'\left(z_{\text{out}}\right) = e^{ik_x d}\left\{\hat{\mathbf{x}}'\left[\hat{\mathbf{x}}'\cdot\hat{\mathbf{e}}'\left(z_{\text{in}}\right)\right] \pm i\hat{\mathbf{y}}'\left[\hat{\mathbf{y}}'\cdot\hat{\mathbf{e}}'\left(z_{\text{in}}\right)\right]\right\} \tag{4.49}$$

However, $\hat{\mathbf{e}}'(z_{in})$ must be expressed in the new coordinate system:

$$\hat{\mathbf{e}}'\left(z_{\text{in}}\right) = \cos\left(\psi\right)\left[\cos\left(\phi\right)\hat{\mathbf{x}}' - \sin\left(\phi\right)\hat{\mathbf{y}}'\right] + \sin\left(\psi\right)\left[\sin\left(\phi\right)\hat{\mathbf{x}}' + \cos\left(\phi\right)\hat{\mathbf{y}}'\right] \tag{4.50}$$

It is easy to verify that circular polarization is recovered when $\psi - \phi = \pm\pi/4$. Finally, one would have to transform $\hat{\mathbf{e}}'(z_{in})$ back into the laboratory coordinate system, but vector algebra is a somewhat clumsy tool for such calculations.

4.3.2.2 Linear Polarizer

A longer look at the retarder equation, Eq. (4.46), suggests how to make a polarization filter: if one could equip the plate with some device which blocks off one of the terms (say the extraordinary term, $\hat{\mathbf{y}}$), then the device would simply project the incoming state $\hat{\mathbf{e}}'(z_{in})$ on the linear state $\hat{\mathbf{x}}$:

$$\mathbf{e}(z_{out}) = e^{ik_x d}\hat{\mathbf{x}}\left[\hat{\mathbf{x}}\cdot\hat{\mathbf{e}}\left(z_{\text{in}}\right)\right] \tag{4.51}$$

In other words, no matter what $\hat{\mathbf{e}}'(z_{in})$ is, the output would always be the linearly polarized state $\hat{\mathbf{x}}$ multiplied with the complex amplitude $\left[\hat{\mathbf{x}}\cdot\hat{\mathbf{e}}\left(z_{\text{in}}\right)\right] = e_x$. Of course the output of a filter is no longer a unit vector, since $\mathbf{e}^*(z_{out})\cdot\mathbf{e}(z_{out}) = |e_x(z_{in})|^2 \neq 1$; we lost some of the beam power in the filter.

One way to produce a linear polarizer is a Glan Thompson prism shown in Fig. 4.5. The retarder plate (made from calcite) is cut into two triangular prisms and glued again together with a cement whose refractive index matches the smaller of $n_{\text{ext}} = 1.486$ and $n_{\text{ord}} = 1.658$. Thus the extraordinary ray is perfectly transmitted. The ordinary ray, however, undergoes total internal reflexion on the calcite-cement-interface, if the cut angle is appropriate. Another way is to make either k_x or k_y in Eq. (4.46) imaginary (say k_x), so that the amplitude of the chosen component decays exponentially with $\exp(-|k_x|\,d)$. A well known example of this principle is a Polaroid sheet made of strongly anisotropic absorbing material. Because of the strong anisotropy, the material exhibits strong linear dichroism, i.e. the extraordinary wave is much more absorbed than the ordinary wave. Modern sheet polarizers are practically as perfect as the Glan-Thompson prism. However, because of the

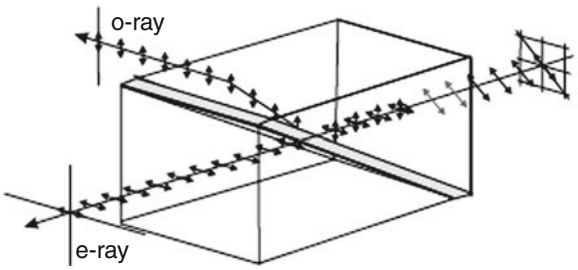


Fig. 4.5 Glan-Thompson polarizer: e-ray stands for extraordinary ray; o-ray for ordinary ray

strong absorption, they can't be used in a laser beam. On the other hand, the sheet form is perfect for polarization imaging.

By combining linear polarizers with retarder plates, one achieves the full control of the polarization state. In particular, one can produce and calculate filters for circularly polarized states. However, as already mentioned, the vector notation is not well suited for such calculations. They can be best done using the elegant tools to be discussed in the next section.

4.3.3 Jones Formalism

In 1940, R. Clark Jones, a researcher with Polaroid Corporation and a Harvard scholar, got annoyed with vector calculus, which prompted him to design an efficient formalism for the analysis of light propagation through optical elements such as, for example, sheets of Polaroid foil [13]. Jones represents optical elements as matrix operators **O**, which act on the polarization states that are represented by 2D complex column vectors **J** (one column matrices, known today as the Jones vectors): $\mathbf{J}_{\text{out}} = \mathbf{O} \cdot \mathbf{J}_{\text{in}}$. There is no doubt that Jones was inspired by the matrix formulation of quantum mechanics. Thus, we shall introduce the Jones formalism using Dirac's bra-ket notation, $\langle bra|$ and $|ket\rangle$. For now, this is useful in order to distinguish the Jones vectors from 3D field vectors such as **E**. Later we shall fill the bra-kets with a more general content, without having to change the notation. Using the linearly polarized base, the Jones vector $|E\rangle$ that corresponds to the field vector **E** reads

$$\mathbf{E} \equiv |E\rangle = \begin{pmatrix} E_x \\ E_y \end{pmatrix} \tag{4.52}$$

Here E_x and E_y are the complex components of the **E**-vector in a local coordinate system whose $\hat{\mathbf{z}}$-axis coincides with the propagation direction $\hat{\mathbf{s}}$. Each ket has a companion bra, a complex conjugate and transpose of the ket:

$$\mathbf{E}^* \equiv \langle E| = \begin{pmatrix} E_x^* & E_y^* \end{pmatrix} \tag{4.53}$$

In the bra-ket-notation, the scalar product of two complex vectors $|a\rangle$ and $|b\rangle$ reads

$$\mathbf{a}^* \cdot \mathbf{b} \equiv \langle a|\, b\rangle = \begin{pmatrix} a_x^* & a_y^* \end{pmatrix} \begin{pmatrix} b_x \\ b_y \end{pmatrix} = a_x^* b_x + a_y^* b_y \tag{4.54}$$

For now, $\langle a|b\rangle$ is merely an abbreviation for $a_x^* b_x + a_y^* b_y$, but later we shall employ a more general definition of the scalar product. A special case of a scalar product is the norm $\langle a|a\rangle \equiv |\mathbf{a}|^2$. A *unit* Jones vector $|e\rangle$ is normalized such that $\langle e|e\rangle = 1$, just like an ordinary 3D unit vector in Euclidean space. Recall that unit vectors are dimensionless. For now, a complex unit Jones vector $|e\rangle$ represents a

polarization state, equivalent to a polarization unit vector $\hat{\mathbf{e}}$: $\mathbf{E} = E\hat{\mathbf{e}} \equiv |E\rangle = E\,|e\rangle$. The following three pairs of polarization states are particularly important:

$$\begin{aligned} |X\rangle &:= \begin{pmatrix} 1 \\ 0 \end{pmatrix} & |Y\rangle &:= \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ |/\rangle &:= \tfrac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} & |\backslash\rangle &:= \tfrac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\ |R\rangle &\equiv \left|\sigma^-\right\rangle := \tfrac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix} & |L\rangle &\equiv \left|\sigma^+\right\rangle := \tfrac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} \end{aligned} \tag{4.55}$$

The states $|X\rangle, |Y\rangle$ and $|/\rangle, |\backslash\rangle$ are linearly polarized; both pairs are orthogonal and form two linearly polarized bases: any Jones vector $|E\rangle$ can be expressed as

$$|E\rangle = E_x \left|X\right\rangle + E_y\,|Y\rangle \tag{4.56}$$

where E_x and E_y are complex scalars. The states $|/\rangle, |\backslash\rangle$ are easily generated by turning the polarizer $\pm 45°$ with respect to x- and y-axes. At first glance it appears superfluous to have two linear bases, but the significance of the turned base will emerge later. The states $|L\rangle, |R\rangle$ form the circularly polarized base.

4.3.4 Optical Elements in Jones Language

An optical element is a Jones matrix $\mathbf{O}$ that acts on an input state $\left|E^{\text{in}}\right\rangle$ and produces an output state $\left|E^{\text{out}}\right\rangle$ according to $\left|E^{\text{out}}\right\rangle = \mathbf{O}\left|E^{\text{in}}\right\rangle$. The general structure of a Jones matrix in the x,y-base is

$$\mathbf{O} = \begin{pmatrix} O_{xx} & O_{xy} \\ O_{yx} & O_{yy} \end{pmatrix} \tag{4.57}$$

The numbering of the matrix elements is such that, e.g., O_{xy} means the contribution of the y-component of the input state to the x-component of the output state:

$$\begin{aligned} E_x^{\text{out}} &= O_{xx}E_x^{\text{in}} + O_{xy}E_y^{\text{in}} \\ E_y^{\text{out}} &= O_{yx}E_x^{\text{in}} + O_{yy}E_y^{\text{in}} \end{aligned} \tag{4.58}$$

4.3.4.1 Projecting Filters

The central concept in the present approach is the projection operator. A projector is constructed by multiplying a column vector from the right by a row vector, which results in a hermitian 2×2 matrix:

$$\hat{\mathbf{f}}\hat{\mathbf{f}}^{*} \equiv |f\rangle\ \langle f| := \begin{pmatrix} f_x \\ f_y \end{pmatrix} \begin{pmatrix} f_x^* & f_y^* \end{pmatrix} = \begin{pmatrix} f_x f_x^* & f_x f_y^* \\ f_y f_x^* & f_y f_y^* \end{pmatrix} \tag{4.59}$$

Such matrices represent polarization filters. Note that a filter state is normalized such that $\langle f|f\rangle = 1$. In optics the projection operator is known under the name "Jones matrix of the filter". Three particularly important pairs of filter matrices are readily constructed from Eq. (4.55):

$$\mathbf{F}_X = |X\rangle\ \langle X| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad \mathbf{F}_Y = |Y\rangle\ \langle Y| = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \tag{4.60}$$

$$\mathbf{F}_/ = |/\rangle\ \langle /| = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \qquad \mathbf{F}_\backslash = |\backslash\rangle\ \langle \backslash| = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \tag{4.61}$$

$$\mathbf{F}_R = |R\rangle\ \langle R| = \frac{1}{2}\begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix} \quad \mathbf{F}_L = |L\rangle\ \langle L| = \frac{1}{2}\begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix} \tag{4.62}$$

$\mathbf{F}_X$, $\mathbf{F}_Y$, $\mathbf{F}_/$ and $\mathbf{F}_\backslash$ represent linear polarizers. Thanks to Polaroid, a linear polarizer is a common device, but circular polarizers $\mathbf{F}_L$ and $\mathbf{F}_R$ are more complicated.

4.3.4.2 Transformers

To transform the polarization from one state to the other, one needs two construction elements: (1) a retarder plate of variable thickness (either fast or slow anisotropy axis will do) to generate a phase shift between the ordinary and extraordinary beams; (2) a rotational stage to transfer the optical power between the two orthogonal states. The Jones matrix of a retarder is readily obtained by translating Eq. (4.46) into the Jones language. The result is a linear combination of two projectors:

$$|e(z_{out})\rangle = e^{ik_x d}\mathbf{F}_x|e(z_{in})\rangle + e^{ik_y d}\mathbf{F}_y|e(z_{in})\rangle = \mathbf{T}_{Ret}\,|e(z_{in})\rangle \tag{4.63}$$

Omitting a common phasor $\exp(ik_x d)$, the retarder matrix reads:

$$\mathbf{T}_{Ret} = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\Delta k d} \end{pmatrix} \quad \text{where} \quad \Delta k = k_y - k_x \tag{4.64}$$

Two kinds of retarders are readily available: λ/4-plates, whose thickness is accurately adjusted so that $\Delta k d = \pm\pi/2$ and λ/2-plates, whose thickness is such that $\Delta k d = \pm\pi$. The sign is + for a slow anisotropy axis ($\Delta\chi > 0$). Assuming that $\hat{\mathbf{a}} \parallel \hat{\mathbf{y}}$, Jones matrices of these two retarders read:

$$\mathbf{T}_{\lambda/4} = \begin{pmatrix} 1 & 0 \\ 0 & \pm i \end{pmatrix} \qquad \mathbf{T}_{\lambda/2} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{4.65}$$

The λ/4-plates are used for transforming the polarization from linear to circular (and vice versa), and λ/2-plates are needed to rotate the plane of linear polarization. The last element we need is a tool to rotate the retarder plates or linear polarizers:

$$\mathbf{R_2}(\phi) = \begin{pmatrix} \cos(\phi) & +\sin(\phi) \\ -\sin(\phi) & \cos(\phi) \end{pmatrix} \qquad \mathbf{R_2^{-1}}(\phi) = \begin{pmatrix} \cos(\phi) & -\sin(\phi) \\ +\sin(\phi) & \cos(\phi) \end{pmatrix} \tag{4.66}$$

The angle ϕ is measured clockwise, when viewing in the direction of propagation. For future algorithmic needs we also recall that the rotations can be expressed in terms of the direction cosines:

$$\mathbf{R_2}(\phi) = \begin{pmatrix} \hat{\mathbf{x}}' \cdot \hat{\mathbf{x}} & \hat{\mathbf{x}}' \cdot \hat{\mathbf{y}} \\ \hat{\mathbf{y}}' \cdot \hat{\mathbf{x}} & \hat{\mathbf{y}}' \cdot \hat{\mathbf{y}} \end{pmatrix} \qquad \mathbf{R_2^{-1}}(\phi) = \begin{pmatrix} \hat{\mathbf{x}} \cdot \hat{\mathbf{x}}' & \hat{\mathbf{x}} \cdot \hat{\mathbf{y}}' \\ \hat{\mathbf{y}} \cdot \hat{\mathbf{x}}' & \hat{\mathbf{y}} \cdot \hat{\mathbf{y}}' \end{pmatrix} \tag{4.67}$$

When applied to a Jones vector whose components are given in laboratory coordinates $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, the operation $|J'\rangle = \mathbf{R}_2 |J\rangle$ expresses this vector in a local system $\hat{\mathbf{x}}'$, $\hat{\mathbf{y}}'$, that is rotated by ϕ around the z-axis (rotation is clockwise when viewed in the direction of propagation). The inverse matrix $\mathbf{R_2^{-1}}$ transforms the components that are known in the rotated system into the components in the laboratory system: $|J\rangle = \mathbf{R_2^{-1}} |J'\rangle$. The subscript $_\mathbf{2}$ is a reminder of the fact that the rotations are 2D: all components of the optical system under consideration are well aligned on the $\hat{\mathbf{z}}$-axis and the anisotropy axis $\hat{\mathbf{a}}$ is normal to $\hat{\mathbf{z}}$. The operation of rotating an optical element is mathematically expressed as

$$\mathbf{O}(\phi) = \mathbf{R_2^{-1}}(\phi)\mathbf{O}(0)\mathbf{R_2}(\phi) \tag{4.68}$$

From right to left the first rotation brings us into the $\hat{\mathbf{x}}'$, $\hat{\mathbf{y}}'$-system; the inverse rotation rotates the output back into the laboratory system. In the same way as a linear retarder, one can also construct a circular retarder made from a chiral medium (see Section 4.2.2.2). The only difference is that the two filters are $\mathbf{F}_L$ and $\mathbf{F}_R$ (instead of $\mathbf{F}_X$ and $\mathbf{F}_Y$). Introducing the abbreviation $\delta k = k_L - k_R$, we obtain the chiral transfer matrix $\mathbf{T}_{\text{Chi}} = e^{ik_L d}\mathbf{F}_L + e^{ik_R d}\mathbf{F}_R$ as

$$\mathbf{T}_{\text{Chi}} = e^{ikd} \begin{pmatrix} \cos(\delta k\, d) & +\sin(\delta k\, d) \\ -\sin(\delta k\, d) & \cos(\delta k\, d) \end{pmatrix} \tag{4.69}$$

This matrix represents the rotation of linear polarization, i.e., "optical activity."

4.3.4.3 Producing Circular Polarization

In order to produce circularly polarized light out of linearly polarized laser light we need two optical elements: (1) an optional λ/2-retarder in a rotary stage, which allows to align the polarization in $\hat{\mathbf{x}}$-direction, if necessary; and (2) a λ/4-retarder whose $\hat{\mathbf{a}}$-axis is aligned in $\pm\pi/4$ with $\hat{\mathbf{x}}$. In this way the linear polarization is split symmetrically into the ordinary and extraordinary modes, arranged in

a V-shaped manner around $\hat{\mathbf{x}}$. The $\pi/2$-phase difference upon passing the $\lambda/4$-plate produces circularly polarized output. In Jones formalism this operations are expressed as $|e(z_{out})\rangle = \mathbf{T}_{\lambda/4}(\pi/4)\,\mathbf{T}_{\lambda/2}(\phi)\,|e(z_{in})\rangle$. The action of the rotating $\lambda/2$-plate $\mathbf{T}_{\lambda/2}(\phi)$ on an arbitrary linearly polarized state $\langle e(z_{\text{in}})| = (\cos(\psi)\ \ \sin(\psi))$ is

$$|e(z_{out})\rangle = \mathbf{T}_{\lambda/2}(\phi)\,|e(z_{in})\rangle = \mathbf{R}_2^{-1}(\phi)\,\mathbf{T}_{\lambda/2}\mathbf{R}_2(\phi)\,|e(z_{in})\rangle \tag{4.70}$$

$$\begin{pmatrix} \cos(2\phi - \psi) \\ \sin(2\phi - \psi) \end{pmatrix} = \begin{pmatrix} \cos(2\phi) + \sin(2\phi) \\ \sin(2\phi) - \cos(2\phi) \end{pmatrix} \begin{pmatrix} \cos(\psi) \\ \sin(\psi) \end{pmatrix} \tag{4.71}$$

Now we align the polarization by setting $2\phi = \psi$ and investigate the action of the $\lambda/4$-plate $\mathbf{T}_{\lambda/4}\left(\pi/4\right)$ on the input state $\langle e(z_{\text{in}})| = (1\ \ 0)$:

$$|e(z_{out})\rangle = \mathbf{T}_{\lambda/4}(\pi/4)\,|e(z_{in})\rangle = \mathbf{R}_2^{-1}(\pi/4)\,\mathbf{T}_{\lambda/4}\mathbf{R}_2(\pi/4)\,|e(z_{in})\rangle \tag{4.72}$$

$$e^{i\pi/4}|R\rangle = e^{i\pi/4}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \end{pmatrix} = e^{i\pi/4}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{4.73}$$

We have now succeeded in producing right-circularly polarized light. The phasor $\exp(i\pi/4)$ is usually irrelevant.

4.3.4.4 Blocking and Filtering Circular Polarization

In diagnostic applications we are interested in the role of circular polarization in the propagation and interaction with tissues. Thus, we must equip the detector with a device which is capable of selectively transmitting one circular polarization state and blocking the other. Having understood the role of the $\lambda/4$-plate as a polarization transformer, the construction is easy: we place into the beam line first a $\lambda/4$-plate $\mathbf{T}_{\lambda/4}(\pi/4)$ and then a linear polarization filter $\mathbf{F}_{X/Y}$; the order is in the direction of light propagation. We take the filter $\mathbf{F}_Y$ and investigate how the combined device acts on the circularly polarized states $\langle R| = \sqrt{1/2}(1\ \ -i)$ and $\langle L| = \sqrt{1/2}(1\ \ i)$. The short-hand formula is

$$|e(z_{\text{out}})\rangle = \mathbf{F}_Y\mathbf{T}_{\lambda/4}(\pi/4)\,|e(z_{\text{in}})\rangle \tag{4.74}$$

Inserting some results from above we get

$$\mathbf{F}_Y\mathbf{T}_{\lambda/4}\left(\pi/4\right) = e^{i\pi/4}\frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix} = e^{i\pi/4}\frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 0 \\ -i & 1 \end{pmatrix} \tag{4.75}$$

Thus the output states for right and left circular polarization are

$$|e(z_{\text{out}})\rangle = e^{i\pi/4}\frac{1}{2}\begin{pmatrix} 0 & 0 \\ -i & 1 \end{pmatrix}\begin{pmatrix} 1 \\ -i \end{pmatrix} = e^{i\pi/4}\begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{4.76}$$

$$|e\,(z_{\text{out}})\rangle \;=\; e^{i\pi/4}\,\frac{1}{2}\begin{pmatrix}0 & 0\\ -i & 1\end{pmatrix}\begin{pmatrix}1\\ -i\end{pmatrix} \;=\; \begin{pmatrix}0\\ 0\end{pmatrix} \tag{4.77}$$

While left circular polarization is completely blocked off, right circular polarization is completely transmitted. Note, however, that the device is not a projecting filter from Eq. (4.62), i.e. not a circular polarizer: the output state is not $|R\rangle$, but rather the linearly polarized state $|Y\rangle$. To make the device into a circular polarizer, we would have to add an appropriately oriented $\mathbf{T}_{\lambda/4}$. Note, however, that the finally measured observable is the power transmitted through the filter and not the polarization vector itself.

4.3.5 Deflectors

Jones' original idea was aimed at the analysis of one dimensional optical systems consisting of polarizing elements that are neatly aligned on a single optical axis. Soon, however, there was a need to extend the concept to include elements that change the direction of light propagation, such as deflecting mirrors, refracting prisms, beam splitters or scattering particles. Since the deflected light still remains a transversal wave (in any case in the far field), we can still employ Jones formalism. One can define a Jones matrix $\mathbf{T}$ for any arbitrary two pairs of base vectors, as indicated on the left hand side of Fig. 4.6. The element $T_{x''y'}$, for example, means then the contribution of the y'-component of the input state into the x'' -component of the output state:

$$\left|E^{out}\right\rangle = \begin{pmatrix}T_{x''x'} & T_{x''y'}\\ T_{y''x'} & T_{y''y'}\end{pmatrix}\left|E^{in}\right\rangle \qquad \text{or explicitly} \qquad \begin{aligned} E^{out}_{x''} &= T_{x''x'}E^{in}_{x'} + T_{x''y'}E^{in}_{y'}\\ E^{out}_{y''} &= T_{y''x'}E^{in}_{x'} + T_{y''y'}E^{in}_{y'}\end{aligned} \tag{4.78}$$

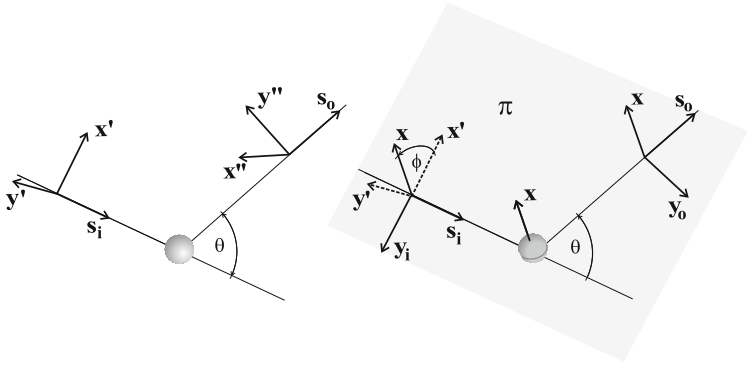


Fig. 4.6 Specifying linearly polarized basis for a deflector

However, a wide spread convention is to use a special choice of the base vectors, which is indicated on the right hand side of Fig. 4.6: one chooses a common $\hat{\mathbf{x}}$-axis that is perpendicular to the plane π spanned by the propagation vectors $\hat{\mathbf{s}}_\mathbf{i}$ and $\hat{\mathbf{s}}_\mathbf{o}$. This plane π is called "scattering plane" or "plane of incidence." This π-system is not always convenient. If there is some symmetry axis in lab frame, like the anisotropy axis $\hat{\mathbf{a}}$ or an overall system axis $\hat{\mathbf{s}}_\mathbf{L}$, then one may prefer a reference plane $\alpha \equiv \text{span}\left(\hat{\mathbf{s}}_\mathbf{i}, \hat{\mathbf{a}}\right)$ or $\zeta \equiv \text{span}\left(\hat{\mathbf{s}}_\mathbf{i}, \hat{\mathbf{s}}_\mathbf{L}\right)$ [14, 15]. However, if both, some $\hat{\mathbf{a}}$ and some $\hat{\mathbf{s}}_\mathbf{L}$, are present, then one better reverts to the π-system.

A good understanding of the geometrical relationships in Fig. 4.6 is essential for polarized Monte Carlo simulations considered in Chapter 7. Deflecting optical elements, such as scatterers, act in two ways: they transform the polarization state, but they also change the direction of propagation by an angle θ so that $\cos(\theta) = \hat{\mathbf{s}}_1 \cdot \hat{\mathbf{s}}_\mathbf{o}$. Thus, a deflection event must be characterized by two transformation matrices: (1) the 2D Jones matrix that operates on the transverse field amplitudes; and (2) the 3D geometrical rotational matrix $\mathbf{R}(\phi, \theta) = \mathbf{R}(\theta)\mathbf{R}(\phi)$. Here $\mathbf{R}(\phi)$ is the rotation by ϕ around $\hat{\mathbf{s}}_\mathbf{i}$, which achieves the transformation from $\hat{\mathbf{x}}'_\mathbf{i}$, $\hat{\mathbf{y}}'_\mathbf{i}$, $\hat{\mathbf{s}}_\mathbf{i}$ into $\hat{\mathbf{x}}_\mathbf{i}$ $\hat{\mathbf{y}}_\mathbf{i}$, $\hat{\mathbf{s}}_\mathbf{i}$. The rotation $\mathbf{R}(\theta)$ around $\hat{\mathbf{x}}_\mathbf{i}$ transforms then from the transverse coordinate system i to the coordinate system $i + 1$ of the next straight path segment. The position and orientation of the elements is specified in the laboratory coordinate system $\hat{\mathbf{x}}_L$, $\hat{\mathbf{y}}_L$, $\hat{\mathbf{z}}_L$, and therefore we must keep track of the overall transformation as it accumulates until i.

4.3.5.1 Snellius and Fresnel Laws in Jones Language

A familiar, yet not quite trivial, example of a deflecting element is an interface between two dielectric media with refractive index mismatch, such as the air-tissue interface. The geometry of the problem is outlined in Fig. 4.7.

An interface between two media with refractive indices n_i (for incoming) and n_t (transmitted) is illuminated by a laser beam with the propagation vector $\hat{\mathbf{s}}_\mathbf{i}$. The interface is optically flat on the length scale that is larger than the diameter of the

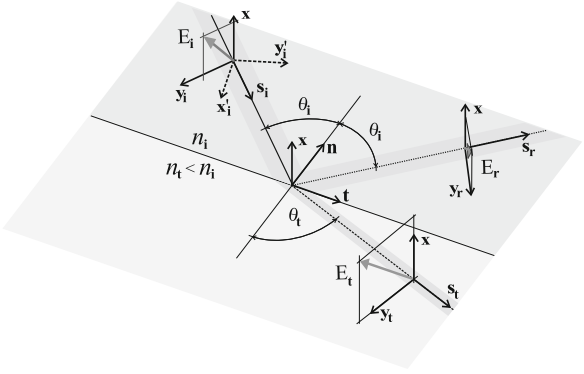


Fig. 4.7 Geometry of Fresnel's laws at an interface

beam. A part of the beam is reflected specularly into the new direction $\hat{\mathbf{s}}_\mathbf{r}$, and a part is transmitted and refracted according to Snellius in the direction $\hat{\mathbf{s}}_\mathbf{t}$. The vectors $\hat{\mathbf{s}}_\mathbf{i}$, $\hat{\mathbf{s}}_\mathbf{r}$ and $\hat{\mathbf{s}}_\mathbf{t}$ lie in the plane π which contains the surface normal $\hat{\mathbf{n}}\,\pi$ (recall Section 4.3.5). To orient this $\hat{\mathbf{n}}$ unambiguously, we shall employ the convention indicated in Fig. 4.7: $\hat{\mathbf{n}}$ points in the direction of the gradient of refractive index. The common axis $\hat{\mathbf{x}}$ is then defined through $\mathbf{x} = \hat{\mathbf{s}}_\mathbf{i} \times \hat{\mathbf{n}} \big/ \left|\hat{\mathbf{s}}_\mathbf{i} \times \hat{\mathbf{n}}\right|$. Note in Fig. 4.7 the surface tangent $\hat{\mathbf{t}}$, such that $\hat{\mathbf{x}}$, $\hat{\mathbf{t}}$ and $\hat{\mathbf{n}}$ span a right handed coordinate system. In the coordinate system $\hat{\mathbf{x}}$,$\hat{\mathbf{t}}$,$\hat{\mathbf{n}}$ the three propagation vectors can be expressed as

$$\begin{aligned}
\hat{\mathbf{s}}_\mathbf{i} &= (0, \quad s_{ti}, \quad s_{ni}) = (0, \quad \sin(\theta_i), \quad \cos(\theta_i)) \\
\hat{\mathbf{s}}_\mathbf{r} &= (0, \quad s_{ti}, \quad -s_{ni}) = (0, \quad -\sin(\theta_i), \quad \cos(\theta_i)) \\
\hat{\mathbf{s}}_\mathbf{t} &= \left(0, \quad \tfrac{n_i}{n_t} s_{ti}, \quad \sqrt{1 - \tfrac{n_i^2}{n_t^2} s_{ti}^2}\right) = \tfrac{n_i}{n_t}\left(0, \quad \sin(\theta_i), \quad \sqrt{\tfrac{n_t^2}{n_i^2} - \sin(\theta_i)^2}\right)
\end{aligned} \tag{4.79}$$

where $s_{ti} = \hat{\mathbf{s}}_\mathbf{i} \cdot \hat{\mathbf{t}}$ and $s_{ni} = \hat{\mathbf{s}}_\mathbf{i} \cdot \hat{\mathbf{n}}$. The expression for the transmitted propagation vector $\hat{\mathbf{s}}_\mathbf{t}$ includes the Snellius law. Exploiting Eq. (4.79) we can immediately write the cosines of the deflection angles $\cos(\theta_{ir}) = \hat{\mathbf{s}}_\mathbf{i} \cdot \hat{\mathbf{s}}_\mathbf{r}$ and $\cos(\theta_{it}) = \hat{\mathbf{s}}_\mathbf{i} \cdot \hat{\mathbf{s}}_\mathbf{t}$. The angles θ_{it}, θ_{ir}, together with ϕ specify completely the 3D transformation matrix (recall Section 4.3.5). We now assume that the rotation $\mathbf{R}(\phi)$ into the π-plane has already been done.

The next step is to obtain the Jones matrix that operates on the polarization. Recall that the field components $E_\parallel \equiv E_y$ and $E_\perp \equiv E_x$, whose polarization are parallel and perpendicular to the plane π, exhibit different reflection and transmission coefficients. The coefficients, found in any textbook, can be combined in the Jones matrices for reflection and transmission. Using the abbreviations $C_i = (\hat{\mathbf{s}}_\mathbf{i} \cdot \hat{\mathbf{n}}) = \cos(\theta_i)$ and $C_t = (\hat{\mathbf{s}}_\mathbf{t} \cdot \hat{\mathbf{n}}) = \cos(\theta_t) = \sqrt{1 - \sin(\theta_t)^2}$ we write

$$\mathbf{T}_\mathbf{r} = \begin{pmatrix} r_\perp & 0 \\ 0 & r \end{pmatrix} \quad \text{where} \quad \begin{aligned} r_\perp &= \tfrac{n_i C_i - n_t C_t}{n_i C_i + n_t C_t} \\ r_\parallel &= \tfrac{n_t C_i - n_i C_t}{n_t C_i + n_i C_t} \end{aligned} \tag{4.80}$$

$$\mathbf{T}_t = \begin{pmatrix} t_\perp & 0 \\ 0 & t \end{pmatrix} \quad \text{where} \quad \begin{aligned} t_\perp &= \tfrac{2 n_i C_i}{n_i C_i + n_t C_t} \\ t_\parallel &= \tfrac{2 n_i C_i}{n_t C_i + n_i C_t} \end{aligned} \tag{4.81}$$

For $\sin(\theta_t)$ we recall the Snellius law $n_t \sin(\theta_t) = n_i \sin(\theta_i)$. A special case is total internal reflection, where $\sin(\theta_t)^2 > 1$ and $\cos(\theta_t)$, thus imaginary. In this case $t_\perp = t_\parallel = 0$ (evanescent waves do not contribute to light propagation) and

$$r_\perp = e^{i\delta_\perp} \qquad \delta_\perp = 2 \arctan\left\{\frac{n_t |C_t|}{n_i C_i}\right\} \tag{4.82}$$

$$r_{||} = e^{i\delta_\perp} \qquad \delta_{||} = 2 \arctan\left\{\frac{n_i |C_t|}{n_t C_i}\right\} \tag{4.83}$$

The corresponding transmittance and reflectance, i.e., transmission and reflection probabilities, are

$$P_r = \frac{\langle E^r | E^r \rangle}{\langle E^i | E^i \rangle} = |r_\perp|^2 |e_{xi}|^2 + |r_{||}|^2 |e_{yi}|^2 = |\mathbf{e_r}|^2 \tag{4.84}$$

$$P_t = \frac{\langle E^t | E^t \rangle}{\langle E^i | E^i \rangle} = \frac{n_t C_t}{n_i C_i} \left\{ |t_\perp|^2 |e_{xi}|^2 + |t_{||}|^2 |e_{yi}|^2 \right\} = \frac{n_t C_t}{n_i C_i} |\mathbf{e_t}|^2 \tag{4.85}$$

where e_{xi} and e_{yi} are the components of the complex unit polarization vector of the input, and $\mathbf{e_r}$, $\mathbf{e_t}$ are the polarization vectors resulting thereof (not unit vectors). Just a little bit more work is needed when it comes to the algorithmic realization of an interface in the course of polarized Monte Carlo simulations (see Chapter 7) The photon state is not known in the convenient π-base $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}_\mathbf{i}$ but in some $\hat{\mathbf{x}}'_\mathbf{i}$, $\hat{\mathbf{y}}'_\mathbf{i}$ which was the outcome of a previous scattering event. Thus, one must first determine $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}_\mathbf{i}$.

4.4 Measuring Polarization

4.4.1 Basic Concepts

The information contained in a Jones vector $|E\rangle$ consists of four numbers, namely the real and imaginary parts of the two complex components E_x and E_y, or their magnitudes $|E_x|, |E_y|$ and their phases ϕ_x, ϕ_y. In optics, the field components are not directly accessible; the measurables are proportional to the complex squares of the field components. However, most of the desired information can be extracted by appropriate filtering. It is important to realize that two kinds of filtering are actually involved: prior application of polarization filters requires that one must first filter a transversal wave out of the impinging light field (see Section 4.5). Here we assume that this spatial filtering is already done. The light field whose polarization is to be measured is a collimated beam that carries the power P, as expressed in Eq. (4.33). The measured signal S is proportional to the energy U or the number of photons $N = U/\hbar\omega$, that cross the detector surface A $\hat{\mathbf{n}}$ during a finite measuring time T. If the detector captures the entire beam, then

$$S \sim U = \int_A \int_T \mathbf{S} \cdot \hat{\mathbf{n}} \, dA dT = \frac{c\epsilon_o}{2} \frac{\langle \kappa_z \rangle}{k^2} \int_T |E(t)|^2 dt = T \langle P \rangle_t \tag{4.86}$$

Here $\mathbf{S}$ is the Poynting vector. On the right hand side we recalled Eq. (4.33) and introduced also the *time averaged power* $\langle P \rangle_t$. Note that we are anticipating the possibility that the field vector $\mathbf{E}(t) = E(t)\hat{\mathbf{e}}(t)$ varies during the measuring time. The measurement involves averaging over time! We shall soon return to this point, but for now we set $\mathbf{E}(t) = \mathbf{E} = \text{const.}$, so that $S \sim |E^2|$. Often this $|E^2|$ is called intensity, denoted with I. "Intensity" is probably the most abused term in physics

and therefore we use S for the measured signal. However, in the present section we neglect all pre-factors, setting S equal to the norm of a Jones vector $|E\rangle$:

$$S = \langle E|E\rangle = |E_x|^2 + \left|E_y\right|^2 \quad \text{where} \quad |\mathrm{E}\rangle = \begin{pmatrix} E_x \\ E_y \end{pmatrix} = \sqrt{S}\begin{pmatrix} e_x \\ e_y \end{pmatrix} \tag{4.87}$$

Upon equipping the detector with a polarization filter $|f\rangle$ the measured signal becomes

$$S_f \;=\; \left\langle E^f \middle| E^f \right\rangle \;=\langle E\,|f\rangle\;\langle f\,|f\rangle\;\langle f\,|E\rangle \;=\; \langle E\,|f\rangle\;\langle f\,|E\rangle \;=\; S\,|\langle f\,|e\rangle|^2 \tag{4.88}$$

Here $|\langle f\,|e\rangle|^2$ is the probability that a photon with polarization $|e\rangle$ will pass through the filter $|f\rangle$. Note that we do not need the Jones matrix of the filter to calculate S_f, but only the scalar product of the state $|e\rangle$ with the filter $|f\rangle$. However, we need a filter set appropriate for the extracting of the full available information on the Jones vector $|E\rangle$. As a guidance we consider the most general filter:

$$|f\rangle \;=\; \begin{pmatrix} \cos(\psi)\; e^{-i\beta/2} \\ \sin(\psi)\; e^{+i\beta/2} \end{pmatrix} \tag{4.89}$$

The signal measured upon transmitting the state $|E\rangle$ through this filter is

$$S_f \;=\; \left\langle E^f \middle| E^f \right\rangle \;=\; |E_x|^2\cos(\psi)^2 \;+\; \left|E_y\right|^2\sin(\psi)^2 + 2\cos(\psi)\sin(\psi)\;\Re\left\{E_x E_y^* e^{i\beta}\right\} \tag{4.90}$$

Upon examining Eq. (4.90) we identify four particularly convenient special cases. Here they are compiled together with the corresponding hardware filter elements:

$$\psi = 0: \qquad |f\rangle = |X\rangle \quad S_X \sim |E_x|^2 \tag{4.91}$$

$$\psi = \pi/2: \qquad |f\rangle = |Y\rangle \quad S_Y \sim \left|E_y\right|^2 \tag{4.92}$$

$$\psi = \pi/4,\ \beta = 0: \qquad |f\rangle = |/\rangle \quad S_/ \sim \frac{1}{2}S_0 + \frac{1}{2}\left[E_x E_y^* + E_x^* E_y\right] \tag{4.93}$$

$$\psi = \pi/4, \beta = \pi/2: \qquad |f\rangle = |R\rangle \quad S_R \sim \frac{1}{2}S_0 - i\frac{1}{2}\left[E_x E_y^* - E_x^* E_y\right] \tag{4.94}$$

where $S_0 \;=\; |E|^2 \;=\; |E_x|^2 \;+\; \left|E_y\right|^2$. Obviously, the components $|E_x|^2$ and $\left|E_y\right|^2$ are extracted using the two orthogonal linear analyzers $|X\rangle$ and $|Y\rangle$. Note also that

$$[E_x E_y^* \;+\; E_x^* E_y] \;=\; 2\Re E_x E_y^* \;=\; 2\,|E_x|\;\left|E_y\right|\cos(\delta) \tag{4.95}$$

$$-\,i[E_x E_y^* \;-\; E_x^* E_y] \;=\; 2\Im E_x E_y^* \;=\; 2\,|E_x|\;\left|E_y\right|\sin(\delta) \tag{4.96}$$

The angle $\delta = \phi_x - \phi_y$ represents the phase difference between the two linear polarizations; $\delta = n_{even}\pi/2$ means linear polarization, $\delta = n_{odd}\pi/2$ circular. Note that two filters are needed to extract δ including its sign, for example, a polarization filter, e.g. $|R\rangle$ (in practice a circular blocker will do), and a linear filter $|/\rangle$ $|\backslash\rangle$. It is easy to solve the equations (4.91-4.94) for the quantities $S_0, |e_x|, |e_y|$, and δ, which fully characterize the polarized harmonic plane wave. Note that the choice of filters in Eq. (4.91) is not unique; one could as well use $|\backslash\rangle$ and $|L\rangle$.

4.4.2 Time Averaging, Stokes Parameters

At this point, we must clarify a subtle but important issue. So far we were concerned with perfectly coherent, pure photon states $|E\rangle$ corresponding to harmonic waves. Such states are always perfectly polarized, i.e., they possess a unique polarization vector $|e\rangle$. This may be a good approximation for a typical laser beam, but certainly not for an old-fashioned light bulb, which produces the so-called "natural unpolarized light." Light reflected from a biological tissue, though not completely depolarized, is still quite different from the laser light. In modern terminology one would say that most of light sources generate state mixtures. In classical electrodynamics the mixture of states can be modeled by replacing the monochromatic wave $\mathbf{E}\exp(-i\omega t)$ with a quasi-monochromatic wave $\mathbf{E}(t)\exp(-i\omega t)$, where the field vector $\mathbf{E}(t)$ fluctuates randomly on a time scale which is much longer than the oscillation period $1/\omega$ but much shorter than the experimental integration time T. The measured quantities are then time averages, for example $\langle S_f \rangle = \langle |\langle f|E\rangle|^2 \rangle$.

Even before the nature of light was fully understood, Sir George Gabriel Stokes realized this fact and suggested organizing the time-averaged information on state mixtures in a set of four real parameters [16]. The four Stokes parameters are:

$$S_0 = I = \langle S_{X/R} \rangle + \langle S_{Y\backslash L} \rangle = \left\langle |E_x|^2 \right\rangle + \left\langle |E_y|^2 \right\rangle \tag{4.97}$$

$$S_1 = Q = \langle S_X \rangle - \langle S_Y \rangle = \left\langle |E_x|^2 \right\rangle - \left\langle |E_y|^2 \right\rangle \tag{4.98}$$

$$S_2 = U = \langle S_/ \rangle - \langle S_\backslash \rangle = \left[\left\langle E_x E_y^* \right\rangle + \left\langle E_x^* E_y \right\rangle \right] \tag{4.99}$$

$$S_3 = V = \langle S_L \rangle - \langle S_R \rangle = -i \left[\left\langle E_x E_y^* \right\rangle - \left\langle E_x^* E_y \right\rangle \right] \tag{4.100}$$

This array of equations is to be interpreted as follows: the first and second columns are two widely used notations for the Stokes parameters. The third column indicates their physical meaning: S_0 is the total time averaged signal $\langle S(t) \rangle$, as it would be measured without any polarization filter. S_1 and S_2, being measured with linear filters, together characterize the degree of linear polarization. S_3 obviously characterizes the degree of circular polarization. The fourth column of the array of equations expresses the parameters in the linearly polarized x, y-basis. All quantities have already appeared in Eqs. (4.91–4.94); the only new aspect is the time

averaging, denoted with the brackets $\langle \ldots \rangle$. Only four measurements are needed to determine all four Stokes parameters. Two of the six measurements indicated in the third column of Eq. (4.97) are superfluous. On the other hand, increasing the number of measurements will improve the accuracy.

To fully appreciate the power of the Stokes approach, we must consider the following derived quantities:

$$m_T = \sqrt{s_1^2 + s_2^2 + s_3^2} \tag{4.101}$$

$$m_L = \sqrt{s_1^2 + s_2^2} \tag{4.102}$$

$$m_c = s_3 \tag{4.103}$$

Here s_i are the normalized versions of the Stokes parameters, i.e. $s_i = S_i/S_0$. The quantity m_T can be interpreted as the total degree of polarization, or, the degree of polarization coherence. It is easy to verify that $m_T = 1$ for a coherent pure state with constant E_x and E_y. In other words, for such pure states $s_1^2 + s_2^2 + s_3^2 = 1$. The quantities m_L and m_C characterize the degree of linear and circular polarization, respectively. Note that $m_L^2 + m_c^2 = m_T^2$. As an opposite example, consider an incoherent mixture of linearly polarized states $\langle e| = (\cos(\psi) \sin(\psi))$, where ψ is uniformly distributed between 0 and 2π. In this case $\left\langle E_x E_y^* \right\rangle = \langle \cos(\psi) \sin(\psi) \rangle = 0$ and $\left\langle E_y^2 \right\rangle = \langle E_x^2 \rangle$. Thus, $m_T = 0$. In general, $m_T < 1$ for partially polarized light.

4.4.3 General Experiment Design, Coherence Matrix

If Stokes had lived 100 years later, he would certainly had defined his parameters in terms of Pauli matrices. It is quite easy to verify that:

$$S_0 := \langle \langle E | \sigma_0 | E \rangle \rangle \quad S_1 := \langle \langle E | \sigma_1 | E \rangle \rangle \quad S_2 := \langle \langle E | \sigma_2 | E \rangle \rangle \quad S_3 := \langle \langle E | \sigma_3 | E \rangle \rangle \tag{4.104}$$

where

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \sigma_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} \tag{4.105}$$

Warning: Our notation of Pauli matrices differs slightly from the standard used in QT. Since optics and QT developed along separate paths, we have three options: (1) a confusion of signs and subscripts; (2) different notation for Stokes parameters; or (3) different notation for the σ-s. Since this chapter concerns optics, we decided on (3). Pauli matrices exhibit certain peculiar and useful properties that can be easily understood when one realizes their relation with the projection operators from Eqs. (4.60), (4.61) and (4.62):

$$\sigma_0 \equiv \mathbf{F}_{X/R} + \mathbf{F}_{Y\backslash L} \quad \sigma_1 \equiv \mathbf{F}_X - \mathbf{F}_Y \quad \sigma_2 \equiv \mathbf{F}_/ - \mathbf{F}_\backslash \quad \sigma_3 \equiv \mathbf{F}_R - \mathbf{F}_L \tag{4.106}$$

To demonstrate the practical usefulness of Pauli matrices, we return to Eq. (4.88), which relates the signal transmitted through a filter $|f\rangle$ with an input state $|E\rangle$. Introducing the time averaging, the measurable becomes

$$\langle S_f \rangle = \langle \langle E|f\rangle \langle f|E\rangle \rangle = \langle \langle f|E\rangle \langle E|f\rangle \rangle \tag{4.107}$$

Note here the projection operators $|f\rangle\langle f|$ or $|E\rangle\langle E|$ (recall Eq. (4.59)). Upon evaluating $|E\rangle\langle E|$ we find that Eq. (4.107) can be expressed in terms of a matrix $\mathbf{C}$ as $\langle S_f \rangle = \langle f|C|f\rangle$, where

$$\mathbf{C} = \langle |E\rangle\langle E| \rangle = \left\langle \begin{pmatrix} E_x \\ E_y \end{pmatrix} \begin{pmatrix} E_x^* E_y^* \end{pmatrix} \right\rangle = \begin{pmatrix} \langle E_x E_x^* \rangle & \langle E_x E_y^* \rangle \\ \langle E_y E_x^* \rangle & \langle E_y E_y^* \rangle \end{pmatrix} \tag{4.108}$$

Note that the complex hermitian "coherency matrix" $\mathbf{C}$ contains the same information as the Stokes parameters S_i. On the diagonal there are two real numbers, and the off-diagonal terms differ only in the sign of the imaginary part. Thus, four measurements are needed to determine $\mathbf{C}$. Now we exploit a marvelous property of Pauli matrices: they form a complete set of independent matrices, i.e., any 2×2 matrix can be written as a linear superposition of σ-s. The reader may verify that the Pauli expansion of the coherency matrix $\mathbf{C}$ reads:

$$\mathbf{C} = \frac{1}{2} \sum_{i=0}^{3} S_i \sigma_i \tag{4.109}$$

Here S_i are the Stokes parameters of the state $|E\rangle$. We insert this expansion into the right hand side of Eq. (4.107), observing thereby the quantities $\langle f|\sigma_i|f\rangle$. Thus, we recall Eq. (4.104) and obtain finally a beautifully simple expression for the measurable $\langle S_f \rangle$:

$$\langle S_f \rangle = \frac{1}{2} \sum_{i=0}^{3} s_{fi} S_i \tag{4.110}$$

where s_{fi} are the Stokes parameters of the filter state $|f\rangle$. The practical use of the small exercise with Pauli matrices for the polarization analysis should be obvious: Eq. (4.110) is only one linear equation for the four unknowns S_0, S_1, S_2 and S_3. Using four linearly independent measurements $\langle S_f \rangle$, one can determine all four Stokes parameters. If one is willing to do more measurements, one obtains an overdetermined equation system, which can be solved by standard techniques, in order to improve the accuracy of the determination of the parameters. The required filter coefficients s_{fi} are easily calculated upon programming the definitions (Eq. (4.104)) in Maple or Mathematica.

4.4.4 Stokes Vectors

To conclude the story about Stokes parameters, we note that the sum in Eq. (4.110) can be interpreted as the scalar product of two 4D Stokes-vectors:

$$\langle S_f \rangle = \frac{1}{2}\vec{s}_f \cdot \vec{S} \quad \text{where} \quad \vec{S} = (S_o,\ S_1,\ S_2,\ S_3) \tag{4.111}$$

The Stokes vectors of the three fundamental pairs of states can be easily obtained by inserting Eq. (4.55) into (4.104):

$$\begin{array}{lll} \vec{s}_X = (1,\ +1,\ 0,\ 0) & \vec{s}_{/} = (1,\ 0,\ +1,\ 0) & \vec{s}_R = (1,\ 0,\ 0,\ +1) \\ \vec{s}_Y = (1,\ -1,\ 0,\ 0) & \vec{s}_{\backslash} = (1,\ 0,\ -1,\ 0) & \vec{s}_L = (1,\ 0,\ 0,\ -1) \end{array} \tag{4.112}$$

One may now realize why those three states are so fundamental: using these three pairs we can define a set of base vectors of the 4D Stokes space:

$$\vec{s}_0 = \frac{1}{2}\left(\vec{s}_{X/R} + \vec{s}_{Y\backslash L}\right) = (1,\ 0,\ 0,\ 0) \tag{4.113}$$

$$\vec{s}_1 = \frac{1}{2}\left(\vec{s}_X - \vec{s}_Y\right) = (0,\ 1,\ 0,\ 0) \tag{4.114}$$

$$\vec{s}_2 = \frac{1}{2}\left(\vec{s}_{/} - \vec{s}_{\backslash}\right) = (0,\ 0,\ 1,\ 0) \tag{4.115}$$

$$\vec{s}_3 = \frac{1}{2}\left(\vec{s}_R - \vec{s}_L\right) = (0,\ 0,\ 0,\ 1) \tag{4.116}$$

The vector $\vec{s}_0$ represents an unpolarized light field, such as would result from an incoherent superposition of any two orthogonal pure states with equal intensities. The remaining three can be called polarization unit vectors of the three fundamental states. Note, however, that the states $\vec{s}_1$, $\vec{s}_2$, and $\vec{s}_3$ do not carry any power. They must be always combined with $\vec{s}_0$.

4.5 Beam Filters and Beam States

The optical system we have in mind is not an array of polarizing elements and deflectors on an optical table but a sample of biological tissue. Consider the situation sketched in Fig. 4.8. The sample is irradiated by a nice coherent beam wave, but what comes out, apart from some specular reflection, is a disordered mixture of quasi-spherical waves. From this mixture one must first filter out a beam wave suitable for polarization analysis. One way to do such wave filtering is to place a very small detector very far in the far field, so that the superposition of the partial waves from the tissue appears as if coming from a point source, far enough to have a flat wave front over the entire detector area. In M. von Laue's words: "... we move the detector point always so far away, that if a spherical wave would originate in

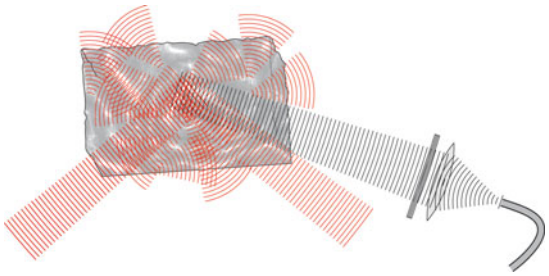


Fig. 4.8 Random light field generated by scattering in a tissue sample and pure beam selection by a single-mode receiver. Note that we avoid specular reflection

this point, it would look like a plane wave in the region where the scattering bodies are contained." [17]. One can show that the plane wave approximation requires that $L^2 \ll \lambda R$, where R is the distance to the detector point and L is the size of the region of interest. This far field assumption is certainly well fulfilled when observing stellar atmospheres, but hard to fulfill on a distance separating two consecutive scattering events during propagation in tissue. This kind of wave filtering is highly impractical (and useless for imaging), and therefore it is only used in textbooks. In optical practice one uses spatial filters consisting of lenses and apertures that select simultaneously the k-vector range and the observed position. This kind of wave filtering works fine, especially for imaging. But it is still not as perfect as polarization filtering with a Glan-Thompson prism. To select a pure beam state analogous to a pure polarization state one would have to use a nearly point-like detector giving nearly zero power. A Glan-Thompson prism does the filtering virtually without any power loss.

While perfect polarizers have been around for more than a hundred years, single-mode beam filters are a relatively new addition to photonic instrumentation. A beam filter is constructed from a single-mode optical fiber (single-mode wave guide), a collimating lens and a polarization filter, as illustrated in Fig. 4.8. The beam filter works both as a transmitter or as a receiver. No matter what light source (e.g. a laser diode of appropriate wavelength) is coupled into the single-mode fiber, there is always the same beam emanating through the collimator lens. Replacing the laser diode with a detector, one obtains a single-mode receiver. No matter what light field is observed, one is sure to always select the same beam that propagates into the fiber. Besides being a practical device for scattering experiments [18, 19], the single-mode receiver also has a theoretical advantage: using the single-mode concepts, one can treat scattering problems (involving interference) in a rigorous yet rather simple fashion [20]. Beam filters, as theoretical concepts and/or practical devices, will be useful for polarization imaging (Section 7.5.6), for the detection of scattered light (Section 7.1.1) as well as for bridging electrodynamics with Monte Carlo simulations.

4.5.1 Mathematical Representation of Beam Filters

Electrodynamic theory provides tools to make the mode selection by a single-mode receiver very explicit: One can define an electromagnetic field $\hat{\mathbf{E}}^o, \hat{\mathbf{B}}^o \equiv |o\rangle$

that possesses all the properties of the classical electromagnetic field which would emanate from a transmitter (recall von Laue's remark), except that its propagation is reversed (time reversal) and that the field is empty. We shall call this field observation mode. The observation mode $\hat{\mathbf{E}}^o$, $\hat{\mathbf{B}}^o$ is a *virtual* field in the original sense of the word: a virtual field does not carry photons, but it has the potential to do so. (In this spirit, one may regard all fields as purely geometrical constructions that provide a path system for the transport of particles.) In Sections 4.3 and 4.4 we pointed out that filtering is synonymous with the projection operation. Projection, in turn, requires the existence of a norm $\langle f | f \rangle$ and of a scalar product $\langle E | f \rangle$. Recall that the norm of a beam state is defined as

$$\langle E | E \rangle = \oint_A \mathbf{S} \cdot d\mathbf{a} \equiv \frac{1}{4} \frac{1}{\mu_o} \oint_\sigma [E \times \mathbf{B}^* + E^* \times \mathbf{B}] \cdot \hat{\mathbf{n}} \, d^2 r \tag{4.117}$$

where A is an arbitrary surface enclosing the source of the beam (i.e. a laser); $\hat{\mathbf{n}}$ is the outer normal to this surface. In the case of an ordinary laser beam, as defined in Eq. (4.29), or in the quasi plane beam wave version in Eq. (4.38), the norm is the power carried in the beam: $\langle E | E \rangle = P$. A virtual observation beam does not carry any power, but otherwise it looks much like an ordinary beam wave. We are mostly interested in polarized paraxial beams in the Rayleigh range, whose electric field is given by Eq. (4.38). Since the virtual observation mode does not carry any power, we define the observation beam as a "powerless" version of Eq. (4.38):

$$\hat{\mathbf{E}}^{\mathbf{o}}(r) = \ldots \sqrt{\frac{2}{c\epsilon_o}} X(r) \, \hat{\mathbf{e}}_{\mathbf{o}} e^{ik\hat{\mathbf{s}}_{\mathbf{o}} \cdot \mathbf{r}} \tag{4.118}$$

The profile $X(\mathbf{r})$ is the beam profile from Eq. (4.36), expressed, however, in an arbitrary oriented coordinate system. A corresponding transformation concerns also the components of the polarization vector $\hat{\mathbf{e}}_{\mathbf{o}}$ and the propagation vector $\hat{\mathbf{s}}_{\mathbf{o}}$ of the observation beam. Upon inserting this powerless $\hat{\mathbf{E}}^o$ into Eq. (4.117), one finds that

$$\langle o | o \rangle = \frac{1}{4} \frac{1}{\mu_o} \oint_\sigma \left[\hat{\mathbf{E}} \times \hat{\mathbf{B}}^* + \hat{\mathbf{E}}^* \times \hat{\mathbf{B}} \right] \cdot \hat{\mathbf{n}} \, d^2 r = 1 \tag{4.119}$$

Note that the "hat" mark over the field amplitudes does not mean that $\hat{\mathbf{E}}$ and $\hat{\mathbf{B}}$ were unit vectors. It is just a reminder of the fact that the norm of a virtual state $|o\rangle$ is 1, just as the norm of a unit Euclidean vector. Consequently, we define the scalar product of an external field $|E\rangle$ with an observation mode $|o\rangle$ as

$$\langle o | E \rangle = \frac{1}{4} \frac{1}{\mu_o} \oint_A \left[\mathbf{E} \times \hat{\mathbf{B}}^{o*} + \hat{\mathbf{E}}^{o*} \times \mathbf{B} \right] \cdot \hat{\mathbf{n}} \, d^2 r \tag{4.120}$$

where A is an arbitrary surface enclosing the sink of the observation beam (a detector) but excluding the source of $|E\rangle$; $\hat{\mathbf{n}}$ is the inner normal to this surface. If both

fields are transversal (in the far field or in Rayleigh range) and their propagation vectors $\hat{\mathbf{s}}_e$ and $\hat{\mathbf{s}}_o$ nearly collinear, then the following expression is a useful approximation of Eq. (4.120):

$$\langle o\,|E\rangle \;\approx\; \frac{c\epsilon_o}{4}\oint_A \left(\hat{\mathbf{E}}^{o*}\cdot\mathbf{E}\right)\left[\hat{\mathbf{s}}_e+\hat{\mathbf{s}}_o\right]\cdot\hat{\mathbf{n}}\,d^2r \tag{4.121}$$

Note here the Fresnel inclination factor $\left[\hat{\mathbf{s}}_e+\hat{\mathbf{s}}_o\right]\cdot\hat{\mathbf{n}}/2$ and the polarization factor $\left(\hat{\mathbf{E}}^{o*}\cdot\mathbf{E}\right)$, both well known from the rigorous diffraction theory. Equation (4.121) is particularly useful for the description of the coupling of laser beams into single-mode fibers. We shall need it in the context of polarization imaging.

In some cases, the external field $|E\rangle$ is too complicated for an explicit expression, but the source of this field, i.e. the oscillating current distribution $\mathbf{j}$ that radiates this field, is simple. For this case, the following (exact) version of Eq. (4.120) can be derived using a magic wand [21] from the tools of electromagnetism (Poynting theorem from Eq. (4.8) and reciprocity theorem [22]):

$$\langle o\,|E\rangle \;=\; -\frac{1}{4}\int_V \hat{\mathbf{E}}^{o*}\cdot\mathbf{j}\,d^3r \tag{4.122}$$

Here V is an arbitrary volume that encompasses the source $\mathbf{j}$ of $|E\rangle$ but excludes the sink of $|o\rangle$. Equation (4.122) will be quite useful for concise treatment of scattering problems that involve interference, in particular in the single scattering limit.

Knowing the scalar product, we can calculate the field in the beam that is selected by the receiver as the projection of the external field on the receiver mode

$$\left|E^{os}\right\rangle \;=\; |o\rangle\,\langle o\,|E\rangle \;=\; |o\rangle\,\mathcal{E} \tag{4.123}$$

The abbreviation $\mathcal{E}\equiv\langle E\,|o\rangle$ is introduced to interpret the scalar product as a *complex amplitude*. In order to get the real field, one simply takes the virtual field from Eq. (4.118) and multiplies it with the complex amplitude $\mathcal{E}$:

$$\mathbf{E}^{os}(\mathbf{r}) \;=\; \hat{\mathbf{E}}^{\mathbf{o}}(\mathbf{r})\mathcal{E} \tag{4.124}$$

Note that $|o\rangle$ and $\langle o|$ are complex conjugate and thus the phase of the receiver mode does not enter the selected field $\mathbf{E}^{os}(\mathbf{r})$. The selected field $\mathbf{E}^{os}(\mathbf{r})$ inherits the phase of the source. Thus we can write the complex amplitude as

$$\mathcal{E}\equiv\langle E\,|o\rangle \;=\; \sqrt{P_{os}}\,e^{i\Phi_s} \tag{4.125}$$

where Φ_s is the inherited phase. The received signal P_{os}, i.e. the power measured with a single-mode receiver, is given by the general expression, Eq. (4.88). Since the observation beam is normalized, we can write

$$P_{os} \;=\; |\,\langle E\,|o\rangle\,|^2 \;=\; |\mathcal{E}|^2 \tag{4.126}$$

Note that $P = |\mathcal{E}|^2$ is completely analogous with $I = |E|^2$. In many situations one does not care about the unmeasurable local details of the electromagnetic field of a beam wave and so the knowledge of $\mathcal{E}$ is all one needs. However, once $\mathcal{E}$ is known, we can calculate the field of the selected beam at any position, in particlar in the far field, according to Eq. (4.124). For the calculation of $\langle E\,|o\rangle$ one can use any of the expressions from Eqs. (4.120), (4.121) or (4.122). Of course, one would choose the form which is most convenient for a particular problem.

4.5.2 Virtual Beams as Base Vectors

In Section 4.3.3 we introduced the Jones formalism in a rather careless traditional way and omitted the details of the structure of the fields. Essentially we pretended that when working with plane waves, only the field amplitude E and the polarization vector $\hat{\mathbf{e}}$ mattered. In laboratory practice, however, one usually works with beams. The measured quantity is the beam power P and not some local intensity $I(\mathbf{r}) \sim |E\,(\mathbf{r})|^2$. Having understood the concept of virtual modes as unit vectors in the space of the solutions of Maxwell Equations, we are now in the position to mend the inconsistency and put the Jones formalism on sound formal footing. To express a laser beam in, for example, XY-polarized base, we write

$$\mathbf{E}^{\mathbf{I}}\,(\mathbf{r}) \;\equiv\; \left|E^{I}\right\rangle \;=\; |X\rangle\;\mathcal{E}_x \;+\; |Y\rangle\;\mathcal{E}_y \tag{4.127}$$

where $|X\rangle$ and $|Y\rangle$ are unit vectors, i.e. virtual beams normalized such that $\langle X\,|X\rangle \;=\; 1$ and $\langle Y\,|Y\rangle \;=\; 1$. Explicit definitions of the two unit vectors are

$$\begin{aligned} |X\rangle \;&\equiv\; \hat{\mathbf{E}}^{\mathbf{x}}\,(\mathbf{r}) \;=\; \ldots\sqrt{\tfrac{2}{c\varepsilon_o}}X\,(\mathbf{r})\;\hat{\mathbf{x}}\;e^{ik\hat{\mathbf{s}}_1\cdot r} \\ |Y\rangle \;&\equiv\; \hat{\mathbf{E}}^{\mathbf{y}}\,(\mathbf{r}) \;=\; \ldots\sqrt{\tfrac{2}{c\varepsilon_o}}X\,(\mathbf{r})\;\hat{\mathbf{y}}\;e^{ik\hat{\mathbf{s}}_1\cdot r} \end{aligned} \tag{4.128}$$

As an exercise one may use Eq. (4.121) to verify that the complex amplitudes $\mathcal{E}_x$ and $\mathcal{E}_y$ are the scalar products defined in previous section. A Jones vector is defined as a two dimensional vector whose components are the complex amplitudes

$$|E\rangle \;:=\; \begin{pmatrix}\mathcal{E}_x\\ \mathcal{E}_y\end{pmatrix} \qquad \langle E| \;:=\; \begin{pmatrix}\mathcal{E}_x^* & \mathcal{E}_y^*\end{pmatrix} \tag{4.129}$$

so that

$$\langle E\,|E\rangle \;=\; |\mathcal{E}_x|^2 \;+\; \left|\mathcal{E}_y\right|^2 \;=\; P \tag{4.130}$$

The virtual beam waves propagate through the optical system just as ordinary waves, which would allow us to extend Jones formalism to optical systems of the

"more familiar type involving lenses, prisms, etc." (Recall the citation from the preamble of Section 4.3.)

4.6 Polarization Analysis of an Unknown Optical System

4.6.1 Designing the Experiment, from Jones to Perrin-Mueller Matrix

Having filtered out a beam wave from the scattered, transmitted or reflected light, we can concentrate on polarization analysis. The geometry is sketched in Fig. 4.9. Note that we are trying to avoid specular reflection, because we would like to obtain information from the depth of the sample. The illumination and the observation channels are equipped with polarization optics as needed to manipulate and filter the polarization states. The analysis is done in the π-base system (recall Section 4.3.5).

How much information is contained in the polarization of the scattered light field and what measurements are needed to retrieve this information? After the exercise in Sections 4.3.3 and 4.4, the answers to these questions should not be too difficult. The transformation of the input state $\left|e^{\text{in}}\right\rangle$ into an output state $\left|E^{\text{out}}\right\rangle$ is usually achieved by an amplitude transfer matrix (Jones matrix) $\mathbf{T}$, which, however, depends on the geometry of illumination and observation:

$$\left|E^{\text{out}}\,(t)\right\rangle = \mathbf{T}\left(\hat{\mathbf{s}}_{\mathbf{i}},\ \hat{\mathbf{s}}_{\mathbf{o}},\ t\right)\ \left|e^{\text{in}}\right\rangle \tag{4.131}$$

Note that in Eq. (4.131) we are using two different symbols for the input and for the output: whereas $\left|e^{\text{in}}\right\rangle$ is a coherent state normalized so that $\left\langle e^{\text{in}}\left|e^{\text{in}}\right.\right\rangle = 1$, the output state $\left|E^{\text{out}}\,(t)\right\rangle$ is in general a mixture. In fact, $\mathbf{T}\left(\hat{\mathbf{s}}_{\mathbf{i}},\ \hat{\mathbf{s}}_{\mathbf{o}},\ t\right)$ exhibits random fluctuations, unless the tissue sample is stiffly frozen. But even if the sample were

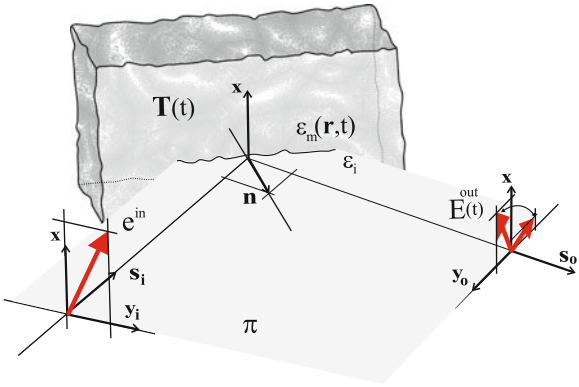


Fig. 4.9 Simplified geometry for polarization diagnostics of a tissue sample

frozen, we are not as much interested in a particular random realization of the tissue structure as in an ensemble average characteristic for the given tissue type. The characterization will always involve averaging of the filtered signals.

We return to Eq. (4.107), inserting therein the output state $\left|E^{\text{out}}(t)\right\rangle$ from Eq. (4.131):

$$\left\langle S_{fe}(\mathbf{T})\right\rangle \;=\; \langle\langle f\,|E\rangle\;\langle E\,|f\rangle\rangle \;=\; \left\langle \langle f\,|\mathbf{T}|\,e\rangle\;\left\langle e\left|\mathbf{T}^{\dagger}\right|f\right\rangle\right\rangle \tag{4.132}$$

On the right hand side † means Hermitian adjoint, i.e. transposed conjugate of $\mathbf{T}$. Obviously, any measured $\left\langle S_{fe}(\mathbf{T})\right\rangle$ is some real linear combination of the correlations $\left\langle T_{ij}T_{kl}^{*}\right\rangle$. Thus, the 16 averages $\left\langle T_{ij}T_{kl}^{*}\right\rangle$ represent the full amount of information on the observed system which can be retrieved from time averaged observations. There are 4 real quantities $\left\langle T_{ij}T_{ij}^{*}\right\rangle$, the remaining 12 are complex conjugate pairs. Thus the information can be retrieved using 16 linearly independent combinations of the filters $|e\rangle$ and $|f\rangle$. The problem is completely analogous with the characterization of an unknown state vector $|E\rangle$, but organizing the 16 measurements in a clever way is a demanding task. Luckily we prepared the needed tools: in Section 4.4.3 we learned how to re-write $\langle\langle f\,|E\rangle\langle E\,|f\rangle\rangle$ in terms of a scalar product of two Stokes vectors: $\left\langle S_{fe}\right\rangle \;\sim\; \vec{s}_f\cdot\vec{S}$, where $\vec{S}$ is the Stokes vector of the output state $|E\rangle$. If we were able to find a 4×4 matrix $\mathbf{M}$ which transforms the Stokes vector $\vec{s}_e$ of the input state $|e\rangle$ into the output vector $\vec{S}$, then we could write $\left\langle S_{fe}\right\rangle \sim \vec{s}_f\cdot\mathbf{M}\cdot\vec{s}_e$. Such a matrix $\mathbf{M}$ does indeed exist and is called Mueller matrix, despite the fact that it had already been introduced by the French physicist Francis Perrin (son of Jean Perrin) in 1942 [23]. To derive the Perrin-Mueller matrix, we shall exploit the wonderful properties of Pauli matrices. First we extract from Eq. (4.132) the identity $|E\rangle\,\langle E|\;\equiv\mathbf{T}\,|e\rangle\,\langle e|\,\mathrm{T}^{\dagger}$. (*For the sake of readability we omit for a while the outer brackets of time averaging.*) Expanding the projectors on both sides in terms of Pauli matrices we obtain a matrix equation that relates the Stokes parameters of the input and output:

$$\sum_k S_k\sigma_k \;=\; \sum_i s_{ei}\mathbf{T}\sigma_i\mathbf{T}^{\dagger} \tag{4.133}$$

These equations look a little bit strange, but they can be converted into a set of four ordinary linear equations using the following properties of the Pauli matrices: (1) the matrices σ_1, σ_2 and σ_3 are traceless, i.e. the sums of their diagonal elements are always zero. Only the identity matrix σ_0 has a non-zero trace, namely 2. (2) the product of two Pauli matrices $\sigma_k\sigma_i$ is always a traceless Pauli matrix, except for $\sigma_j\sigma_j=\sigma_0$. In other words, Trace $(\sigma_k\sigma_j)=2\delta_{kj}$, where $\delta_{kj}=1$ if $j=k$ and $\delta_{kj}=0$ if $j\neq k$. Thus, upon multiplying both sides of Eq. (4.133) with σ_j and carrying out the Trace operation we obtain:

$$S_j \;=\; \sum_i s_{ei}\,M_{ji}, \quad \text{where} \quad M_{ji} \;=\; \frac{1}{2}\text{Trace}\left(\mathbf{T}\sigma_i\mathbf{T}^{\dagger}\sigma_j\right) \tag{4.134}$$

Here M_{ji} are the elements of the Mueller matrix, j is the row index. Since the matrix elements can be easily calculated using an algebra program, we refrain from reproducing the explicit results. The expressions look somewhat ugly (see e.g. [24]), but the formal beauty of the Perrin-Mueller approach becomes apparent in the following expression for the measured signal

$$\langle S_{fe} \rangle = \frac{1}{2} \vec{s}_f \cdot \mathbf{M} \cdot \vec{s}_e \tag{4.135}$$

Upon carrying out the multiplications in Eq. (4.135) for one input state $\vec{s}_e$ and one filter $\vec{s}_f$, one obtains one of the 16 equations which are needed to determine the 16 elements of an unknown Mueller matrix from the measurements $\langle S_{fe} \rangle$. One may rightly guess that the sixteen measurements can be done by permuting the special states $|X\rangle$, $|Y\rangle$, $|/\rangle$ and $|R\rangle$, one of those elements in the illumination channel and one in the observation. Of course these 4×4 measurements are not the only way how to set up the equation matrix. A little bit of patience will improve the accuracy as well as the understanding. We suggest doing at least 6×4 measurements: the three fundamental pairs for the input, and at least 4 measurements to determine each of the output Stokes vectors $\vec{O}_e = M\vec{s}_e$. The Stokes vectors $\vec{s}_e$ of the three fundamental pairs are given in Eq. (4.112). One immediately sees that

$$\begin{array}{lll} \vec{O}_X = \vec{M}_0 + \vec{M}_1 & \vec{O}_/ = \vec{M}_0 + \vec{M}_2 & \vec{O}_R = \vec{M}_0 + \vec{M}_3 \\ \vec{O}_Y = \vec{M}_0 - \vec{M}_1 & \vec{O}_\backslash = \vec{M}_0 - \vec{M}_2 & \vec{O}_L = \vec{M}_0 - \vec{M}_3 \end{array} \tag{4.136}$$

where $\vec{M}_j$ are the columns of the Mueller matrix $\mathbf{M} = \left(\vec{M}_0, \vec{M}_1, \vec{M}_2, \vec{M}_3\right)$. Thus, by adding and subtracting the Stokes outputs of the fundamental pairs, one determines entire columns of the Mueller matrix. A nice compilation of other different possibilities for organizing the measurements, as well as practical hints, can be found in [25]. In fact, there are infinite possibilities. Contemporary computation tools allow solutions of an arbitrary large set of over-determined equations and good algorithms complain reliably if the system is under-determined (SVD). The bottle neck is the measurement. As with the determination of an unknown Stokes vector, the more measurements one is willing to do, the more accurate is the result.

4.6.2 Interpreting the Perrin-Mueller Matrix, Alternative Codings

Having understood the column structure of the Perrin-Mueller matrix, we can attempt an interpretation. This can be best done by applying the matrix to linear combinations of the polarization unit vectors defined in Eq. (4.113). The unpolarized unit vector $\vec{s}_0$ picks the column $\vec{M}_0$. This column represents re-polarization of an unpolarized state mixture: if any of the elements $M_{k0},\ k = 1,\ 2,\ 3$ is non zero, then the output is at least partially polarized. (Recall for example the polarization of sunlight by scattering in the atmosphere.) The remaining three columns (together with column 0) describe the transfer of the power from one of the polarization

states to the others. Note, for example, that in order to make circular polarization out of linear states $\vec{s}_x$ or $\vec{s}_/$ the elements M_{30}, M_{31} and/or M_{32} must be non-zero. For the elements of the amplitude transfer matrix **T** this means either non-zero off-diagonal terms T_{xy}, T_{yx} so that e.g. $\Im\left\langle T_{xx}T_{yx}^*\right\rangle \neq 0$ (in M_{30} and M_{31}), or complex diagonal terms with phase shift between T_{xx} and T_{yy} so that $\Im\left\langle T_{xx}T_{yy}^*\right\rangle$ (in M_{32}). On the other hand, to make linear polarization out of $\vec{s}_R$ or $\vec{s}_L$ it is sufficient that $\left\langle |T_{xx}|^2\right\rangle \neq \left\langle |T_{yy}|^2\right\rangle$ (in element M_{10}).

Apparently, the information contained in **M** is not as easily interpretable as in the case of the Stokes vector. It may well be that this particular coding rather obscures the sought diagnostic information instead of enhancing it. Therefore, instead of brooding too long over **M**, one should always try an alternative coding. For example, one can survey the real and imaginary parts of the elements $\left\langle T_{ij}T_{kl}^*\right\rangle$ of the 4D coherence matrix, which can be arranged in a real matrix **P**, and which contain the same information as the Mueller matrix. Finally, one can survey directly the matrix **S** containing the 16 independent observables $\left\langle S_{fe}\right\rangle$ instead of converting **S** into the **P** or **M** matrices. For example, without birefringence it is difficult to make circular polarization out of linear. Thus the measurements $\langle S_{RX}\rangle$ and/or $\langle S_{LX}\rangle$ may be suited to diagnose birefringence. We have at least three matrices, **S**, **P**, and **M**, which contain exactly the same information, only in a different coding. Which coding would provide the most discernable signature of, say, the chirality of the medium? It would be nice to find a set of 16 really existing or imaginable optical devices (filters, blockers, depolarizers, transformers, attenuators, ...) whose 16 matrices would form a complete set for an intelligible decomposition of **S**, **P** or **M** matrices. In any case, when designing data analysis procedures for tissue diagnostics one should be guided not only by canonical formalisms but also by scientific imagination. The research on the interpretation of the Mueller and other matrices continues [26].

4.6.3 Mueller Versus Jones

Carrying the reasoning from the previous section a little bit further, one discovers that by adding new elements into the optical path the game with Mueller matrices and Stokes vectors can be continued. Propagation of polarized light through the optical system can be analyzed with the so called Mueller formalism, i.e. in terms of Mueller matrices and Stokes vectors, as easily as with the Jones formalism which we employed in Section 4.3. Any Jones matrix, be it a filter **F**, rotation matrix **R** or some retarder transfer matrix **T**, can be converted into its Mueller matrix using the trace-formula in Eq. (4.134), and any Jones vector $|E\rangle$ can be converted into a Stokes vector using Eq. (4.104). Thus the following two expressions are almost completely equivalent:

$$\left|E^{\text{out}}\right\rangle = \mathbf{F}\mathbf{R}^{-1}\mathbf{T}_{\lambda/2}\mathbf{R}\left|e^{\text{in}}\right\rangle \equiv \vec{S}_{\text{out}} = \mathbf{M}_F\mathbf{M}_{-R}\mathbf{M}_{\lambda/2}\mathbf{M}_R\cdot\vec{s}_{\text{in}} \tag{4.137}$$

Thus, we have at least two possibilities for rationalizing the work on propagation of polarized light through biological tissues: (1) Jones calculus that operates with 2×2 complex matrices and propagates the *amplitudes*; and (2) Stokes-Mueller calculus, that operates with 4×4 real matrices and propagates the *probabilities*. Hans Müller, an MIT scientist, published his calculus in 1948, only a few years after the Harvardian Jones [27]. As it tends to happen between MIT and Harvard, the two approaches are often felt as a sort of competition, which manifests itself in the following statement that is found obligately in most texts on the polarization: "Jones formalism can't treat unpolarized light." In our opinion, this is a misunderstanding. There is no such thing as "unpolarized light." But there are light fields where the polarization vector is a randomly fluctuating quantity. One can always do the Jones calculus and calculate probabilities or averages at any step of the light path. In a way, the Mueller-matrix characterizes an ensemble of Jones matrices, as has been proven formally by Kim, Mandel and Wolf [28], who felt a need to "reconcile" Harvard and MIT.

We share the opinion of those pragmatic scientists, who regard the two algorithms as useful complements. Jones' approach is certainly more fundamental, i.e. closer to the underlying quantum world. However, it is a bit tedious in usage when it comes to the ensemble averaged observables. This is because the observables are quadratic in field amplitudes; but it is difficult to keep together the pairs of amplitudes to be averaged as a product. Here Stokes, Perrin and Mueller come in: in the Stokes vectors and Perrin-Mueller matrices such products are neatly organized, ready to be averaged. But there is a price to be paid: the common phase of the two components of the state vector is gone. Thus, Stokes-Mueller calculus is suited for problems where amplitude interference between different photon paths can be neglected, or is willingly neglected. Chandrasekhar's treatment of polarized radiative transfer is the most prominent example [14]: the linear relationships between the probabilities greatly simplify the complicated analytics. On the other hand, the most interesting and relevant problems require numerical calculations, and here the difference between Jones and Mueller has faded. Today's computing hardware and software handles 2×2 complex matrices as easily as 4×4 real matrices. In fact, it takes 64 multiplications to calculate a product of two Perrin-Mueller matrices as compared with only 32 multiplications in a product of two Jones matrices.

In Chapter 7 we shall employ the Jones approach for the development of polarized scattering and polarized Monte Carlo simulations. This amplitude based formalism is closer to the fundamental physics, preserves the phase and thus allows one to include path interference. Such interference effects are exploited by several practical analytical techniques, such as imaging speckle interferometry [29] or diffusing wave spectroscopy [30]. These techniques are beyond the scope of the present book, but we would like to keep the path open for future developments. However, we shall use the Stokes-Perrin-Müller concepts (or **S**,**P**-matrix concepts) whenever appropriate, that is, when it comes to averaging and organizing the measured observables in a rational and efficient way.

References

1. Duncan DD, Hollinger JO, and Jacques SL (eds) *Laser-tissue interaction XII: Photochemical, photothermal, and photomechanical*, Vol. 3914 of Proc. SPIE., SPIE (2000).
2. Duncan DD, Jacques SL, and Johnson PC (eds) *Laser-tissue interaction XII: Photochemical, photothermal, and photomechanical*, Vol. 4257 of Proc. SPIE., SPIE, Polarized light interactions (2001).
3. Jacques SL, Ramella-Roman JC, and Lee K. Imaging skin pathology with polarized light. *J. Biomed. Opt.*, 7(3):329–340 (2002).
4. Jackson JD. *Classical electrodynamics*. Wiley, New York, 2nd edition (1975).
5. Landau LD and Lifshitz EM. *Electrodynamics of continuous media*. Pergamon Press, Oxford (1984).
6. Russakov G. A derivation of the macroscopic Maxwell equations. *Am. J. Phys.*, 38:1188–1195 (1970).
7. Lahoz DG and Graham GM. Experimental decision on the electromagnetic momentum expression. *J. Phys. A: Math. Gen.*, 15:303–318 (1982).
8. Richter F, Florian M, and Henneberger K. Poynting's theorem and energy conservation in the propagation of light in media. *Eur. Phys. Lett.*, 81:6705 (2008).
9. Markel VA. Correct definition of the poynting vector in electrically and magnetically polarizable medium reveals that negative refraction is impossible. *Opt. Express*, 16:19152–19168 (2008).
10. Mehrgrübeogl M, Kehtarnavaz N, Rastegar S, and Wang LV. Effect of molecular concentrations in tissue-simulating phantoms onimages obtained using diffuse reflectance polarimetry. *Opt. Express*, 3:286–297 (1998).
11. Kauert M, Frenz M, and Ricka J. Characterization and preparation of focused laser beams. *Opt. Eng.*, 47(1):014201 (2008).
12. Meier M, Romano V, and Feurer T. Material processing with pulsed radially and azimuthally polarized laser radiation. *Appl. Phys. A, Mater. Sci. Process.*, 86(3):329–334 (2007).
13. Jones RC. A new calculus for the treatment of optical systems I. description and discussion of the calculus. *J. Opt. Soc. Am.*, 31(7):488–493 (1941).
14. Subrahmanyan C. *Radiative transfer*. Dover, New York (1960).
15. Sekera Z. Scattering matrices and reciprocity relationships for various representations of state of polarization. *J. Opt. Soc. Am.*, 56(12):1732–1740 (1966).
16. Stokes GG. On the composition and resolution of streams of polarized light from different sources. *Trans. Cambridge Philos. Soc.*, IX:399 (1852).
17. von Laue M. *Rntgenstrahl-Interferenzen*. Becker und Erler, Leipzig (1941).
18. Gisler T, Rüger H, Egelhaaf SU, Tschumi J, Schurtenberger P, and Rička J. Mode-selective dynamic light scattering: Theory vs. experimental realization with single-mode and multimode receivers. *Appl. Opt.*, 34:3546–3553 (1995).
19. Könz F, Rička J, Frenz M, and Fankhauser F II. Single-mode fiber based apparatus for dynamic light scattering in the vitreous. *Opt. Eng.*, 34:2390–2395 (1995).
20. Rička J. Dynamic light scattering with single-mode and multimode receivers. *Appl. Opt.*, 32:2860–2875 (1993).
21. Rowling JK. *Harry Potter*. Bloomsbury, UK (1997).
22. Dehoop AT. A reciprocity relation between transmitting and receiving properties of an antenna. *Appl. Sci. Res.*, 19(2):90–140 (1968).
23. Perrin F. Theory of light scattering by macroscopically isotropic bodies. *J. Chem. Phys.*, 10:415–421 (1942).
24. Bohren CF and Huffman DR. *Absorption and scattering of light by small particles*. Wiley, New York, 2nd print, edition (1998).
25. Bickel WS and Bailey WM. Stokes vectors, Mueller matrices, and polarized scattered-light. *Am. J. Phys.*, 53(5):468–478 (1985).

26. Smith MH. Interpreting Mueller matrix images of tissues. In: DD Duncan, SL Jacques, and PC Johnson (eds) *Laser-tissue interaction XII: Photochemical, photothermal, and photomechanical*, Vol. 4257 of Proc. SPIE, pp. 82–89, (2001).
27. Mueller H. The foundation of optics. *J. Opt. Soc. Am.*, 38:661 (1948).
28. Kim K, Mandel L, and Wolf E. Relationship between Jones and Mueller matrices for random-media. *J. Opt. Soc. Am. A*, 4(3):433–437 (1987).
29. Tuchin VV, Zimnykov DA, Galanzha EI, Ulianov SS, Sinichkin UP, Mishin AB, Chaussksy AA, and Ryabukov VP. Dynamical and structural diagnostics of living tissues using speckle techniques. In: VV Tuchin, JA Izatt, and JG Fujimoto (eds) *Coherence domain optical methods in biomedical science and clinical applications IV*, Vol. 3915 of Proc. SPIE, pp. 137–151 (2000).
30. Pine DJ, Weitz DA, Chaikin PM, and Herbolzheimer E. Diffusing-wave spectroscopy. *Phys. Rev. Lett.*, 60:1134–1137 (1988).

Chapter 5
Monte Carlo Modeling of Light Transport in Tissue (Steady State and Time of Flight)

Steven L. Jacques

5.1 Introduction

Monte Carlo simulations are a fundamental and versatile approach toward modeling light transport in tissues. While diffusion theory for light transport is a fast and convenient way to model light transport, it fails when close to sources or boundaries and when absorption is strong compared to scattering; in other words, whenever conditions cause the gradient of fluence rate (or photon concentration) to not be simply linear but to have some curvature. Monte Carlo steps in to treat problems when diffusion theory fails. Figure 5.1 illustrates a Monte Carlo simulation.

In the general Monte Carlo simulation, "photons" are inserted into tissue at a location defined by *x, y, z* coordinates with a trajectory defined by directional cosines (projection of trajectory onto *x, y* and *z* axes). The random distance traveled before the photon interacts with the tissue is based upon the selection of a random number [0,1] and the local attenuation coefficient of the medium. At the end of each photon step, the weight of the photon is reduced by absorption. The remaining non-absorbed weight is redirected according to a scattering (or "phase") function that describes the angular dependence of single scattering by the particular tissue. Once a new trajectory is specified, the photon is again moved a random distance. The details of the photon path, absorption, scattering, reflection, and refraction are described in the following paragraphs.

Refraction at mismatched boundaries and even changes in local optical properties can be included in the simulation. The heat generated, S [W/cm^3], within a small volume depends upon the total photon weight absorbed in the volume, the total number of photons, and power of the laser beam. Total energy absorbed [J/cm^3] can be computed when the energy [J] of the light source is given.

The Monte Carlo method is a widely used approach toward sampling probability density functions for simulating a wide range of problems. The first use of the

S.L. Jacques (✉)
Biomedical Engineering Department, Oregon Health and Science University, Portland, OR, USA
e-mail: jacquess@ohsu.edu

A.J. Welch, M.J.C. van Gemert (eds.), *Optical-Thermal Response of Laser-Irradiated Tissue*, 2nd ed., DOI 10.1007/978-90-481-8831-4_5,

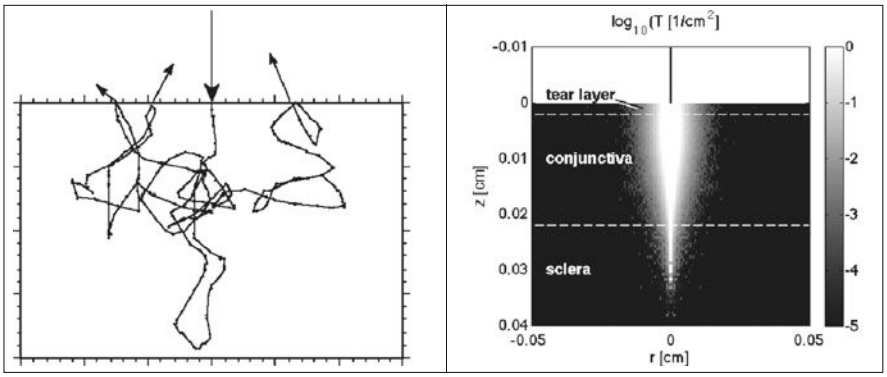


Fig. 5.1 A Monte Carlo simulation of photon transport. (**a**) Movement of a single photon in a light-scattering tissue. When the photon is totally internally reflected at the air/tissue surface, a fraction of the photon weight is allowed to escape as reflectance and the remaining photon weight continues to propagate in the tissue. In this example for the sake of illustration, the third time the photon encounters the surface, the photon is allowed to fully escape. (**b**) Superposition of millions of photons recorded as relative fluence rate, [W/cm^2] per W delivered, which is equivalently stated as transport T [1/cm^2]. The example is the penetration of a CO_2 laser (10.6 μm wavelength) into the conjunctiva of the eye, which illustrates how Monte Carlo can also described the scattering in a situation where absorption dominates over scattering, a situation not well described by diffusion theory

Monte Carlo method for photon transport in biological materials was Adams and Wilson (1983), which considered isotropic scattering [1]. Keijzer et al. (1987) introduced anisotropic scattering into the Monte Carlo simulation of biological tissues, implementing a simulation that propagated photons using cylindrical coordinates, which introduced the Hop/Drop/Spin nomenclature for organizing the program [2]. Prahl et al. (1989) reformulated the program using photon propagation based on Cartesian coordinates, which made the program much simpler to convey in written form [3]. Wang and Jacques (1993) adapted and augmented the work of Keizer and Prahl to write the Monte Carlo Multi-Layered (`MCML`) program that considers tissues with many planar layers with different optical properties [4]. `MCML` is a well-organized program with a simple text input file that the user can modify to specify different problems, which allows various problems to be run without needing to recompile the program each time. `MCML` has been widely promulgated via the web as source code [5–6], and modified by various groups to handle a variety of problems. Jacques (1998) reported on using Monte Carlo to specify the point spread function for light in tissue in planar, cylindrical and spherical coordinates from a plane source, line source or point source, respectively [7], using a minimal Monte Carlo program derived from `MCML` called `mc321.c` that is listed in the Appendix of this chapter for reference and is available on the web [8]. Students have usually modified `mc321.c` to build their own specialized programs. Jacques (2003) prepared a Monte Carlo subroutine, `mcsub.c`, which allows routine calls to Monte Carlo runs

from other programs, discussed in [9], with updated versions listed on the web [10]. Ramella-Roman et al. (2005) modified `mc321.c` to simulate polarized light transport by propagating the Stokes Vector state of a photon [11], and posted the program on the web [12]. This paragraph does not offer a complete review of all contributions to Monte Carlo simulations, but provides a brief history of the programming code that is most widely used in biomedical optics and is freely available on the web.

A small note: throughout this chapter, simple math equations are mixed with programming language, and often an equal sign, "=", is used when an assignment, "←", is the appropriate symbol. The meaning will be obvious, but for those readers familiar with scientific grammar, the poor grammar is intentional.

5.2 Basic Monte Carlo Sampling

The Monte Carlo simulation of light propagation in tissue requires random selection of photon step size, scattering angle and reflection or transmission at boundaries. This is accomplished by a random number [0,1] assigned to the value of a random variable x, such as photon step size. The relationship is established through the density function $p(x)$ and distribution function $F(x)$ (see Chapter 3). Given $p(x)$, the value of the probability distribution function at a particular value x_1 of the random variable x is

$$F(x_1) = \int_0^{x_1} p(x)dx = \text{function}(x_1) \tag{5.1}$$

This $F(x_1)$ is equated with a random number, RND_1, in the interval [0,1]:

$$RND_1 = F(x_1) \tag{5.2}$$

Then Eq. (5.2) is rearranged to solve for x_1 in terms of RND_1. The resulting expression allows a series of values RND_1 to specify a series of values x_1. The histogram of x_1 values will conform to the probability density function $p(x)$. Figure 5.2 illustrates the procedure.

This process is illustrated in the following section by the random selection of photon step size, s.

5.2.1 Predicting the Step Size of a Photon

The first example is predicting the step size of a photon between interaction events (absorption or scattering) as the photon multiply scatters within a tissue. A photon takes a step (s [cm]) whose length is exponentially distributed, and depends on the value of the total attenuation coefficient $\mu_t[\text{cm}^{-1}]$ equal to $\mu_a + \mu_s$, where μ_a is

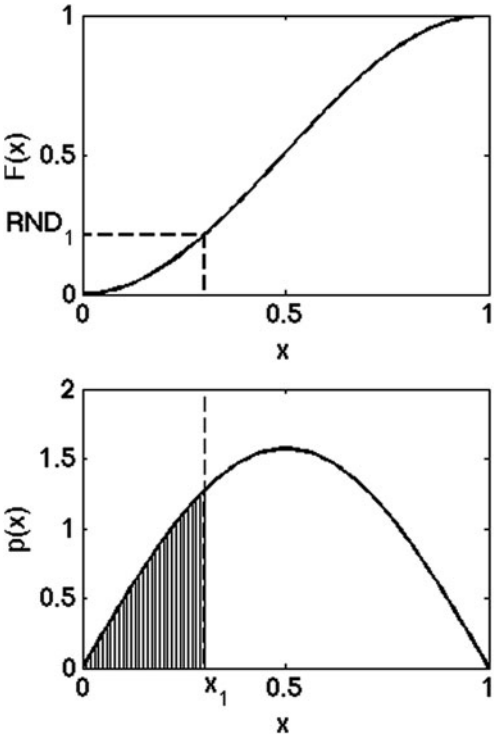


Fig. 5.2 A random number generator [0,1] selects a value RND_1 which is set equal to the distribution function $F(x)$ which then specifies the value x_1. The cross hatched area is equal to RND_1

the absorption coefficient and μ_s is the scattering coefficient. The value $1/\mu_t$ is the photon's mean free pathlength before either absorption or scattering occurs.

The first step is to specify a properly normalized probability density function that describes the probability of a particular photon step size s:

$$p(s) = \exp(-\mu_t s)/\mu_t \tag{5.3}$$

such that

$$\int_0^\infty p(s)ds = \int_0^\infty \frac{e^{-\mu_t s}}{\mu_t} ds = 1 \tag{5.4}$$

The second step is to determine the integral of $p(s)$ evaluated at a particular s_1, which defines the probability distribution function:

$$F(s_1) = \int_0^{s_1} p(s)ds = \int_0^{s_1} \frac{e^{-\mu_t s}}{\mu_t} ds = 1 - e^{-\mu_t s_1} \tag{5.5}$$

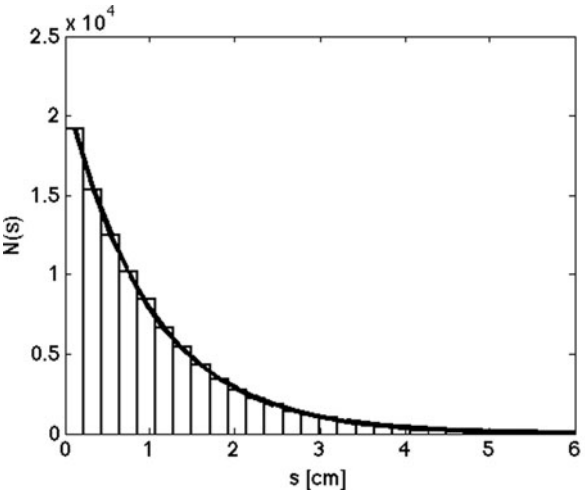


Fig. 5.3 Histogram of step sizes predicted by 1000 random numbers using Monte Carlo sampling (Eq. (5.7)). $N(s)$ is the number of step sizes per bin over a range of s value bins. The value of μ_t is 1 cm^{-1}. The bin size of the histogram is $ds_1 = 0.275$ cm. The curved line is $N(s) = 1000 \exp(-\mu_t s_1) ds_1 / \mu_t$

The third step is to set $F(x)$ equal to a random number RND_1.

$$RND_1 = F(s_1) = 1 - e^{-\mu_t s_1} \tag{5.6}$$

and solve for s_1:

$$s_1 = \frac{-\ln(1 - RND_1)}{\mu_t} = \frac{-\ln(RND_1)}{\mu_t} \tag{5.7}$$

The terms (1 – RND_1) and (RND_1) are equal in a probabilistic sense because of the uniform distribution of the random number between [0,1]. This is the final answer. A series of random numbers RND_1 will generate a series of step sizes s_1 that follow the probability density function $p(s)$ of Eq. (5.3).

To illustrate, Fig. 5.3 shows a histogram of s_1 values specified by 1000 RND_1 values. The value of μ_t is 1 cm^{-1}. The bin size of the histogram is $ds_1 = 0.275$ cm. Also plotted is the curve $N(s) = 1000 \exp(-\mu_t s_1) ds_1$.

5.2.2 *Predicting the Photon Launch Point for a Circular Flat-Field Beam*

The second example is predicting where a photon should be launched at a tissue surface so as to mimic a circular flat-field (i.e., a uniform irradiance) beam of light. Assume the beam of light is delivered perpendicular to the tissue surface. Also assume that we are working in cylindrically symmetric Cartesian coordinates. The Monte Carlo simulation will propagate the photon in 3D using x, y and z coordinates, but we will report out our results only as a function of r and z. Therefore, we need only specify the radial distance r from the center of the beam where a photon

should enter the tissue, and the photon can be launched along the x axis by letting $x = r$. Assume the beam of light has a radius a.

The number of photons per unit area should be constant. For each radial position r there is an annular ring of area $(2\pi r\, dr)$. Therefore, the number of photons to be launched at each choice of r should vary as $(2\pi r\, dr)$. The total area of the light beam is πa^2. The first step is to specify the properly normalized probability density function $p(r)$:

$$p(r) = \frac{2\pi r}{\pi a^2} = \frac{2}{a^2} r \tag{5.8a}$$

such that

$$\int_0^a p(r)dr = \int_0^a \frac{2}{a^2} r dr = 1 \tag{5.8b}$$

The second step is to integrate this $p(r)$, evaluated for a specific r_1, and equate to RND_1.

$$\int_0^{r_1} \frac{2}{a^2} r dr = \frac{r_1^2}{a^2} = RND_1 \tag{5.9}$$

The third step is to solve for r_1 in terms of RND_1:

$$r_1 = a\sqrt{RND_1} \tag{5.10}$$

Now to illustrate, Fig. 5.4 shows the histogram of 1000 r_1 values predicted by 1000 random numbers.

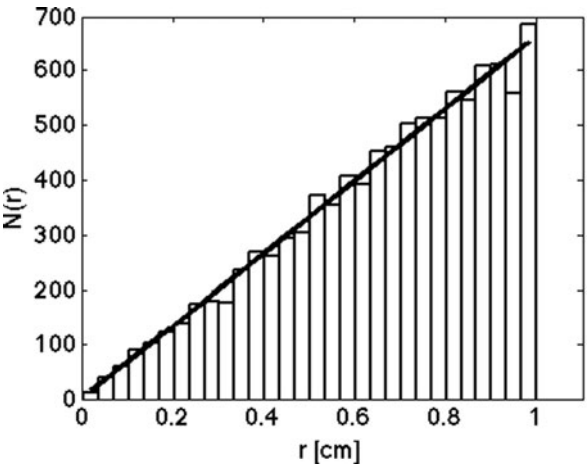


Fig. 5.4 Histogram of radial position of photon launch to achieve a circular flat-field beam of light, perpendicularly illuminating a tissue surface. The histogram is predicted by 1000 random numbers using Monte Carlo sampling (Eq. (5.10)). $N(r)$ is the number of predicted r positions per bin over a range of r value bins. The value of the beam radius a is 1 cm. The bin size of the histogram is $ds_1 =$ 0.0327 cm. The diagonal line is $N(s) = (1000\ 2r\, dr)/a^2$

5.3 The Steady-State Monte Carlo Propagation of Photons in a Tissue

This section presents the basic form of the steady-state Monte Carlo simulation of photon propagation in tissues with optical scattering and absorption properties. The term steady-state means that a stable distribution of light has been achieved, and there are no time dynamics. Section 5.4 will discuss time-resolved Monte Carlo.

It is convenient to consider the Monte Carlo simulation as yielding the fractional density matrix of incident light absorbed, A [1/cm^3], in response to one unit of delivered power or energy. The corresponding heat source matrix S [W/cm^3] or radiant energy density matrix W [J/cm^3] is obtained by scaling A by the beam power P or radiant energy Q, respectively, such that

$$S = PA \tag{5.11a}$$

or

$$W = QA \tag{5.11b}$$

where $\boldsymbol{S}(\boldsymbol{r})$ [W/cm^3] is the local density of power deposition at $r = (x,\ y,\ z)$ in response to a delivered power P [W], and $\boldsymbol{W}(\boldsymbol{r})$ [J/cm^3] is the local density of energy deposition in response to a delivered energy Q [J].

Also, the simulation keeps track of the escaping flux density out of the front surface to which light is delivered, $\boldsymbol{R}_r$ [1/cm^2], and out of the rear surface, $\boldsymbol{T}_r$ [1/cm^2]. The $\boldsymbol{A}$ [1/cm^3] is converted into the fractional transport $\boldsymbol{T}[iz][ir]$[1/cm^2] within the tissue, where iz and ir are the indices of the array T, such that the fluence rate ϕ [W/cm^2] and the fluence ψ [J/cm^2] are

$$\phi(\boldsymbol{r}) = PT(\boldsymbol{r}) \tag{5.12a}$$

$$\psi(\boldsymbol{r}) = QT(\boldsymbol{r}) \tag{5.12b}$$

The following sections will present the algorithm for implementing the Monte Carlo simulation. The overall algorithm is described by the flow diagram in Fig. 5.5.

5.3.1 The Input File

The parameters that govern a Monte Carlo simulation must be passed to the program before it can run. These parameters can be included in the programming at the beginning of a simple Monte Carlo code (as in `mc321.c`, see Appendix), or passed to a Monte Carlo subroutine from another program (as in `mcsub.c` [8, 9]) or read from an input file (as in `MCML` [4–6]). In Fig. 5.5, these possibilities are referred to as the "input file".

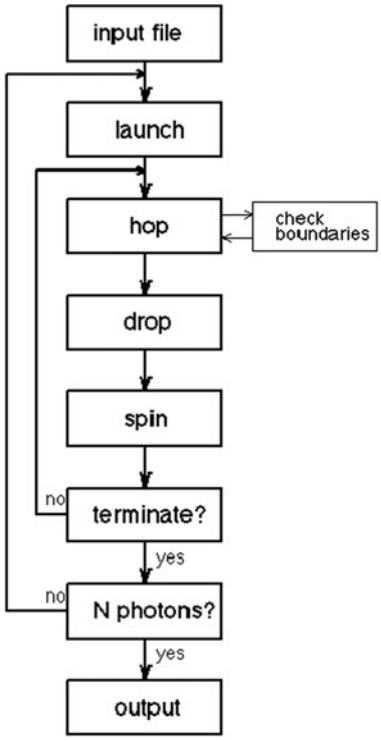


Fig. 5.5 The flow-diagram for a steady-state Monte Carlo simulation

In this simple summary of a generic Monte Carlo simulation, a simple homogeneous tissue is specified. The tissue parameters to be specified are:

$$\begin{array}{lll} \mu_a & [\mathrm{cm}^{-1}] & \text{absorption coefficient of tissue} \\ \mu_s & [\mathrm{cm}^{-1}] & \text{scattering coefficient of tissue} \\ g & [-] & \text{anisotropy of scattering of tissue} \\ n & [-] & \text{refractive index of tissue} \end{array} \tag{5.13}$$

where [–] denotes [dimensionless]. The above parameters specify an infinite optically homogeneous medium, and the simulation will track the 3D movement of a photon over an infinite spatial range. To specify boundaries, for example the front and rear surfaces of a slab of tissue, the following parameters should be specified:

$$\begin{array}{lll} D & [\mathrm{cm}] & \text{thickness of the tissue slab} \\ n_f & [-] & \text{refractive index of external medium outside front of tissue} \\ \mathrm{n_r} & [-] & \text{refractive index of external medium outside rear of tissue} \end{array} \tag{5.14}$$

The position $z = 0$ is aligned with the front surface of the tissue slab, and the rear surface is aligned with $z = \mathrm{D}$. As the program runs, the history of the photon

is recorded by deposition of "photon weight" into spatially distributed bins, either $A(x,\ y,\ z)$ for full 3-D accumulations or $A(r,\ z)$ for circular symmetric laser beams. In the example of this chapter, the bins are denoted $A[iz][ir]$ [photon weight/bin]. The grid of bins for recording this A is independent of the photon propagation, which is implemented in *x, y,* z coordinates. Once a photon is absorbed at *x, y, z*, its position in this example is recorded as z, where $r = \sqrt{x^2 + y^2}$, and some "photon weight" is deposited in $A[iz][ir]$. The parameters to be specified for recording output are:

$$\begin{array}{ll} N_z\ [-] & \text{number of } z \text{ bins (depth)} \\ N_r\ [-] & \text{number of } r \text{ bins (radial position)} \\ dz\ \ [\text{cm}] & \text{size of } z \text{ bins} \\ dr\ \ [\text{cm}] & \text{size of } r \text{ bins} \end{array} \tag{5.15}$$

where the bins cover the extent of the tissue,

$$dz = \frac{D}{N_z} \tag{5.16}$$

Finally, the number of photons that will be launched must be specified,

$$N_{\text{photons}} \quad [-] \quad \text{number of photons to be launched} \tag{5.17}$$

Typically about 10^4–10^8 photons are launched, depending on how much computation time is appropriate or available. The run time will vary with the optical properties. As will be discussed in later sections, each photon will take an average step size of $1/(\mu_s + \mu_a)$ and at each step the photon weight will drop via multiplication by the fraction $\mu_s/(\mu_s + \mu_a)$ called the albedo. The weight is initially 1 upon launching the photon and drops until it reaches a threshold value, THRESHOLD, which is typically 10^{-4}. The time required to reach THRESHOLD is proportional to the number of steps, N_{steps}, taken by the photon

$$\text{THRESHOLD} = \left(\frac{\mu_s}{\mu_s + \mu_a}\right)^{N_{\text{steps}}} \tag{5.18}$$

and the number of steps required for each photon is

$$N_{\text{steps}} = \frac{\ln\,(\text{THRESHOLD})}{\ln\left(\frac{\mu_s}{\mu_s + \mu_a}\right)} \tag{5.19}$$

As the absorption μ_a decreases, N_{steps} increases. As the scattering μ_s increases, N_{steps} increases. The values of g and n have no effect. Therefore, the time required to run N_{photons} is

$$t = N_{\text{photons}}\, t_{\text{per.step}}\, N_{\text{steps}} \tag{5.20}$$

where $t_{\text{per.step}}$ is the time required per photon step in the simulation. One may run several example simulations with known values of THRESHOLD, μ_s, μ_a, and N_{photons}, and record the time required for each simulation. Then fit the data to solve for $t_{\text{per.step}}$:

$$t_{\text{per.step}} = \frac{t}{N_{\text{photons}} N_{\text{steps}}} \tag{5.21}$$

For example, if it takes 23.0 s to run 10^5 photons with optical properties $\mu_a = 1\ \text{cm}^{-1}$, $\mu_s = 100\ \text{cm}^{-1}$, using THRESHOLD $= 10^{-4}$, then $t_{\text{per.step}}$ is 249 ns. Now, if one desires a simulation that takes exactly $t = 10$ min (600 s) for properties $\mu_a = 0.4\ \text{cm}^{-1}$ and $\mu_s = 65\ \text{cm}^{-1}$, then the choice of N_{photons} is

$$\begin{aligned} N_{\text{photons}} &= \frac{t}{t_{\text{per.step}} N_{\text{steps}}} = \frac{t}{t_{\text{per.step}}} \frac{\ln\left(\dfrac{\mu_s}{\mu_s + \mu_a}\right)}{\ln(\text{THRESHOLD})} \\ &= \frac{600s}{249 \times 10^{-9}} \frac{\ln\left(\dfrac{65}{65 + 0.4}\right)}{\ln\left(10^{-4}\right)} = 1.60 \times 10^6 \end{aligned} \tag{5.22}$$

So one would launch 1.60 million photons for a 10-min simulation using the above optical properties. In this manner, the choice of N_{photons} for a given simulation can be made on the basis of available time, regardless of the optical properties in the simulation.

Now the Monte Carlo simulation can proceed to run.

5.3.2 Launching Photons

Photons are launched with an initial weight (w) set to 1.0. As the photon propagates, this weight will be decremented as "photon weight" is "dropped" into the bins $A[iz][ir]$, where iz and ir are index pointers to the bins for the depth z and radial position r, respectively. A large number of photons will be launched (N_{photons}) and later at the end of the program all the weight deposited in $A[iz][ir]$, as well as any weight that escaped the tissue as transmission or reflectance, will be normalized by N_{photons}, such that the final results are reported as the fraction of all delivered light that is either absorbed, reflected or transmitted.

The position of the photon is specified in Cartesian coordinates, (x,y,z), and all the propagation is done in full 3D. The recording of $A[iz][ir]$ is in cylindrical coordinates, but one could implement $A[iz][ix][iy]$ if desired. If one is using 100×100 bins for $A[iz][ir]$, one needs only to fill 10^4 bins with photon weight. If one is using $100 \times 100 \times 100$ bins for $A[iz][ix][iy]$, one needs to fill 10^6 bins with photon weight. Since the signal-to-noise in any bin is roughly proportional to $\sqrt{A}/A$, it takes many more photons to attain good signal-to-noise filling $A[iz][ix][iy]$ than filling

$A[iz][ir]$. This is why cylindrical coordinates are often chosen for use. But there is no reason not to record data as $A[iz][ix][iy]$. The choice of recording does not affect the photon propagation.

The trajectory of the photon will be specified by the trajectory cosines (ux, uy, uz):

$$\begin{aligned} ux &= \cos\theta \cos\varphi \\ uy &= \cos\theta \sin\varphi \\ uz &= \cos\theta \end{aligned} \tag{5.23}$$

where θ is the angle that the trajectory makes with respect to the z axis, and φ is the angle that the trajectory makes with the x axis.

So let's launch some photons.

5.3.2.1 Collimated Launch

In this example, photons are launched perpendicular to the tissue and enter the tissue exactly at the origin $(x, y, z) = (0, 0, 0)$. The photon position and trajectory are

$$\begin{aligned} x &= 0 \\ y &= 0 \\ z &= 0 \end{aligned} \tag{5.24a}$$

and

$$\begin{aligned} ux &= 0 \\ uy &= 0 \\ uz &= 1 \end{aligned} \tag{5.24b}$$

The photon is directed straight downward, so $uz = 1$. The values of ux and uy are zero because no component of the trajectory is directed in the x or y directions.

The case of collimated launch as a flat-field beam, i.e., uniform irradiance, where the beam has a radius a, was discussed in Section 5.2.2. Photons can be placed along the x axis according to Eq. (5.10) with $y = 0$, $z = 0$.

5.3.2.2 Isotropic Point Source

To launch a photon isotropically, i.e., with no preferential direction, at a position located within the tissue, for example at $x = 0$, $y = 0$, and $z = 0.1$ cm, the launch specification is:

$$\begin{aligned} x &= 0 \\ y &= 0 \\ z &= 0.1 \end{aligned} \tag{5.25a}$$

and

$$\begin{aligned} \cos\theta &= 2RND - 1 \\ \sin\theta &= \sqrt{(1-\cos^2\theta)} \\ \varphi &= 2\,\pi\ RND \end{aligned} \tag{5.25b}$$

if ($\varphi < \pi$)

$$\sin\varphi = \sqrt{(1-\cos^2\ \varphi)}$$

else

$$\sin\varphi = -\sqrt{(1-\cos^2\varphi)}$$

such that

$$\begin{aligned} ux &= \sin\theta\cos\varphi \\ uy &= \sin\theta\sin\varphi \\ uz &= \cos\theta \end{aligned} \tag{5.25c}$$

The z position is 0.1 cm. The trajectory (ux,uy.uz) is randomly selected such that

$$\sqrt{ux^2 + uy^2 + uz^2} = 1 \tag{5.26}$$

It is necessary for the total length of the trajectory vector to be unity.

The photon is launched at $x = 0$ and $y = 0$ because we wish to retain cylindrical symmetry. If one launched at any other x,y position the cylindrical symmetry would be broken. Then the results would have to be recorded as $A[iz][ix][iy]$. The results recorded as $A[iz][ir]$ would respond as if photons were launched in a ring of radius $\sqrt{x^2 + y^2}$.

5.3.2.3 Collimated Gaussian Beam

Consider a collimated Gaussian beam that irradiates a tissue, which has a $1/e$ radius of b. The probability density function for the radial position of launch is

$$p(r) = \frac{e^{-(r/b)^2} 2\pi r}{\pi b^2} \tag{5.27}$$

which can be sampled using the Monte Carlo sampling method by

$$r = b\sqrt{-\ln(RND)} \tag{5.28}$$

To launch a Gaussian beam when using cylindrically symmetric results, one can launch the photon at $x = r$. The launch parameters are:

$$\begin{aligned} x &= b\sqrt{-\ln(RND)} \\ y &= 0 \\ z &= 0 \end{aligned} \tag{5.29a}$$

and

$$\begin{aligned} ux &= 0 \\ uy &= 0 \\ uz &= 1 \end{aligned} \tag{5.29b}$$

5.3.2.4 Focused Gaussian Beam

Consider the same Gaussian beam but focused to a focal point within the tissue at depth z_{focus}. Let the focus have a radial Gaussian distribution with a $1/e$ radius of w. Then the launch parameters are calculated:

$$\begin{aligned} x &= w\sqrt{-\ln(RND)} \\ y &= 0 \\ z &= 0 \end{aligned} \tag{5.30a}$$

and

$$\begin{aligned} x_{\text{focus}} &= w\sqrt{-\ln(RND)}\,\text{sign}(2RND - 1) \\ \text{temp} &= \sqrt{((x - x_{\text{focus}})^2 + {z_{\text{focus}}}^2)} \\ \sin\theta &= -(x - x_{\text{focus}})/\text{temp} \\ \cos\theta &= \text{z}_{\text{focus}}/\text{temp} \end{aligned} \tag{5.30b}$$

such that

$$\begin{aligned} ux &= \sin\theta \\ uy &= 0 \\ uz &= \cos\theta \end{aligned} \tag{5.30c}$$

There is a tendency in Monte Carlo simulations using cylindrical coordinates to not get sufficient photons in the bins along the z axis, because the size of each bin is $(2\pi r\ dr\,dz)$, so the size of bins near $r = 0$ is very small, and the likelihood of photon deposition in such a small bin is quite low. During this focused Gaussian launch, one should launch toward both $+x_{\text{focus}}$ and $-x_{\text{focus}}$ positions, which forces the photons to cross the z axis. Doing this causes the bins along the central axis to not be so neglected, and is accomplished by the extra term sign(2 *RND*–1) in the expression for x_{focus}. The function sign () has a value of +1 or −1.

5.3.3 Hop

Now the photon is launched along a trajectory, and takes a step along this trajectory. The standard procedure for taking the step is described as the "Standard Hop" in Section 5.3.3.1. But if the tissue has a front and/or rear boundary, then a second step to "Check boundaries" is taken, as described in Section 5.3.3.2, and if the photon is attempting to escape the tissue, a procedure is used to decide whether the photon escapes or is reflected back into the tissue.

5.3.3.1 Standard Hop

The step size of the photon's step (or hop) must be determined. The step size is calculated:

$$s = -\ln(RND)/\mu_t \tag{5.31}$$

as was described in Section 3.2.1. Now the position of the photon is updated:

$$\begin{aligned} x &= x + s\,ux \\ y &= y + s\,uy \\ z &= z + s\,uz \end{aligned} \tag{5.32}$$

5.3.3.2 Check Boundaries

As part of the Hop step, there is a side box labeled "check boundaries" in Fig. 5.5, which is used when there are front and surface boundaries in the problem. This boundary check is denoted by a side box to emphasize that it is part of the Hop step.

As the photon moves toward the front surface of the tissue and attempts to cross the boundary to escape the tissue, there is a possibility that the photon will be reflected by the surface boundary where the air/tissue interface (or *external medium*/tissue interface) presents a mismatch in refractive indices ($n_f \neq n$ or $n_r \neq n$, for front and rear boundaries, respectively). The method chosen here for handling the boundary is to let a fraction of the photon weight escape the tissue as observable reflectance, and let the remaining fraction of photon weight reflect back into the tissue and continue to propagate.

After taking the step, the position of the photon is checked to see if the photon is still within the tissue or has escaped the tissue:

$$\begin{aligned} &\text{If } z < 0 \\ &\qquad \text{photon is trying to escape} \\ &\text{else} \\ &\qquad \text{photon still within tissue} \end{aligned} \tag{5.33}$$

If the photon is trying to escape, then a partial step is taken along the escaping trajectory that will just reach the boundary surface. The size of the partial step s_1 is

$$s_1 = \text{abs}(z/uz) \tag{5.34}$$

where abs() denotes the absolute value function. First, the photon retracts the full step it took in Hop, which led to escape,

$$\begin{aligned} x &= x - s\,ux \\ y &= y - s\,uy \\ z &= z - s\,uz \end{aligned} \tag{5.35a}$$

and then the photon takes the partial step s_1,

$$\begin{aligned} x &= x + s_1\,ux \\ y &= y + s_1\,uy \\ z &= z + s_1\,uz \end{aligned} \tag{5.35b}$$

Now the photon is located at the boundary surface. Next, a decision is made about how much of the photon weight escapes or how much is reflected.

The probability of reflectance at the boundary is a function of the angle of incidence, encoded as the value uz, and the refractive indices n_1 and n, where n_1 denotes the external medium and equals either n_f or n_r for the front and rear boundaries, respectively. Consider a photon striking the front surface. The internal reflectance, R_i, is calculated using the Fresnel reflection equation:

$$\begin{aligned} R_i = &\frac{(\sin\theta_1\cos\theta_2 - \cos\theta_1\sin\theta_2)^2}{2} \\ &\times \frac{\left((\cos\theta_1\cos\theta_2 + \sin\theta_1\sin\theta_2)^2 + (\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2)^2\right)}{\left((\sin\theta_1\cos\theta_2 + \cos\theta_1\sin\theta_2)^2\,(\cos\theta_1\cos\theta_2 + \sin\theta_1\sin\theta_2)^2\right)} \end{aligned} \tag{5.36}$$

where

n_1	refractive index of incident medium	$= n_{\text{tissue}}$
n_2	refractive index of transmitted medium	$= n_f$
$\cos\theta_1$	incident trajectory	= uz
$\sin\theta_1$	incident trajectory	$= (1 - uz^2)^{1/2}$
$\sin\theta_2$	transmitted trajectory	$= \sin\theta_1(n_1/n_2)$
$\cos\theta_2$	transmitted trajectory	$= (1 - \sin\theta_1^2)^{1/2}$

Once the reflectance R_i is computed, a fraction $(1 - R_i)$ of the current photon weight is allowed to escape and the remaining fraction of weight is reflected back into the tissue to continue propagating. The escape is recorded by placing its remaining weight in the reflectance array bin, $R_r[ir]$,

$$R_r = R_r[ir] + (1 - R_i)w \tag{5.37}$$

where the radial position of escape is r:

$$\begin{aligned} r &= \sqrt{x^2 + y^2} \\ ir &= \text{round}(r/dr) + 1 \end{aligned} \tag{5.38}$$

The function round(r/dr) denotes taking the integer value of the value r/dr, rounding down, such that a photon somewhere within the location of one bin will be assigned to that particular bin. (Note: In this summary, ir extends from 1 to N_z, and ir is not allowed to equal zero. Sometimes, programs let ir extend from 0 to N_z–1. Not mentioned here but important in programming is that a check should be made that ir does not exceed Nr, the allocated size of $Rr[ir]$.)

The remaining fraction of the photon weight is reflected by the boundary,

$$w = R_i w \tag{5.39}$$

and the trajectory with respect to the z axis is reversed,

$$uz = -uz \tag{5.40}$$

The ux and uy components of the trajectory are not changed. Then the photon position is updated by taking the remaining portion of the original step, which equals $s - s_1$:

$$\begin{aligned} y &= (s - s_1)ux \\ y &= (s - s_1)uy \\ z &= (s - s_1)uz \end{aligned} \tag{5.41}$$

Some of the photon weight has escaped the boundary and some has been reflected by the boundary back into the tissue to continue propagating.

The procedure for testing if the photon is attempting to escape the tissue through its rear surface boundary and contributing to $T[iz][ir]$, and for modifying the position and trajectory accordingly, is very similar to the above procedure for the front surface, and is not outlined here. If one wished to place other boundaries in the problem, like a lateral cylindrical boundary as if the tissue were held within a tube (e.g., a glass test tube, or a pipe), this "check boundaries" box is the proper place within the program to implement such special boundaries.

5.3.4 Drop

Arriving at its new position, the photon must interact with the tissue. Upon interaction, a fraction μ_a/μ_t of the photon's weight is absorbed and the remaining μ_s/μ_t fraction of the photon's weight is scattered and continues to propagate. The absorbed fraction is placed in the bin that encloses the current photon position. This process is summarized by the following calculation steps:

$$\begin{aligned} r &= \sqrt{x^2 + y^2} \\ ir &= \text{round}(r/dr) + 1 \\ iz &= \text{round}(z/dz) + 1 \\ A[iz][ir] &= A[iz][ir] + w(\mu_a/\mu_t) \\ w &= w(\mu_s/\mu_t) \end{aligned} \tag{5.42}$$

where ir and iz are index values and dr and dz are bin width and depth, respectively. This finishes the absorption event.

Note that ir and iz should not be allowed to exceed N_r and N_z, respectively, lest one exceed the allocated size of $A[iz][ir]$. Always check this. Often, the last bin is used as an overflow bin and any photon weight that is deposited outside the array is simply accumulated in the last iz,ir bin. At the end of the simulation, the values of weigth in $Rr[ir]$, $Tr[ir]$ and $A[iz][ir]$ are the fraction of total delivered photon weight. The sum of $Rr[ir]$, $Tr[ir]$ and $A[iz][ir]$ over all bins divided by N_{photons} will equal unity. In Section 5.3.7, the proper normalization of these fractional weights by the size of the bins will convert Rr and Tr into the fraction of delivered power or energy escaping per unit surface area and convert A into the fraction of delivered power or energy deposited per unit volume. But the values of Rr, Tr and A determined from the overflow bins are meaningless after normalization.

Next, is the scattering event.

5.3.5 Spin

The photon is scattered into a new trajectory according to some scattering function. The two angles of scatter are θ and φ , the deflection and azimuthal scattering angles, respectively. This section describes how to calculate the new trajectory after sampling the probabilities for θ and φ.

The most commonly used function for the deflection angle θ is the Henyey-Greenstein (HG) function (1941), which was proposed for describing the scattering of light from distant galaxies by galactic dust. The original paper does not offer any explanation for the function, but simply asserts its use. But the HG function is actually very interesting. The function is here expressed as a function of the angle of deflection, θ, and the anisotropy of scattering, g. Using the definitions in Chapter 3 (Section 3.4.3.1) where the probability density functions for $p(\varphi)$ and $p(\theta)$ are considered independently:

$$p(\theta) = \frac{1}{2} \frac{1 - g^2}{\left(1 + g^2 - 2\,g\cos\theta\right)^{3/2}} \tag{5.43}$$

which has the properties that

$$\int_0^{\pi} p\,(\theta)\, 2\pi\, \sin\theta\, d\theta \;=\; 1 \tag{5.44a}$$

and

$$\int_0^{\pi} p(\theta)\cos\theta\, 2\pi\, \sin\theta d\theta = g \tag{5.44b}$$

The last equation is the definition of g. Hence, the HG function is an identity with respect to the definition of g. If you choose a value g to define $p(\theta)$ using Eq. (5.43), the definition of g in Eq. (5.44b) will yield exactly g.

The Monte Carlo sampling of the HG function is specified by the following sequence of calculations:

$$\cos(\theta) = \frac{1 + g^2 - \left(\frac{1 - g^2}{(1 - g + 2gRND)}\right)^2}{2\,g} \tag{5.45}$$

If g is 0, then use $\cos\theta = 2\ RND - 1$. If g is 1.0, then simply let $\cos\theta = 1.0$. Otherwise, use the Monte Carlo sampling in Eq. (5.45).

The azimuthal angle is calculated:

$$\varphi = 2\pi RND \tag{5.46}$$

To update the trajectory based on the values of $\cos\theta$ and φ specified using random numbers, use the following calculations:

$$\begin{aligned}
\sin\theta &= \sqrt{\left(1 - \cos\theta^2\right)} \\
\text{temp} &= \sqrt{\left(1 - uz^2\right)} \\
uxx &= \sin\theta\,(ux\,uz\,\cos\varphi - uy\sin\varphi)/\text{temp} + ux\cos\theta \\
uyy &= \sin\theta\,(uy\,uz\,\cos\varphi + ux\sin\varphi)/\text{temp} + uy\cos\theta \\
uzz &= -\sin\theta\,\cos\varphi\,\text{temp} + uz\cos\theta
\end{aligned} \tag{5.47a}$$

If the trajectory is extremely close to alignment with the z axis, *i.e.* nearly $(uz,\ uy,\ uz) = (0, 0, \pm 1)$, do not use Eq. (5.24) above, but instead use:

$$\begin{aligned}
&uxx = \sin\theta\cos\varphi \\
&uyy = \sin\theta\sin\varphi \\
\text{if } uz \geq 0 \quad &uzz = \cos\theta \\
\text{else} \quad &uzz = -\cos\theta
\end{aligned} \tag{5.47b}$$

Finally, one updates the trajectory:

$$\begin{aligned} ux &= uxx \\ uy &= uyy \\ uz &= uzz \end{aligned}$$

The photon is now oriented along a new trajectory, and ready to take a new step s (see Fig. 5.5).

There are alternative scattering functions. Mie theory is an important scattering function to consider. Perhaps an experiment has yielded a particular scattering function, and one wishes to run a simulation using this function. This chapter will not discuss these alternatives, but as long as the criteria of Eqs. (5.44a, b) are followed, most any scattering function for the deflection angle θ can be used. Sometimes the scattering function does not lend itself to a solution of $\cos\theta$ in terms of a random number, as in Eq. (5.3). Also, sometimes one wishes to consider an azimuthal scattering angle φ that depends on the deflection angle θ, as in Mie scattering of polarized light. In such cases, the "rejection method" is a useful approach [10].

5.3.6 Terminate?

The photon will continue propagating as its weight becomes progressively smaller. How can one stop the photon yet properly conserve energy? The "Roulette Method" is used to terminate the photon. A threshold value (THRESHOLD) is chosen, typically 10^{-4}. When the photon's weight drops below this threshold value, the roulette procedure is employed. A random number (*RND*) is generated; and if this random number is less than a small fraction called CHANCE, typically 0.10, then the photon weight is increased by dividing w by CHANCE. For CHANCE = 0.10, this would be a 10-fold increase in w. Otherwise, the photon is terminated. Consequently, 9 out of 10 times the photon is terminated, but 1 out of 10 times the photon's weight is increased 10-fold and the photon continues to propagate. The result is that photons are usually terminated, but energy is conserved by the occasional surviving photon being given extra weight. Since millions of photons are run, the statistically averaged result is correct. In summary, the roulette method is implemented by the following:

$$\begin{aligned} &\text{if}(w < \text{THRESHOLD}) \\ &\quad \text{if}(RND <= \text{CHANCE}) \\ &\qquad\qquad w = w/\text{CHANCE} \\ &\quad \text{else} \\ &\qquad \text{terminate the photon} \end{aligned} \tag{5.49}$$

Once the photon is terminated, a new photon can be launched. One checks to see if the total number of photons has already reached the maximum number

(N_{photons}) requested by the input file. If not, then a new photon is launched. If yes, the simulation is complete, and it is time to prepare the results for output.

5.3.7 Normalizing Results for Output

Now, all the photons (N_{photons}) have been run, and it is time to save the output of the simulation. The data has been stored in the array A, as $A[iz][ir]$ or perhaps $A[iz][ix][iy]$, in units of [photon weight/bin]. Either way, the key parameter is the volume V [cm^3] associated with each bin. For $A[iz][ir]$, the volumes vary with the value of ir,

$$V = 2\pi(ir - 0.5)dr^2 dz \tag{5.50a}$$

and for $A[iz][ix][iy]$, the volumes are all equal,

$$V = dx\, dy\, dz \tag{5.50b}$$

Then the values A [photon weight/bin] are normalized by the appropriate V and by the value N_{photons} to yield the absorbed fraction, A [1/cm^3], for each pixel:

$$A[ir,\, iz] = \frac{A\,[ir,\, iz]}{V\,[ir,\, iz]\, N_{\text{photons}}} \tag{5.51}$$

The fractional transport, T [1/cm^2], is then calculated as

$$T = \frac{A}{\mu_a} \tag{5.52}$$

Recall from Eqs. (5.12a, b) that fluence rate ϕ [W/cm^2] equals the incident power P [W] times T, $\phi = PT$, and the fluence ψ [J/cm^2] equals the incident energy Q [J] times T, $\psi = QT$.

Keep in mind that the values of A, V and μ_a are specific to each bin, when calculating $T[iz][ir]$ or $T[iz][ix][iy]$. In this summary of a simple implementation of the Monte Carlo method, the μ_a was assumed to be uniform, as well as the scattering properties, so every bin had the same value of μ_a. But in `MCML`, for example, bins at different depths can have different values of μ_a.

The light fluxes that have escaped at the front and rear surface boundaries are similarly normalized, but in this case the surface area AREA rather than the volume V is used. The value of AREA[ir] is $2\pi(ir - 0.5)dr^2$. The array of escaping photons, $R_r[ir]$ [photon weight/bin], is converted to the fractional escaping flux density, $R_r[ir][1/\text{cm}^2]$, by the expression:

$$R_r = \frac{R_r}{AREA\, N_{\text{photons}}} \tag{5.53}$$

This completes the discussion of the steady-state Monte Carlo simulation. The output is the fractional transport, $T\,[iz][ir][1/\text{cm}^2]$, and the fractional escaping flux

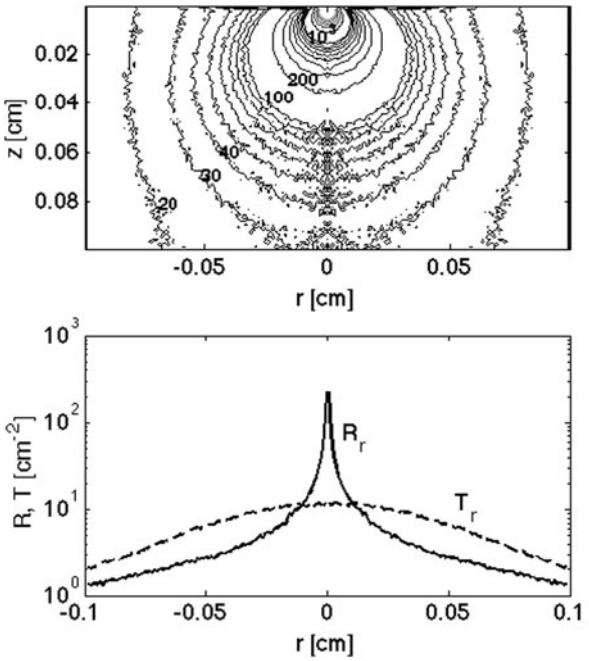


Fig. 5.6 Example steady-state Monte Carlo simulation. The optical properties were $\mu_a = 1\,\text{cm}^{-1}$, $\mu_s = 100\,\text{cm}^{-1}$, $g = 0.90$, $n = 1.4$, for a 1-mm-thick tissue slab with air/tissue boundaries at front and rear surfaces. There were 10^6 photons propagated during a run time of 3 min 49 s on a laptop computer (2 GHz processor). (**a**) Iso-T contours for 20, 30, 40, 100, 200 and 10^3 1/cm^2], T denotes the fractional transport. The noise is evident along the central axis near $r = 0$ because the bins are smaller and collect fewer photons. Running more photons improves the signal-to-noise. (**b**) The reflectance (R_r) and transmission (T_r) as flux densities of escape [1/cm^2] versus radial position r at front and rear surfaces, respectively

densities, $R_r[ir][1/\text{cm}^2]$ and $T_r[ir][1/\text{cm}^2]$. This discussion ends with one example calculation shown in Fig. 5.6, showing T, R_r and T_r for a 1-mm-thick slab of tissue, with light delivered as a pencil beam of collimated light at the origin: $(x, y, z) = (0, 0, 0)$, $(ux, uy, uz) = (0, 0, 1)$.

5.4 Time Resolved Monte Carlo Propagation

Time-resolved Monte Carlo simulation is almost identical to the steady-state simulation discussed above, except for some minor changes. There are actually several ways to implement time-resolved Monte Carlo, and this section shows one approach that illustrates the basic idea.

In this example, the photon is allowed to propagate with no absorption, and the total path of the photon is accumulated after each step s throughout the propagation:

$$L = L + s \tag{5.54}$$

Since there is no absorption, $\mu_a = 0$ and $\mu_t = \mu_s$. Since the speed of light is c, the time duration of the propagation is:

$$t = \frac{L}{c} \tag{5.55}$$

where c is the speed of light in the tissue, $c = c_o/n$ where c_o is the speed of light in vacuo. The time for a photon to escape out the front surface is determined from its total pathlength L at the moment of escape divided by c.

Assume that one is interested in a set of 10 time-points, $t[it]$, where $it = 1 - 10$ is an index that refers to the desired time-point. Let the first time-point, $t[1]$, be 100 ps. Propagation is allowed to continue until the next photon step causes L to exceed the pathlength corresponding to 100 ps,

$$L + s > t[it]/c \tag{5.56}$$

where $t[1] = 100$ ps for this example. At this point, a partial step size, s_1, is taken,

$$s_1 = t[it]/c - L \tag{5.57}$$

The photon is now located exactly at the time-point of 100 ps. The current photon weight w is deposited in the bin $A[iz][ir][it]$, where $[it]$ selects a full 2-D $A[iz][ir]$ array associated with each particular time point. The entire photon weight ($w = 1$) is placed in the local bin, but the weight of the photon is not decremented. The bin $A[iz][ir][it]$ takes a snapshot of the photon's location and weight, but does not affect the photon. There is no absorption. The photon is allowed to continue propagating. The remainder of the step size, $s - s_1$, is taken by the photon. The photon continues to propagate as usual, until a next step causes L to exceed the 2nd time point, $t[2]$. The process of taking a partial step, depositing w into $A[ir][iz][it]$ without changing the w of the photon, completing the step, and resuming propagation is executed. The process continues until L passes the last desired time point, then the photon is terminated and a second photon is launched.

During propagation, when a photon strikes one of the boundaries, a fraction $(1 - R_i)$ of the current photon weight will escape, and the photon weight will decrement. The new photon weight, $R_i w$, will internally reflect and continue propagating. The escaping photon weight will be placed in the bin $R_r[ir][jt]$,

$$R_r[ir][ji] = R_r[ir][ji] + (1 - R_i)w \tag{5.58}$$

where jt is an index that encodes the time of escape, and may be divided into equal time steps, dt [s], that cover the time duration of interest. The current jt is computed:

$$jt = \text{round}(t/dt) + 1 \tag{5.59}$$

As an example of evenly divided time bins, if the time duration of interest was from 10 ps to 1 ns, then *dt* would be 10 ps and *jt* would extend from 1 to 100. An alternative approach is to have progressively larger *dt* bin sizes, so the time base can extend from very short times to very long times, which is not discussed here.

Finally, after N_{photons} have been launched and terminated, it is time to normalize the bins A and R_r for final output. For each time point, $[it]$, conservation of energy in terms of photon weight is summarized:

$$\frac{1}{N_{\text{photons}}}\left(\sum_{iz}\sum_{ir}\boldsymbol{A}[iz][ir][it]+\sum_{jt=1}^{it}R_r[ir][jt]\right)=1 \tag{5.60}$$

The above calculation is not routinely needed, but is only a check that energy is conserved. The final normalization is

$$T=c\frac{A}{V\,N_{\text{photons}}} \tag{5.61}$$

which has units of $[1/(\text{cm}^2\ \text{s})]$. The time-resolved fluence rate, $[\text{W/cm}^2]$ in response to an impulse of energy Q [J] delivered at time zero is:

$$\phi(t)=QT \tag{5.62}$$

The escaping flux density, $R_r[1/(\text{cm}^2\ \text{s})]$, is normalized:

$$R_r=\frac{R_r}{\text{AREA}\ N_{\text{photons}}\,dt} \tag{5.63}$$

Both $A[iz][ir][it]$ and $R_r[ir][it]$ will have good signal-to-noise in regions near the source where most photons spend their time, and poor signal-to-noise in regions far from the source. It is difficult to get good results far from the source even when a large number of photons are launched. Usually, time-resolved Monte Carlo simulation is used to specify results close to a source, and time-resolved diffusion theory is used to specify results at far distances from the source.

Now that the Monte Carlo simulation is completed, absorption can be added to the problem. The attenuation due to absorption is specified by Beer's law which says that photon survival equals $\exp(-\mu_a ct)$, since the ct equals the photon's total pathlength L at any particular time t. Therefore,

$$\phi(t)=QT(t)e^{-\mu_a ct} \tag{5.64}$$

and

$$R_r(t)=\frac{R_r(t)}{\text{AREA}\ N_{\text{photons}}\,dt}e^{-\mu_a ct} \tag{5.65}$$

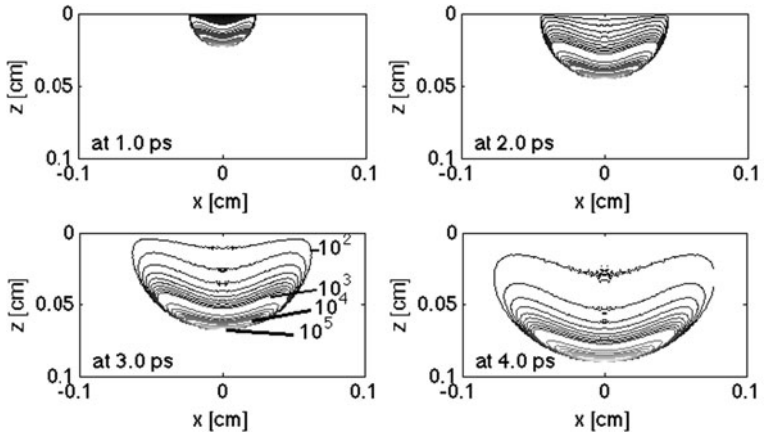


Fig. 5.7 Time-resolved propagation of a collimated laser impulse launched at $r = 0,\ z = 0$ into a 1-mm thickness of a standard tissue. The thickness represents ~10 optical depths for optical properties of $\mu_a = 1\ \text{cm}^{-1},\ \mu_s = 100\ \text{cm}^{-1},\ g = 0.90,\ n = 1.4$. The four time points are 1, 2, 3 and 4 ps. The maps are iso-T contours, where T is the time-resolved fractional transport [1/(cm^2 s)]. The influence of μ_a is minor, for example, $\exp(-\mu_a ct)$ is only 0.92 when $t = 4$ ps

In this way, any value of μ_a can be introduced to learn its influence on the time-resolved distribution of $T(t)$ or $R_r(t)$.

To illustrate time-resolved Monte Carlo, an example of snapshots of $T[iz][ir]$ at 4 time-points is presented in Fig. 5.7, which illustrates the early movement of an incident laser pulse into a tissue. The narrow beam laser pulse broadens with time due to scattering.

Figure 5.8 shows an optical fiber embedded within a tissue and terminating at a depth of 500 μm, delivering light toward the surface. As the light reaches the surface, the time-resolved escape of fractional flux density, $R_r(t, r)$ [1/(cm^2 s)] is shown.

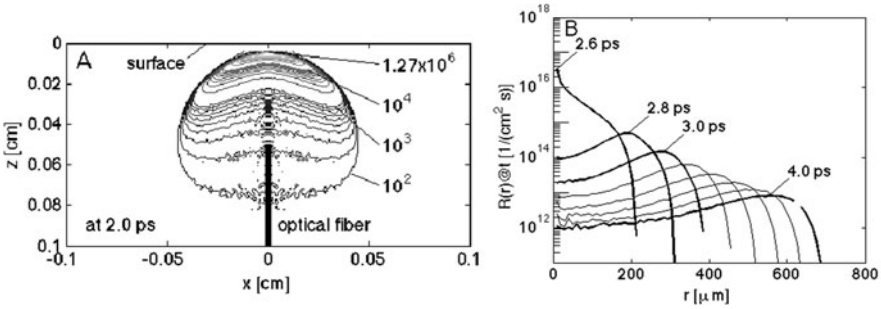


Fig. 5.8 Time-resolved escape of a laser pulse from within tissue. The optical properties were $\mu_s = 100\ \text{cm}^{-1},\ g = 0.90,\ n_{\text{tissue}} = 1.33$. The impulse is delivered toward the surface from an optical fiber terminated at a depth of 500 μm within the tissue and pointed toward the surface. (**a**) Impulse at time 2 ps, shown as T [1/(cm^{2}s)], maximum is 1.27×10^6. (**b**) The escaping fractional flux density versus radial position at different times, $R_r(t, r)$ [1/(cm^2 s)]

5.5 Converting Time-Resolved Results to Frequency-Domain

When the intensity of a light source is modulated at very high frequencies, the ability of the frequency of modulation to transport to some position of observation is described as *frequency domain* light transport. Time-resolved information generated by a time-resolved Monte Carlo simulation can be converted by Fourier Transform into frequency domain information.

Consider the time-resolved escape of fractional flux density, $R_r(t, r)$ [1/(cm^2 s)], when light is delivered as a collimated impulse to position $(r, z) = (0, 0)$ on a tissue surface. The light enters the tissue, but due to scattering begins to escape from the tissue after some delay. This time-resolved $R_r(t, r)$ is shown in Fig. 5.9a, for a typical tissue. As the position of observation moves from 1 mm to 6 mm distance from the source, the time delay before onset of escaping flux increases, and the amount of light escaping decreases.

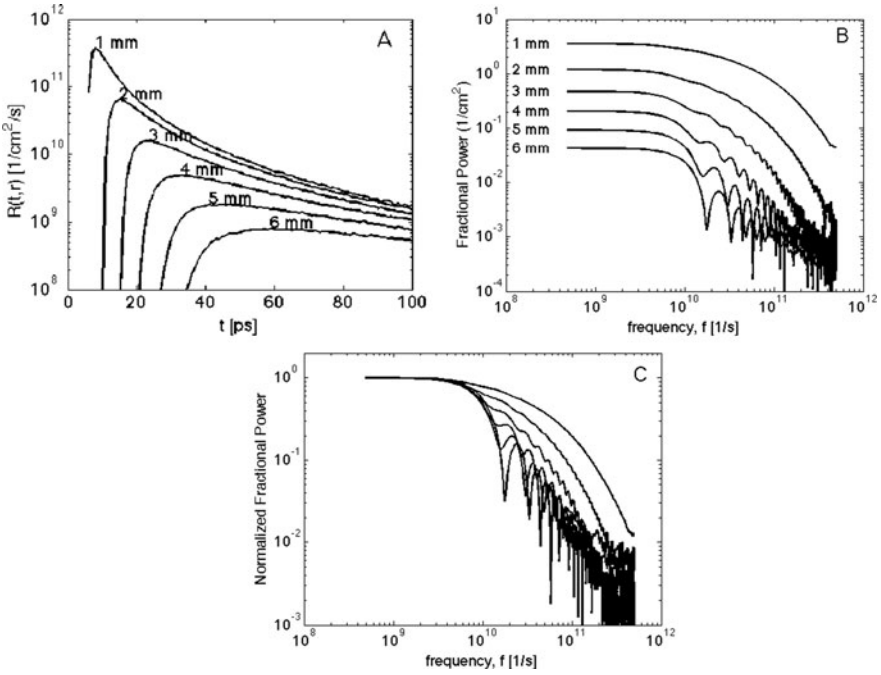


Fig. 5.9 Frequency domain light transport, showing how light modulated at different frequencies and delivered as a collimated beam to position $(r, z) = (0, 0)$, will escape from the tissue as a function of radial position. (**a**) The time-resolved escape of light from the tissue, R_r [1/(cm^2 s)] for r = 1 − 6 mm. (**b**) The fractional power spectrum [1/cm^2] versus frequency f [1/s]. The optical properties of the tissue were $\mu_a = 1\text{cm}^{-1}$, $\mu_s = 100\,\text{cm}^{-1}$, $g = 0.90$, $n = 1.4$, and 10^7 photons were launched during a 35-min simulation. The data at the higher frequencies for the most distant radial positions were based on low photon weights and show artifactual oscillations. (**c**) Normalized fractional power versus frequency

The corresponding frequency domain information is obtained by using a Fast Fourier Transform (*fft*) to convert the time-resolved $R_r(t, r)$ escaping at a particular radial position r, denoted as $R_r(t)$, into the frequency domain:

$$F(f) = \text{abs}(fft\,(R_r(t)dt)) \tag{5.66}$$

where dt is the time-step of the time-resolved data. The absolute value converts the imaginary values generated by the *fft*() into real values that correspond to the power spectrum, expressed as the fractional power $F\,[1/cm^2]$. In other words, if the source was modulated at a frequency f [1/s], the function F would specify the fractional escaping flux, $R_r\,[1/cm^2]$, that was still modulated at frequency f. Figure 5.9b shows this power spectrum. The limiting values toward low frequencies of modulation correspond to the steady-state reflectance $R_r(r)\,[1/cm^2]$, which is why the factor dt was included in the above equation.

To illustrate the above equation more specifically, the equivalent programming code written in MATLAB^{TM} notation is listed:

```
mua = 1;    % absorption coefficient [cm^-1]
mus = 100;  % scattering coefficient [cm^-1]
g   = 0.90; % anisotropy of scattering [dimensionless]
n   = 1.5;  % refractive index [dimensionless]
dt = 1e-12; % time step of time-resolved data [s], in this case dt = 1 ps.
t = (1:100)'*dt;       % time base of Monte Carlo data [s], up to 100 ps
r = (1:100)'/100*1.0;  % radial position of Monte Carlo data [cm], up to 1 cm
Rr = getRrMonteCarlo(t,r,mua,mus,g,n); % get Monte Carlo data [1/cm2/s],
      not shown
N = 2048;  % adds zeros to end of time-resolved data, for padding the transform
f = (1:N/2)'/N/dt; % the x-axis frequency of the power spectrum
ir = 10;   % selects one radial position r(ir)
F = abs(fft(Rr(:,ir)*dt, N));  % Rr(:,ir) is the time-resolved Rr(t) at r(ir)
plot(f, F)  % the plot command yields the power spectrum in [1/cm^2]
```

The original data $R_r(t, r)$ has 100 time-points, but the *fft*() operates best when the number of data points is a multiple of 2. Therefore, zeros are added to the end of the data, which is called *padding*. Adding more zeros causes the result to have more points, so the curves look smoother. In this case 1948 zeros were added to yield a final 2048 data points. The above program yielded Fig. 5.9b. Figure 5.9c normalizes the data by the DC value of R_r, so as to emphasize the shape of the power spectra.

Note in this example that light escaping at 1 mm from the source has higher frequency content than light escaping at 6 mm from the source. In practical frequency domain measurements, the detection of light transport is usually made at least 10 mm from the source, where diffusion theory suffices to describe light transport, and the frequency content of interest is in the hundreds of MHz [1/s]. This example shows how time-resolved Monte Carlo data can be converted to the frequency domain to address questions where diffusion theory is inadequate.

5.6 Summary

Monte Carlo simulations are a relatively simple and flexible method for exploring the behavior of light transport in biological tissues or other media with average absorption and scattering properties. The simulations are like experiments involving a number of photons (N_{photons}) and hence the simulations take time to execute. Some simulations take only a minute, and others take hours or even days. If you are trying to fill small bins with rare photons, the simulations take a lot of time. But common problems take about 10 min or less to get a good result.

The strength of Monte Carlo simulations is to treat situations where diffusion theory, or some other analytic expression of light transport, fails. It is not the proper tool for every job.

5.7 Appendix: `mc321.c`

A simplest version of a Monte Carlo simulation, called `mc321.c`, is listed below. The program can be downloaded from the web [8], but since websites may change, the program is listed here as an archival example, and as an easy reference while reading this chapter. This program does not consider escape across boundaries, but shows the basic format of photon propagation, recording and normalization. The output of this program is plotted in Fig. 5.10:

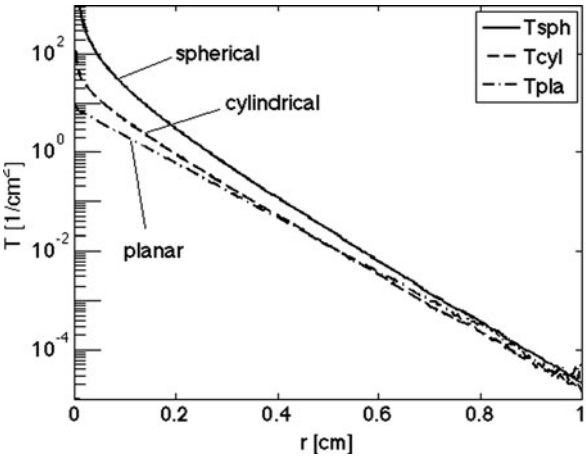


Fig. 5.10 The data in `mc321.out`, produced by `mc321.c`, plotted using a separate graphics program. The spherical transport from a point source, the cylindrical transport from a line source, and the planar transport from a planar source are shown versus the distance, *r*, from the source

```
/*****************************************
 *  mc321.c    , in ANSI Standard C programming language
 *
 *  Monte Carlo simulation yielding spherical, cylindrical, and planar
 *    responses to an isotropic point source (equivalent to a plane source,
```

```
 *    line source, and point source, respectively) in an infinite homogeneous
 *    medium with no boundaries. This program is a minimal Monte Carlo
 *    program scoring photon distributions in spherical, cylindrical,
 *    and planar shells.
 *
 *  by Steven L. Jacques based on prior collaborative work
 *    with Lihong Wang, Scott Prahl, and Marleen Keijzer.
 *    partially funded by the NIH (R29-HL45045, 1991-1997) and
 *    the DOE (DE-FG05-91ER617226, DE-FG03-95ER61971, 1991-1999).
 *
 *  A published report illustrates use of the program:
 *    S. L. Jacques: "Light distributions from point, line, and plane
 *    sources for photochemical reactions and fluorescence in turbid
 *    biological tissues," Photochem. Photobiol. 67:23-32, 1998.
 **********/

#include <math.h>
#include <stdio.h>

#define Nbins           500
#define Nbinsp1         501
#define PI          3.1415926
#define LIGHTSPEED     2.997925E10 /* in vacuo speed of light [cm/s] */
#define ALIVE       1                   /* if photon not yet terminated */
#define DEAD        0                   /* if photon is to be terminated */
#define THRESHOLD   0.01                /* used in roulette */
#define CHANCE      0.1                 /* used in roulette */
#define COS90D      1.0E-6
     /* If cos(theta) <= COS90D, theta >= PI/2 - 1e-6 rad. */
#define ONE_MINUS_COSZERO 1.0E-12
     /* If 1-cos(theta) <= ONE_MINUS_COSZERO, fabs(theta) <= 1e-6 rad. */
     /* If 1+cos(theta) <= ONE_MINUS_COSZERO, fabs(PI-theta) <= 1e-6 rad. */
#define SIGN(x)           ((x)>=0 ? 1:-1)
#define InitRandomGen    (double) RandomGen(0, 1, NULL)
     /* Initializes the seed for the random number generator. */
#define RandomNum        (double) RandomGen(1, 0, NULL)
     /* Calls for a random number from the randum number generator. */

/* DECLARE FUNCTION */
double RandomGen(char Type, long Seed, long *Status);
     /* Random number generator */
main() {

/* Propagation parameters */
double  x, y, z;    /* photon position */
```

```
double  ux, uy, uz; /* photon trajectory as cosines */
double  uxx, uyy, uzz;  /* temporary values used during SPIN */
double  s;          /* step sizes. s = -log(RND)/mus [cm] */
double  costheta;   /* cos(theta) */
double  sintheta;   /* sin(theta) */
double  cospsi;     /* cos(psi) */
double  sinpsi;     /* sin(psi) */
double  psi;        /* azimuthal angle */
double  i_photon;   /* current photon */
double  W;          /* photon weight */
double  absorb;     /* weighted deposited in a step due to absorption */
short   photon_status;  /* flag = ALIVE=1 or DEAD=0 */

/* other variables */
double  Csph[Nbinsp1];  /* spherical   photon concentration CC[ir=0..100] */
double  Ccyl[Nbinsp1];  /* cylindrical photon concentration CC[ir=0..100] */
double  Cpla[Nbinsp1];  /* planar      photon concentration CC[ir=0..100] */
double  Fsph;       /* fluence in spherical shell */
double  Fcyl;       /* fluence in cylindrical shell */
double  Fpla;       /* fluence in planar shell */
double  mua;        /* absorption coefficient [cm^-1] */
double  mus;        /* scattering coefficient [cm^-1] */
double  g;          /* anisotropy [-] */
double  albedo;     /* albedo of tissue */
double  nt;         /* tissue index of refraction */
double  Nphotons;   /* number of photons in simulation */
short   NR;         /* number of radial positions */
double  radial_size;  /* maximum radial size */
double  r;          /* radial position */
double  dr;         /* radial bin size */
short   ir;         /* index to radial position */
double  shellvolume;  /* volume of shell at radial position r */
double  CNT;        /* total count of photon weight summed over all bins */

/* dummy variables */
double  rnd;        /* assigned random value 0-1 */
short   i, j;       /* dummy indices */
double  u, temp;    /* dummy variables */
FILE*   target;     /* point to output file */

/**** INPUT
   Input the optical properties
   Input the bin and array sizes
   Input the number of photons
*****/
```

```
mua          = 1.673;      /* cm^-1 */
mus          = 312.0;  /* cm^-1 */
g            = 0.90;
nt           = 1.33;
Nphotons     = 10000; /* set number of photons in simulation */
radial_size = 2.0;   /* cm, total range over which bins extend */
NR           = Nbins;      /* set number of bins.  */
/* IF NR IS ALTERED, THEN USER MUST ALSO ALTER THE ARRAY DECLARATION TO A
     SIZE = NR+1. */
dr           = radial_size/NR;  /* cm */
albedo       = mus/(mus + mua);

/**** INITIALIZATIONS
*****/
i_photon = 0;
InitRandomGen;
for (ir=0; ir<=NR; ir++) {
   Csph[ir] = 0;
   Ccyl[ir] = 0;
   Cpla[ir] = 0;
   }


/**** RUN
   Launch N photons, initializing each one before progation.
*****/
do {

/**** LAUNCH
   Initialize photon position and trajectory.
   Implements an isotropic point source.
*****/
i_photon += 1;  /* increment photon count */
if ( fmod(i_photon, Nphotons/10)  == 0)
    printf("%0.0f%% done\n", i_photon/Nphotons*100);

W = 1.0;                    /* set photon weight to one */
photon_status = ALIVE;      /* Launch an ALIVE photon */

x = 0;                      /* Set photon position to origin. */
y = 0;
z = 0;

/* Randomly set photon trajectory to yield an isotropic source. */
costheta = 2.0*RandomNum - 1.0;
```

```
sintheta = sqrt(1.0 - costheta*costheta);    /* sintheta is always positive */
psi = 2.0*PI*RandomNum;
ux = sintheta*cos(psi);
uy = sintheta*sin(psi);
uz = costheta;

/* HOP_DROP_SPIN_CHECK
   Propagate one photon until it dies as determined by ROULETTE.
*******/
do {

/**** HOP
   Take step to new position
   s = step size
   ux, uy, uz are cosines of current photon trajectory
*****/
  while ((rnd = RandomNum) <= 0.0);   /* yields 0 < rnd <= 1 */
  s = -log(rnd)/(mua + mus);          /* Step size.  Note: log() is base e */
  x += s * ux;                        /* Update positions. */
  y += s * uy;
  z += s * uz;
/**** DROP
   Drop photon weight (W) into local bin.
*****/
   absorb = W*(1 - albedo);      /* photon weight absorbed at this step */
   W -= absorb;                  /* decrement WEIGHT by amount absorbed */

   /* spherical */
   r = sqrt(x*x + y*y + z*z);    /* current spherical radial position */
   ir = (short)(r/dr);           /* ir = index to spatial bin */
   if (ir >= NR) ir = NR;        /* last bin is for overflow */
   Csph[ir] += absorb;           /* DROP absorbed weight into bin */

   /* cylindrical */
   r = sqrt(x*x + y*y);          /* current cylindrical radial position */
   ir = (short)(r/dr);           /* ir = index to spatial bin */
   if (ir >= NR) ir = NR;        /* last bin is for overflow */
   Ccyl[ir] += absorb;           /* DROP absorbed weight into bin */

   /* planar */
   r = fabs(z);                  /* current planar radial position */
   ir = (short)(r/dr);           /* ir = index to spatial bin */
   if (ir >= NR) ir = NR;        /* last bin is for overflow */
   Cpla[ir] += absorb;           /* DROP absorbed weight into bin */
```

```
/**** SPIN
   Scatter photon into new trajectory defined by theta and psi.
   Theta is specified by cos(theta), which is determined
   based on the Henyey-Greenstein scattering function.
   Convert theta and psi into cosines ux, uy, uz.
*****/
  /* Sample for costheta */
  rnd = RandomNum;
     if (g == 0.0)
        costheta = 2.0*rnd - 1.0;
     else {
        double temp = (1.0 - g*g)/(1.0 - g + 2*g*rnd);
        costheta = (1.0 + g*g - temp*temp)/(2.0*g);
        }
  sintheta = sqrt(1.0 - costheta*costheta); /* sqrt() is faster than sin(). */
  /* Sample psi. */
  psi = 2.0*PI*RandomNum;
  cospsi = cos(psi);
  if (psi < PI)
    sinpsi = sqrt(1.0 - cospsi*cospsi);     /* sqrt() is faster than sin(). */
  else
    sinpsi = -sqrt(1.0 - cospsi*cospsi);

  /* New trajectory. */
  if (1 - fabs(uz) <= ONE_MINUS_COSZERO) {      /* close to perpendicular. */
    uxx = sintheta * cospsi;
    uyy = sintheta * sinpsi;
    uzz = costheta * SIGN(uz);   /* SIGN() is faster than division. */
    }
  else {                                        /* usually use this option */
    temp = sqrt(1.0 - uz * uz);
    uxx = sintheta * (ux * uz * cospsi - uy * sinpsi) / temp + ux * costheta;
    uyy = sintheta * (uy * uz * cospsi + ux * sinpsi) / temp + uy * costheta;
    uzz = -sintheta * cospsi * temp + uz * costheta;
    }

  /* Update trajectory */
  ux = uxx;
  uy = uyy;
  uz = uzz;

/**** CHECK ROULETTE
   If photon weight below THRESHOLD, then terminate photon using Roulette
      technique.
```

```
   Photon has CHANCE probability of having its weight increased by factor of
      1/CHANCE,
   and 1-CHANCE probability of terminating.
*****/
if (W < THRESHOLD) {
   if (RandomNum <= CHANCE)
      W /= CHANCE;
   else photon_status = DEAD;
   }
} /* end STEP_CHECK_HOP_SPIN */
while (photon_status == ALIVE);

  /* If photon dead, then launch new photon. */
} /* end RUN */
while (i_photon < Nphotons);

/**** SAVE
   Convert data to relative fluence rate [cm^-2] and save to file called
     "mcmin321.out".
*****/
target = fopen("mc321.out", "w");

/* print header */
fprintf(target, "number of photons = %f\n", Nphotons);
fprintf(target, "bin size = %5.5f [cm] \n", dr);
fprintf(target, "last row is overflow. Ignore.\n");

/* print column titles */
fprintf(target, "r [cm] \t Fsph [1/cm2] \t Fcyl [1/cm2] \t Fpla [1/cm2]\n");

/* print data:  radial position, fluence rates for 3D, 2D, 1D geometries */
for (ir=0; ir<=NR; ir++) {
    /* r = sqrt(1.0/3 - (ir+1) + (ir+1)*(ir+1))*dr; */
    r = (ir + 0.5)*dr;
    shellvolume = 4.0*PI*r*r*dr; /* per spherical shell */
    Fsph = Csph[ir]/Nphotons/shellvolume/mua;
    shellvolume = 2.0*PI*r*dr;   /* per cm length of cylinder */
    Fcyl = Ccyl[ir]/Nphotons/shellvolume/mua;
    shellvolume = dr;            /* per cm2 area of plane */
    Fpla =Cpla[ir]/Nphotons/shellvolume/mua;
    fprintf(target, "%5.5f \t %4.3e \t %4.3e \t %4.3e \n", r, Fsph, Fcyl, Fpla);
    }

fclose(target);
```

```
} /* end of main */

/* SUBROUTINES */

/*****************************************************************************
 *      RandomGen
 *      A random number generator that generates uniformly
 *      distributed random numbers between 0 and 1 inclusive.
 *      The algorithm is based on:
 *      W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P.
 *      Flannery, "Numerical Recipes in C," Cambridge University
 *      Press, 2nd edition, (1992).
 *      and
 *      D.E. Knuth, "Seminumerical Algorithms," 2nd edition, vol. 2
 *      of "The Art of Computer Programming", Addison-Wesley, (1981).
 *
 *      When Type is 0, sets Seed as the seed. Make sure 0<Seed<32000.
 *      When Type is 1, returns a random number.
 *      When Type is 2, gets the status of the generator.
 *      When Type is 3, restores the status of the generator.
 *
 *      The status of the generator is represented by Status[0..56].
 *
 *      Make sure you initialize the seed before you get random
 *      numbers.
 ****/
#define MBIG 1000000000
#define MSEED 161803398
#define MZ 0
#define FAC 1.0E-9

double RandomGen(char Type, long Seed, long *Status){
  static long i1, i2, ma[56];   /* ma[0] is not used. */
  long        mj, mk;
  short       i, ii;

  if (Type == 0) {              /* set seed. */
    mj = MSEED - (Seed < 0 ? -Seed : Seed);
    mj %= MBIG;
    ma[55] = mj;
    mk = 1;
    for (i = 1; i <= 54; i++) {
      ii = (21 * i) % 55;
      ma[ii] = mk;
      mk = mj - mk;
      if (mk < MZ)
```

```
        mk += MBIG;
      mj = ma[ii];
    }
    for (ii = 1; ii <= 4; ii++)
      for (i = 1; i <= 55; i++) {
        ma[i] -= ma[1 + (i + 30) % 55];
        if (ma[i] < MZ)
          ma[i] += MBIG;
      }
    i1 = 0;
    i2 = 31;
  } else if (Type == 1) {       /* get a number. */
    if (++i1 == 56)
      i1 = 1;
    if (++i2 == 56)
      i2 = 1;
    mj = ma[i1] - ma[i2];
    if (mj < MZ)
      mj += MBIG;
    ma[i1] = mj;
    return (mj * FAC);
  } else if (Type == 2) {       /* get status. */
    for (i = 0; i < 55; i++)
      Status[i] = ma[i + 1];
    Status[55] = i1;
    Status[56] = i2;
  } else if (Type == 3) {       /* restore status. */
    for (i = 0; i < 55; i++)
      ma[i + 1] = Status[i];
    i1 = Status[55];
    i2 = Status[56];
  } else
    puts("Wrong parameter to RandomGen().");
  return (0);
}
#undef MBIG
#undef MSEED
#undef MZ
#undef FAC
```

References

1. Wilson BC and Adam G. A Monte Carlo model for the absorption and flux distributions of light in tissue. *Med. Phys.* 10:824–830 (1983).
2. Keijzer M, Jacques SL, Prahl SA, and Welch AJ. Light distributions in artery tissue: Monte Carlo simulations for finite-diameter laser beams. *Lasers Surg. Med.* 9:148–154 (1989).

3. Prahl, SA, Keijzer, Jacques SL, and Welch AJ. A Monte Carlo model of light propagation in tissue. In: G Müller and D Sliney (eds) *Dosimetry of laser radiation in medicine and biology*, SPIE Series, Vol. IS 5, pp. 102–111 (1989).
4. Wang L-H, Jacques SL, and Zheng L-Q. MCML – Monte Carlo modeling of photon transport in multi-layered tissues. *Comput. Methods Programs Biomed.*, 47:131–146 (1995).
5. Jacques, SL, http://omlc.ogi.edu/software/mc/mcml, Oregon Health & Science University, 2010. This site includes a 178-page manual on MCML. Also, a convolution program, CONV, is available for convolving the point spread functions generated by MCML.
6. Wang LV. Monte Carlo Simulation Package. Modeling of Photon Transport in Multi-layered Tissues (Release 5: MCML 1.2.2 & CONV 1.1). http://labs.seas.wustl.edu/bme/Wang/mc.html, 2010.
7. Jacques SL. Light distributions from point, line, and plane sources for photochemical reactions and fluorescence in turbid biological tissues. *Photochem. Photobiol.* 67:23–32 (1998).
8. Jacques SL. mc321.c. Simple steady-state Monte Carlo program in spherical, cylindrical and planar coordinates (using ANSI standard C). http://omlc.ogi.edu/software/mc/mc321/index.html, 2007.
9. Jacques SL. Monte Carlo simulations of fluorescence in turbid media, Ch. 6. In: MA Mycek and BW Pogue (eds) *Handbook of biomedical fluorescence*. Marcel-Dekker, New York, NY (2003).
10. Jacques, SL, http://omlc.ogi.edu/software/mc/mcsub, Oregon Health & Science University, 2010. This site lists the subroutine mcsub() that can be called by c programs to run a Monte Carlo simulation.
11. Ramella-Roman JC, Prahl SA, and Jacques SL. Three Monte Carlo programs of polarized light transport into scattering media: part I. *Opt. Express* 13(12):4420–4438(2005).
12. Ramella-Roman JC. Polarized light Monte Carlo. http://omlc.ogi.edu/software/polarization/index.html, 2005.

Chapter 6
Diffusion Theory of Light Transport

Willem M. Star

6.1 Introduction

In this chapter light is principally described as particles with energy $h\nu$ and velocity c. These particles are scattered or absorbed by structures in turbid media such as biological tissues and are reflected at boundaries between media with different refractive index, according to the laws of Fresnel. We will only consider monochromatic light, which covers most practical situations in medical and biological laser applications. The formulations are easily extended to polychromatic light as long as scattering effects are elastic. The theory becomes more complicated for inelastic scattering, like in fluorescence. However, even then the diffusion theory is a straightforward extension of the discussions presented in the following sections. Finally, we neglect polarization and interference. To include polarization one would need four diffusion equations instead of one [1]. Polarization of incident light is usually lost in highly scattering media within a millimeter from the surface [2] and is therefore of little concern in diffusion theory. In geometries where single scattering dominates, polarization effects can be important [3], but single scattering is beyond the scope of this chapter.

This chapter starts with a simple derivation of the diffusion equation with simple boundary conditions. This shows the basic physics of the problem. Subsequently, diffusion theory is derived as an approximate solution of the transport equation. This yields relationships between measurable quantities, such as the diffusion constant and the absorption and scattering coefficients. Transport theory also yields expressions for the source terms in the diffusion equation and shows the degree of approximation and how this can be improved.

The formalism of diffusion theory is rather simple. Nevertheless, diffusion theory gives a relatively accurate description of light propagation in highly scattering media. We will see that in simple geometries with only one variable coordinate,

W.M. Star (✉)
Center for Optical Diagnostics and Therapy, Erasmus Medical Center, Rotterdam, The Netherlands
e-mail: wmstar@xs4all.nl

A.J. Welch, M.J.C. van Gemert (eds.), *Optical-Thermal Response of Laser-Irradiated Tissue*, 2nd ed., DOI 10.1007/978-90-481-8831-4_6,

closed-form exact solutions of the diffusion equation can be obtained. An advantage of such solutions is that the effect of changing variables can be easily studied, which helps in understanding the physics of the problem. The general three-dimensional problem can only be solved numerically. However, many numerical methods and computer programs are available. These have been successfully applied to problems of light propagation in biological tissues [4–6] The fact that the three-dimensional problem can be solved at all is a clear advantage of diffusion theory.

One may wonder whether light propagation in biological tissues can be adequately treated by photons whose propagation, including scattering and absorption, is described by transport theory, rather than wave theory. Diffusion theory can be derived from general principles using only macroscopic tissue properties and is therefore expected to hold with the restrictions involved in the approximation (Sections 6.3.4, 6.3.5, and 6.3.6).

6.2 Simple Derivation of Diffusion Theory

6.2.1 Fick's Law

The simplest way to describe the net propagation (net flux) of energy carrying particles in a scattering and absorbing medium is as a diffusion process [7]. Requirements for the approximate validity of the diffusion model are that the optical mean free path ($1/\mu_t$) is much smaller than the typical dimensions of the problem considered, e.g., the distance to boundaries and sources. Furthermore, a photon should be scattered many times before it is absorbed or leaves the medium. The density of photons at position $\boldsymbol{r}$, moving per unit of solid angle in the direction of unit vector $\hat{\boldsymbol{s}}$, is denoted $N(\boldsymbol{r},\hat{\boldsymbol{s}})[\mathrm{m}^{-3}\mathrm{sr}^{-1}]$. Under the conditions mentioned above $N(\boldsymbol{r},\hat{\boldsymbol{s}})$ is nearly uniform in all directions $\hat{\boldsymbol{s}}$ (see also Sections 6.3.5 and 6.3.6).

A net flux of photons results from a gradient in the photon density $N(\boldsymbol{r})[\mathrm{m}^{-3}]$ defined as

$$N_0(\boldsymbol{r}) = \int_{4\pi} N(\boldsymbol{r},\hat{\boldsymbol{s}})\, d\omega \tag{6.1}$$

where $d\omega$ is an element of solid angle about $\hat{\boldsymbol{s}}$. Consequently, a net energy flux results from a gradient in the fluence rate $\phi(\boldsymbol{r}) = N(\boldsymbol{r})hvc$ (see Chapter 3, Section 3.2.2) where h is Planck's constant, c is the speed of light and ν is the photon frequency. Fick's law can be written as

$$F(\boldsymbol{r}) = -D\nabla\phi(\boldsymbol{r})[+\boldsymbol{S}_1(\boldsymbol{r},t)] \tag{6.2}$$

where $\boldsymbol{F}(\boldsymbol{r})$ is the net energy flux vector $[\mathrm{Wm}^{-2}]$ and D is a proportionality constant called the diffusion constant. The vector between brackets is a source term that is usually not included in Fick's law. However, photons from an external light source

may be considered a source of diffuse flux at the point where the first scattering occurs. The actual form of this source term follows from transport theory and can also be obtained from physical arguments (Sections 6.3.3 and 6.3.4).

6.2.2 Energy Conservation and the Diffusion Equation

The diffusion equation follows from Eq. (6.2) and the law of conservation of energy. The energy absorbed per m^3 and per second is $\mu_a(\boldsymbol{r})\phi(\boldsymbol{r})$ (Chapter 3, Section 3.4.3.3). Since the energy lost per m^3 and per second by diffusion equals the divergence of $\boldsymbol{F}$, or $\nabla \cdot \boldsymbol{F}$, the rate of change of energy density is given by

$$\frac{\partial(N_0(\boldsymbol{r},t)h\nu)}{\partial t} = \frac{1}{c}\frac{\partial\phi(\boldsymbol{r},t)}{\partial t} = -\nabla\cdot\boldsymbol{F}(\boldsymbol{r},t) - \mu_a(\boldsymbol{r})\phi(\boldsymbol{r},t) + S_2(\boldsymbol{r},t) \tag{6.3a}$$

where t is the time and $S_2(\boldsymbol{r},t)$ is an additional source term (scalar), representing the contribution to the change in energy density from either a light source distribution within the medium or scattered light from an external light source. $S_2(\boldsymbol{r},t)$ follows from the transport equation but can also be obtained from a simple physical argument (Sections 6.3.3 and 6.3.4). Substitution of Eq. (6.2) into Eq. (6.3a) yields

$$\frac{1}{c}\frac{\partial\phi(\boldsymbol{r},t)}{\partial t} = D\nabla^2\phi(\boldsymbol{r},t) - \mu_a(\boldsymbol{r})\phi(\boldsymbol{r},t) + S_0(\boldsymbol{r},t) \tag{6.3b}$$

This is the diffusion equation where

$$S_0(\boldsymbol{r},t) = -\nabla\cdot\boldsymbol{S}_1(\boldsymbol{r},t) + S_2(\boldsymbol{r},t) \tag{6.4}$$

In this chapter we consider only stationary problems, so that $\partial\phi(\boldsymbol{r},t)\partial t$ is zero and the parameter t is dropped in the equations.

6.2.3 Relationships Between Hemispherical Fluxes, Radiance and Fluence Rate

To have a net transport of energy, the magnitude of the net flux must be greater than zero. We therefore assume that the distribution $N(\boldsymbol{r},\hat{\boldsymbol{s}})$ is slightly anisotropic. That is, the angular dependence of the density of photons as a function of their direction of movement is almost constant thanks to the large number of scattering events. This slight anisotropy is described by a small shift $N_0(\boldsymbol{r})\varepsilon/4\pi\,(\varepsilon << 1)$ of the distribution in a direction given by the unit vector $\hat{\boldsymbol{u}}$ (see Fig. 6.1). The photon density distribution is then given by

$$N(\boldsymbol{r},\hat{\boldsymbol{s}}) = N_0(\boldsymbol{r})[1 + \varepsilon(\boldsymbol{r})\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{u}}]/4\pi \tag{6.5a}$$

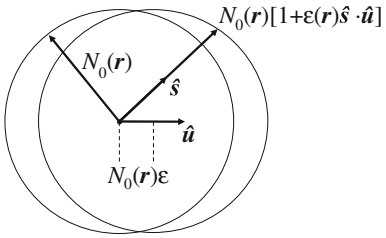


Fig. 6.1 Transport of energy at point $\boldsymbol{r}$. $N(\boldsymbol{r},\hat{\boldsymbol{s}})$ is represented by vectors with the direction $\hat{\boldsymbol{s}}$ and length N. If $N(\boldsymbol{r},\hat{\boldsymbol{s}})$ is isotropic, the same number of photons moves per unit solid angle in all directions $\hat{\boldsymbol{s}}$ and there is no net transport of energy. The distribution can be represented graphically by a sphere with radius $N_0(\boldsymbol{r})$ (we omit the factor 4π in this figure). The black dot is the center of that sphere. If $N(\boldsymbol{r},\hat{\boldsymbol{s}})$ is a nearly isotropic function of $\hat{\boldsymbol{s}}$, there is a net transport of energy in the direction of unit vector $\hat{\boldsymbol{u}}$, according to Eq. (6.5a). This anisotropic distribution can be represented by another sphere, also with radius $N_0(\boldsymbol{r})$, about a new center, displaced by $N_0\varepsilon$ with respect to the center of the isotropic distribution (shift indicated by *dashed lines*)

where the factor $(4\pi)^{-1}$ results from the definition in Eq. (6.1). The radiance is defined as (see Chapter 3, Section 3.2.2.2)

$$L(\boldsymbol{r},\hat{\boldsymbol{s}}) = N(\boldsymbol{r},\hat{\boldsymbol{s}})hvc \quad [\mathrm{Wsr}^{-1}\mathrm{m}^{-2}] \tag{6.5b}$$

Hence, combining Eqs. (6.5a) and (6.5b) yields

$$L(\boldsymbol{r},\hat{\boldsymbol{s}}) = L_0(\boldsymbol{r})[1 + \varepsilon(\boldsymbol{r})\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{u}}] \tag{6.6}$$

Here, $L_0(\boldsymbol{r})$ is defined as the radiance for a perfectly isotropic distribution, that is $L(\boldsymbol{r},\hat{\boldsymbol{s}}) = L_0(\boldsymbol{r})$ for all $\hat{\boldsymbol{s}}$. This is related to $N_0(\boldsymbol{r})$ by $4\pi L_0(\boldsymbol{r}) = N_0(\boldsymbol{r})hvc$. Equation (6.6) basically represents the first two terms of an expansion of $L(\boldsymbol{r},\hat{\boldsymbol{s}})$ in Legendre polynomials (see Appendix 2) and is called the Eddington approximation [8], which is additional to the diffusion approximation given by Eq. (6.2). See also Sections 6.2.4 and 6.3.6. The net energy flux is obtained by integrating $L(\boldsymbol{r},\hat{\boldsymbol{s}})\hat{\boldsymbol{s}}$ over 4π steradians (Chapter 3, Section 3.2.2.4):

$$\boldsymbol{F}(\boldsymbol{r}) = \int_{4\pi} L(\boldsymbol{r},\hat{\boldsymbol{s}})\hat{\boldsymbol{s}}\, d\omega \tag{6.7}$$

Substituting for $L(\boldsymbol{r},\hat{\boldsymbol{s}})$ using Eq. (6.6) we obtain

$$\boldsymbol{F}(\boldsymbol{r}) = \int_{4\pi} L_0(\boldsymbol{r})\hat{\boldsymbol{s}}[1 + \varepsilon(\boldsymbol{r})\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{u}}]d\omega = L_0(\boldsymbol{r})\int_{4\pi}\hat{\boldsymbol{s}}\, d\omega + L_0(\boldsymbol{r})\varepsilon(\boldsymbol{r})\int_{4\pi}\hat{\boldsymbol{s}}(\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{u}})\, d\omega \tag{6.8}$$

The first of the last two integrals is zero and the second equals $(4\pi/3)\hat{\boldsymbol{u}}$. The latter can be verified by integrating each component of the vector separately and

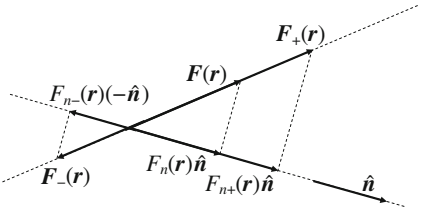


Fig. 6.2 Illustration of the definition of hemispherical and net energy fluxes and their components. $\boldsymbol{F}(\boldsymbol{r})$ is the net energy flux vector, $\boldsymbol{F}_+(\boldsymbol{r})$ and $\boldsymbol{F}_-(\boldsymbol{r})$ are the hemispherical flux vectors in the direction of $\boldsymbol{F}(\boldsymbol{r})$ (i.e. unit vector $\hat{\boldsymbol{u}}$ in Eq. (6.9)). $F_n(\boldsymbol{r})\hat{\boldsymbol{n}}$ is the projection of $\boldsymbol{F}(\boldsymbol{r})$ along the unit vector $\hat{\boldsymbol{n}}$, $F_{n+}(\boldsymbol{r})\hat{\boldsymbol{n}}$ and $F_{n-}(\boldsymbol{r})(-\hat{\boldsymbol{n}})$ are the projections of the hemispherical flux vectors $\boldsymbol{F}_+(\boldsymbol{r})$ and $\boldsymbol{F}_-(\boldsymbol{r})$ respectively. The relationships between $F_n(\boldsymbol{r})$, $F_{n+}(\boldsymbol{r})$ and $F_{n-}(\boldsymbol{r})$ are given by Eqs. (6.13a), (6.13b) and (6.14), respectively (see also Chapter 3, Section 3.2.2.4)

noting that the only non-zero term in the integral is $\int_{4\pi} \hat{s}_x \hat{s}_x \hat{u}_x \, d\omega = (4\pi/3)\hat{u}_x$ and similarly for the y and z components (for vector integrals see also Chapter 1). So Eq. (6.8) reduces to

$$\boldsymbol{F}(\boldsymbol{r}) = (4\pi/3) L_0(\boldsymbol{r}) \varepsilon(\boldsymbol{r}) \hat{\boldsymbol{u}} \tag{6.9}$$

Note that $\boldsymbol{F}(\boldsymbol{r})$ is in the direction $\hat{\boldsymbol{u}}$. Substitution of $L_0(\boldsymbol{r})\varepsilon(\boldsymbol{r})\hat{\boldsymbol{u}}$ in Eq. (6.6) yields

$$L(\boldsymbol{r}, \hat{\boldsymbol{s}}) = L_0(\boldsymbol{r}) + (3/4\pi)\boldsymbol{F}(\boldsymbol{r}) \cdot \hat{\boldsymbol{s}} \tag{6.10}$$

where $4\pi L_0(\boldsymbol{r})$ equals the fluence rate $\phi(\boldsymbol{r})$ since $\phi(\boldsymbol{r}) = \int_{4\pi} L(\boldsymbol{r}, \hat{\boldsymbol{s}}) \, d\omega = L_0(\boldsymbol{r}) \int_{4\pi} d\omega = 4\pi L_0(\boldsymbol{r})$.

Hemispherical energy fluxes in the positive and negative directions of an arbitrary unit vector $\hat{\boldsymbol{n}}$ are defined in Chapter 3, Section 3.2.2.4. See also Fig. 6.2 and the explanation in its legend. The equations for $F_{n+}(\boldsymbol{r})$ and $F_{n-}(\boldsymbol{r})$ are

$$F_{n+}(\boldsymbol{r}) = \int\limits_{\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{n}} > 0} L(\boldsymbol{r}, \hat{\boldsymbol{s}})(\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{n}}) \, d\omega \tag{6.11a}$$

and

$$F_{n-}(r) = -\int\limits_{\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{n}} < 0} L(\boldsymbol{r}, \hat{\boldsymbol{s}})(\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{n}}) \, d\omega \tag{6.11b}$$

Substituting the radiance (Eq. (6.10)) into Eqs. (6.11a) and (6.11b) produces

$$F_{n+}(\boldsymbol{r}) = \int\limits_{\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{n}} > 0} L_0(\boldsymbol{r})(\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{n}}) \, d\omega + \int\limits_{\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{n}} > 0} (3/4\pi)\boldsymbol{F}(\boldsymbol{r}) \cdot \hat{\boldsymbol{s}}(\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{n}}) \, d\omega \tag{6.12a}$$

and

$$F_{n-}(\boldsymbol{r}) = -\int\limits_{\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{n}}<0} L_0(\boldsymbol{r})(\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{n}})\,d\omega - \int\limits_{\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{n}}<0} (3/4\pi)\boldsymbol{F}(\boldsymbol{r})\cdot\hat{\boldsymbol{s}}(\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{n}})\,d\omega \tag{6.12b}$$

where $\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{n}} = \cos\theta$ and $d\omega = \sin\theta d\theta d\varphi$ if the z axis is chosen parallel to $\hat{\boldsymbol{n}}$ (see Chapter 3 for the notation of the angles). Noting that the hemispherical integral $\int\limits_{\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{n}}>0} (\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{n}})\,d\omega = \pi$ (and $= -\pi$ when integrating over $(\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{n}}) < 0$) and that $\int\limits_{\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{n}}>0} (\boldsymbol{A}\cdot\hat{\boldsymbol{s}})(\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{n}}) = +(2/3)(\boldsymbol{A}\cdot\hat{\boldsymbol{n}})$ (the same result upon integration over $(\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{n}}) < 0$), Eqs. (6.12a) and (6.12b) become, respectively,

$$F_{n+}(\boldsymbol{r}) = \pi L_0(\boldsymbol{r}) + \boldsymbol{F}(\boldsymbol{r})\cdot\hat{\boldsymbol{n}}/2 = \phi(\boldsymbol{r})/4 + \boldsymbol{F}(\boldsymbol{r})\cdot\hat{\boldsymbol{n}}/2 \tag{6.13a}$$

and

$$F_{n-}(\boldsymbol{r}) = \pi L_0(\boldsymbol{r}) - \boldsymbol{F}(\boldsymbol{r})\cdot\hat{\boldsymbol{n}}/2 = \phi(\boldsymbol{r})/4 - \boldsymbol{F}(\boldsymbol{r})\cdot\hat{\boldsymbol{n}}/2 \tag{6.13b}$$

Subtracting Eq. (6.13b) from Eq. (6.13a) gives, as expected,

$$F_{n+}(\boldsymbol{r}) - F_{n-}(\boldsymbol{r}) = F(\boldsymbol{r})\cdot\hat{\boldsymbol{n}} = F_n(\boldsymbol{r}) \tag{6.14}$$

Adding Eqs. (6.13a) and (6.13b) gives a relation between the fluxes and the fluence rate when no external sources are present, that is the fluence rate for scattered light $\phi_s(\boldsymbol{r})$:

$$\phi_s(\boldsymbol{r}) = 4\pi L_0(\boldsymbol{r}) = 2[F_{n+}(\boldsymbol{r}) + F_{n-}(\boldsymbol{r})] \tag{6.15}$$

This relationship, which is characteristic of diffusion theory in the Eddington approximation, is independent of the direction of the vector $\hat{\boldsymbol{n}}$.

6.2.4 *Solution of the Source Free Diffusion Equation in a Simple Geometry*

In a stationary situation far from boundaries and sources Eq. (6.3b) reduces to

$$D\nabla^2\phi(\boldsymbol{r}) = \mu_a(\boldsymbol{r})\phi(\boldsymbol{r}) \tag{6.16}$$

Consider a semi infinite medium illuminated by a wide parallel light beam along the z-direction, with $z = 0$ at the surface. Deep inside the medium, say ten penetration depths (Chapter 3, Section 3.4.1.1), Eq. (6.16) applies with only one variable coordinate z and the solution is

$$\phi(z) = \phi(0)\exp(-(\mu_a/D)^{1/2}z) \tag{6.17}$$

where $\phi(0)$ is determined by the intensity if the incident beam, the optical properties of the tissue and by the boundary condition at the surface where the beam enters the medium (see also Sections 6.2.5 and 6.3.8). In the exponential of Eq. (6.17) the effective attenuation coefficient is defined as $\mu_{\text{eff}} = (\mu_a/D)^{1/2}$ (see also Section 6.3.4, Eq. (6.35)). From Eqs. (6.2), (6.9) and (6.17) we obtain an expression for ε in Eq. (6.6), i.e. $\varepsilon = 3(D\mu_a)^{1/2}$ so that Eq. (6.6) becomes

$$L(z, \mu) = L_0(z)[1 + 3(D\mu_a)^{1/2}\mu] \tag{6.18}$$

where $\mu = \cos\theta$ (Fig. 3.5) and $L_0(z) = \phi(z)/(4\pi)$. According to Eq. (6.18), far away from boundaries and sources, the angular dependence of the radiance does not depend on depth. This angular dependence is called the diffusion pattern [9]. In the Eddington approximation the diffusion pattern is only slightly anisotropic (Eq. (6.18)). For isotropic and linear anisotropic scattering (Chapter 3, Section 3.4.3.2) the diffusion pattern can be solved exactly (Section 12 and Appendix III of Ref. [1]). Even for anisotropic scattering the diffusion pattern can be nearly isotropic if the albedo is close to one. On the other hand, even for isotropic scattering the diffusion pattern can be strongly anisotropic if the albedo is well below one. In tissue optics, the fluence rate is much more important than the radiance or the diffusion pattern.

6.2.5 Boundary Conditions

Equation (6.3) is sufficient for approximate calculations of light distributions in tissues, if the source term $S_0(\boldsymbol{r})$ is known and proper boundary conditions are incorporated. The latter can be derived from the law of energy conversation. Let $\hat{\boldsymbol{n}}$ be a unit vector pointing inward at a point on the boundary and perpendicular to it (Fig. 6.3). Initially let us assume that the scattering medium is surrounded by a non-scattering medium with the same refractive index ($n_2 = n_1$). We further assume that diffuse light arises from scattering in the medium and that there are no external sources of diffuse light. By definition, in Eq. (6.11a) the energy flux of scattered light entering the medium at point $\boldsymbol{r}$ of the boundary equals $F_{n+}(\boldsymbol{r})$. Therefore, since no scattered light is assumed to enter the medium, energy conservation requires that $F_{n+}(\boldsymbol{r}) = 0$ at the boundary.

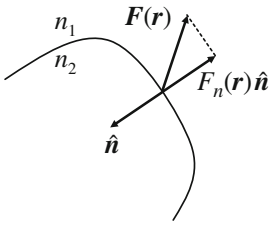


Fig. 6.3 Illustration of the boundary condition. See also Fig. 6.2 for the symbols. The clear medium has refractive index n_1 and the scattering medium has refractive index $n_2 \geq n_1$. $F_n(\boldsymbol{r}) = F_{n+}(\boldsymbol{r}) - F_{n-}(\boldsymbol{r}) = -(1 - r_{21})F_{n-}(\boldsymbol{r})$

One often deals with an air-tissue or water-tissue boundary, where the refractive indices do not match. Here, a hybrid treatment is necessary. Within the medium, light is described by particles. At the boundary, reflection and refraction occur (Chapter 3, Section 3.3), which are wave phenomena. Let the refractive index of the clear medium be n_1 and that of the scattering medium (tissue) n_2, with $n_2 > n_1$. If we define $n = n_2/n_1$, the critical angle for light in the tissue traveling to the surface is $\theta_c = \arcsin(1/n)$. A scattered photon striking the boundary at an angle θ with the normal larger than θ_c is totally reflected. For $\theta < \theta_c$ only a few percent of the photons is reflected (except if $\theta \approx \theta_c$, see also Fig. 6.5). Therefore, at an index-mismatched boundary $F_{n+}(\boldsymbol{r}) \neq 0$. We introduce a reflection factor r_{21} for scattered light in medium 2 striking the boundary with medium 1. The boundary condition is written as: $F_{n+}(\boldsymbol{r}) = r_{21}F_{n-}(\boldsymbol{r})$. The form of r_{21} and the conditions for the description of boundary conditions by reflection factors will be discussed in detail in Sections 6.3.8 and 6.3.9.

Exercise 6.1 Suppose human skin is illuminated by light from an overcast sky. In that case the incident light is largely diffuse. For simplicity, apply a 633 nm filter. Let the irradiance be E_0. Equation (6.17) then holds for all $z \geq 0$. Using the boundary condition above, Eq. (6.2) without source, and Eq. (6.15), derive a formula for $\phi(0)$. Hint: $F_+(0) = E_0 + r_{21}F_-(0)$. Calculate the value of $\phi(0)$ in terms of E_0 if $r_{21} = 0.49$, $\mu_{\text{eff}} = 0.27\ \text{mm}^{-1}$ and $D = 0.087$ mm (optical properties derived from Ref. [15]).

Result: $\phi(0) = 4E_0/[(1 - r_{21})(1 + 2A(\mu_a D)^{1/2})] = 4.14E_0$, where $A = (1 + r_{21})/(1 - r_{21})$

6.2.6 Discussion

In the previous sections the diffusion equation has been derived from general macroscopic principles. This means that diffusion theory is probably a good model if the requirements regarding μ_a and μ_s are satisfied. These requirements follow from the derivation of diffusion theory from the transport equation in the following sections (see Sections 6.3.4, 6.3.5, and 6.3.6). The validity of diffusion theory near boundaries depends on a proper formulation of the boundary conditions.

6.3 Diffusion Approximation in Transport Theory

6.3.1 Introduction

In the previous sections diffusion theory was introduced in a simple way to illustrate the physics of the problem. For further discussion it is useful to treat diffusion theory as an approximate solution of the transport equation. This will yield a relationship between the diffusion constant D and the optical properties of the medium, as well as expressions for the source terms $\boldsymbol{S}_1(\boldsymbol{r})$ and $S_2(\boldsymbol{r})$ (Section 6.3.4).

6.3.2 Transport Equation

The transport equation is known as the Boltzmann equation in statistical mechanics [10, 11]. When used to treat the propagation of light, it is often called the equation of radiative transfer, or equation of transfer [1]. The transport equation is an integro-differential equation for the radiance. A brief derivation is given in Chapter 3, Section 3.4.3. Its general stationary form, neglecting polarization, is

$$\hat{\boldsymbol{s}} \cdot \nabla L(\boldsymbol{r}, \hat{\boldsymbol{s}}) + \mu_t(\boldsymbol{r}) L(\boldsymbol{r}, \hat{\boldsymbol{s}}) = \mu_s \int_{4\pi} p(\hat{\boldsymbol{s}}, \hat{\boldsymbol{s}}') L(\boldsymbol{r}, \hat{\boldsymbol{s}}')\, d\omega' + S(\boldsymbol{r}, \hat{\boldsymbol{s}}) \tag{6.19}$$

Here, $\mu_t = \mu_a + \mu_s$ is the total attenuation coefficient [m^{-1}], $p(\hat{\boldsymbol{s}}, \hat{\boldsymbol{s}}')$ is the scattering or phase function (Chapter 3, Section 3.4.3.1) and $S(\boldsymbol{r}, \hat{\boldsymbol{s}})$ is a source term [$\mathrm{Wm}^{-1} \cdot \mathrm{sr}$]. We will assume that the scattering probability depends only on the (cosine of) the angle between $\hat{\boldsymbol{s}}$ and $\hat{\boldsymbol{s}}'$, that is $p(\hat{\boldsymbol{s}}, \hat{\boldsymbol{s}}') = p(\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{s}}')$. Particular forms of the source term will be discussed in Sections 6.4.1 and 6.4.2. Remember that $d\omega$ symbolizes an element of solid angle (scalar) about $\hat{\boldsymbol{s}}$. We will assume that the materials considered are homogeneous, so that the optical properties are independent of position $\boldsymbol{r}$.

6.3.3 Separation of Scattered from Non-scattered Light; The Source Term

The most common light sources in medical laser applications are an external collimated beam or an optical fiber touching the tissue or a fiber with either a spherical or a cylindrical diffusing tip. This tip may be positioned in a cavity such as the bladder or may be interstitially implanted in the tissue. The radiance at any point in tissue consists of a "primary" part, which is non-attenuated light from the external or internal sources (denoted by subscript p), and scattered light (denoted by subscript s):

$$L(\boldsymbol{r}, \hat{\boldsymbol{s}}) = L_p(\boldsymbol{r}, \hat{\boldsymbol{s}}) + L_s(\boldsymbol{r}, \hat{\boldsymbol{s}}) \tag{6.20}$$

By definition, the fluence rate also consists of a primary and a scattered part:

$$\phi(\boldsymbol{r}) = \int_{4\pi} L(\boldsymbol{r}, \hat{\boldsymbol{s}})\, d\omega = \int_{4\pi} L_p(\boldsymbol{r}, \hat{\boldsymbol{s}})\, d\omega + \int_{4\pi} L_s(\boldsymbol{r}, \hat{\boldsymbol{s}})\, d\omega = \phi_p(\boldsymbol{r}) + \phi_s(\boldsymbol{r}) \tag{6.21}$$

A photon is removed from the primary beam by absorption or scattering. After the first scattering event, a photon is transferred from the primary beam to the scattered part, that is, from L_p to L_s. These "first scattering" events form the source of scattered light. Primary light is attenuated according to $\exp(-\mu_t \ell)$, where ℓ is the path length of non-attenuated photons in tissue between the point of entrance and position $\boldsymbol{r}$. See Fig. 6.4 for geometry and explanation of notation. The local

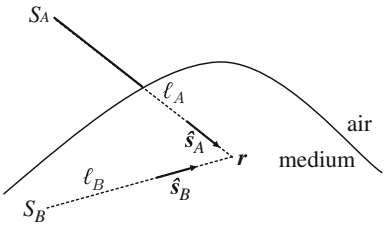


Fig. 6.4 Attenuation of primary light generated outside or inside a medium. The direction of propagation is given by unit vectors $\hat{\boldsymbol{s}}_A$ and $\hat{\boldsymbol{s}}_B$. A source S_A outside the medium may be a discrete source or a laser beam. If the unattenuated irradiance of S_A at point $\boldsymbol{r}$ is $E_A(\boldsymbol{r},\hat{\boldsymbol{s}}_A)$, the attenuated irradiance is $E_A(\boldsymbol{r},\hat{\boldsymbol{s}}_A)\exp(-\mu_t\ell_A)$, where ℓ_A is the path length in tissue (dashed) and μ_t is the attenuation coefficient. If S_B is a discrete source or an element of a source inside the medium and the unattenuated irradiance at point $\boldsymbol{r}$ is $E_B(\boldsymbol{r},\hat{\boldsymbol{s}}_B)$, the attenuated irradiance at point $\boldsymbol{r}$ is similarly $E_B(\boldsymbol{r},\hat{\boldsymbol{s}}_B)\exp(-\mu_t\ell_B)$. If S_B is an isotropic point source with power P, $E_B = P/(4\pi\ell_B^2)$. Note that refraction and reflection may occur at the surface if the refractive indices do not match. In that case the line from S_A to $\boldsymbol{r}$ is not straight and reflection must be included in the formulas for the attenuated irradiance

non-scattered and non-absorbed irradiance is $E(\boldsymbol{r},\hat{\boldsymbol{s}}_0) = E_0(\boldsymbol{r},\hat{\boldsymbol{s}}_0)\exp(-\mu_t\ell)$. Here, the subscript 0 replaces subscripts A and B in Fig. 6.4, $E_0(\boldsymbol{r},\hat{\boldsymbol{s}}_0)$ is the irradiance at point $\boldsymbol{r}$ in the absence of tissue and $\hat{\boldsymbol{s}}_0$ is a unit vector along the line connecting the source (or an element of the source) and point $\boldsymbol{r}$. The quantity $E(\boldsymbol{r},\hat{\boldsymbol{s}}_0)$ thus contributes in two ways to the fluence rate at point $\boldsymbol{r}$: first as the primary fluence rate, that is $\phi_p(\boldsymbol{r}) = E(\boldsymbol{r},\hat{\boldsymbol{s}}_0)$; second as a source of scattered light in the transport equation. The probability that a photon moving in the direction $\hat{\boldsymbol{s}}_0$ is scattered into the direction $\hat{\boldsymbol{s}}$ is equal to $\mu_s p(\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{s}}_0)$. The source term in the transport equation associated with $E(\boldsymbol{r},\hat{\boldsymbol{s}}_0)$ will therefore take the form

$$S(\boldsymbol{r},\hat{\boldsymbol{s}}) = \mu_s E(\boldsymbol{r},\hat{\boldsymbol{s}}_0)p(\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{s}}_0) \tag{6.22}$$

For example, if the source S_A in Fig. 6.4 is a collimated light beam with irradiance E_0, Eq. (6.22) equals $S(\boldsymbol{r},\hat{\boldsymbol{s}}) = \mu_s E_0 p(\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{s}}_A)\exp(-\mu_t\ell_A)$. If S_B is an isotropic point source emitting a power P_0, Eq. (6.22) becomes $S(\boldsymbol{r},\hat{\boldsymbol{s}}) = \mu_s P_0 p(\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{s}}_B)\exp(-\mu_t\ell_B)/(4\pi r^2)$, where r is the distance between S_B and point $\boldsymbol{r}$. If there is more than one source, Eq. (6.22) must be summed over all directions $\hat{\boldsymbol{s}}_0$ or integrated.

When Eq. (6.22) is used as the source term in the transport equation (6.19), primary light has been accounted for so that the radiance in Eq. (6.19) must be represented by $L_s(\boldsymbol{r},\hat{\boldsymbol{s}})$ which describes only scattered light. The transport equation then becomes

$$\hat{\boldsymbol{s}}\cdot\nabla L_s(\boldsymbol{r},\hat{\boldsymbol{s}}) + \mu_t(\boldsymbol{r})L_s(\boldsymbol{r},\hat{\boldsymbol{s}}) = \mu_s\int_{4\pi} p(\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{s}}')L_s(\boldsymbol{r},\hat{\boldsymbol{s}}')\,d\omega' + \mu_s p(\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{s}}_0)E(\boldsymbol{r},\hat{\boldsymbol{s}}_0) \tag{6.23}$$

6.3.4 Diffusion Theory Derived from the Transport Equation

The derivation of the steady state diffusion equation (see Eq. (6.3b)) takes two steps which involve energy balance and flux balance equations based on the transport equation (Eq. (6.23)). The first step is integrating Eq. (6.23) over 4π solid angle. The second step is multiplying Eq. (6.23) by $\hat{\boldsymbol{s}}$ and integrating over 4π solid angle. The two resulting equations are combined to form the diffusion equation, Eq. (6.3), with $\partial\phi/\partial t = 0$.

Since ∇ only operates on $\boldsymbol{r}$, the first term in Eq. (6.23) upon integration can be written

$$\int_{4\pi} \hat{\boldsymbol{s}} \cdot \nabla L_s(\boldsymbol{r}, \hat{\boldsymbol{s}})\, d\omega = \int_{4\pi} \nabla \cdot (L_s(\hat{\boldsymbol{r}}, \hat{\boldsymbol{s}})\hat{\boldsymbol{s}})\, d\omega = \nabla \cdot \Big(\int_{4\pi} L_s(\boldsymbol{r}, \hat{\boldsymbol{s}})\hat{\boldsymbol{s}}\, d\omega \Big) = \nabla \cdot \boldsymbol{F}(\boldsymbol{r}) \tag{6.24}$$

where Eq. (6.7) has been used. The second term in Eq. (6.23) yields $\mu_t \phi_s(\boldsymbol{r})$, using Eq. (6.21). The first term on the right hand side of Eq. (6.23) upon integration yields $\mu_s \phi_s(\boldsymbol{r})$. This can be seen by first integrating over $\hat{\boldsymbol{s}}$ and then over $\hat{\boldsymbol{s}}'$ and noting that $\int_{4\pi} p(\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{s}}')\, d\omega = 1$. The source term of Eq. (6.23) upon integration becomes

$$S_2(\boldsymbol{r}, \hat{\boldsymbol{s}}_0) = \int_{4\pi} \mu_s p(\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{s}}_0) E(\boldsymbol{r}, \hat{\boldsymbol{s}}_0)\, d\omega = \mu_s E(\boldsymbol{r}, \hat{\boldsymbol{s}}_0) \tag{6.25}$$

which is the scalar source term S_2 of Eq. (6.3a). Hence, the result of the integration of Eq. (6.23) over 4π solid angle is

$$\nabla \cdot \boldsymbol{F}(\boldsymbol{r}) = -\mu_a \phi_s(\boldsymbol{r}) + \mu_s E(\boldsymbol{r}, \hat{\boldsymbol{s}}_0) \tag{6.26}$$

This exact equation is identical to Eq. (6.3a), the law of energy conservation for steady state light propagation, $\partial\phi/\partial t = 0$. The physical meaning of $S_2(\boldsymbol{r}, \hat{\boldsymbol{s}}_0)$ in Eq. (6.3a) is now obvious, because the change in the energy density of scattered light at point $\boldsymbol{r}$ should be equal to the local non-scattered irradiance multiplied by the scattering coefficient.

For the second step we multiply Eq. (6.23) by $\hat{\boldsymbol{s}}$ and again integrate over 4π solid angle. The first term will be considered later. The second term on the left side of Eq. (6.23) becomes $\mu_t \boldsymbol{F}(\boldsymbol{r})$, because of the definition of the flux vector, Eq. (6.7). The first term on the right hand side of Eq. (6.23) becomes

$$\mu_s \int_{4\pi} \Big[\int_{4\pi} p(\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{s}}')\hat{\boldsymbol{s}}\, d\omega \Big] L_s(\boldsymbol{r}, \hat{\boldsymbol{s}}')\, d\omega' = \mu_s g \int_{4\pi} L_s(\boldsymbol{r}, \hat{\boldsymbol{s}}')\hat{\boldsymbol{s}}'\, d\omega' = \mu_s g \boldsymbol{F}(\boldsymbol{r}) \tag{6.27}$$

To understand Eq. (6.27) we choose the z-axis in the direction of $\hat{\boldsymbol{s}}'$ and first integrate over ω. The vector $\boldsymbol{v} = \int_{4\pi} p(\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{s}}')\hat{\boldsymbol{s}} d\omega$ has no component perpendicular to $\hat{\boldsymbol{s}}'$, because $p(\hat{\boldsymbol{s}}, \hat{\boldsymbol{s}}')$ was assumed to depend only on the angle between $\hat{\boldsymbol{s}}$ and $\hat{\boldsymbol{s}}'$. The component of $\boldsymbol{v}$ in the direction of $\hat{\boldsymbol{s}}'$ is $\int_{4\pi} p(\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{s}}')(\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{s}}')d\omega = g$, according

to Chapter 3, Eq. (3.47). Therefore, $\boldsymbol{v} = g\hat{\boldsymbol{s}}'$. The last term in Eq. (6.27) then follows from the definition of the flux in Eq. (6.7). The source term in Eq. (6.23) becomes

$$\mu_s E(\boldsymbol{r}, \hat{\boldsymbol{s}}_0) \int_{4\pi} p(\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{s}}_0)\hat{\boldsymbol{s}}\, d\omega = \mu_s g E(\boldsymbol{r}, \hat{\boldsymbol{s}}_0)\hat{\boldsymbol{s}}_0 \tag{6.28}$$

Multiplication of the transport equation by $\hat{\boldsymbol{s}}$ and integration over ω and ω′ thus yields

$$\int_{4\pi} [\hat{\boldsymbol{s}} \cdot \nabla L_s(\boldsymbol{r}, \hat{\boldsymbol{s}})]\hat{\boldsymbol{s}}\, d\omega + \mu_t \boldsymbol{F}(\boldsymbol{r}) = \mu_s g \boldsymbol{F}(\boldsymbol{r}) + \mu_s g E(\boldsymbol{r}, \hat{\boldsymbol{s}}_0)\hat{\boldsymbol{s}}_0 \tag{6.29a}$$

Equation (6.29) is still exact. Again, the source term can be interpreted by noting that Eq. (6.29) is essentially a balance equation for the net flux. The first term at the left is the net flux lost by diffusion (flow). One can see this more easily when this term is written in the form

$$\int_{4\pi} [\hat{\boldsymbol{s}} \cdot \nabla L_s(\boldsymbol{r}, \hat{\boldsymbol{s}})]\hat{\boldsymbol{s}}\, d\omega = \int_{4\pi} [\nabla \cdot (L_s(\boldsymbol{r}, \hat{\boldsymbol{s}})\hat{\boldsymbol{s}}]\hat{\boldsymbol{s}}\, d\omega \tag{6.29b}$$

Here, $L_s(\boldsymbol{r}, \hat{\boldsymbol{s}})\hat{\boldsymbol{s}}$ is an elemental flux vector in the direction $\hat{\boldsymbol{s}}$, so that $[\nabla \cdot (L(\boldsymbol{r}, \hat{\boldsymbol{s}})\hat{\boldsymbol{s}}]\hat{\boldsymbol{s}}$ is the loss of elemental flux in the direction $\hat{\boldsymbol{s}}$. The integral sums these losses and the result is the total loss of net flux by diffusion. The second term is the net flux lost by absorption and scattering. The first term on the right is the net flux gained by scattering of scattered light and the last term is the net flux gained by scattering of non-attenuated light.

Equation (6.29) becomes identical to Eq. (6.2), if we assume that

$$(1/\mu_{tr}) \int_{4\pi} [\hat{\boldsymbol{s}} \cdot \nabla L_s(\boldsymbol{r}, \hat{\boldsymbol{s}})]\hat{\boldsymbol{s}}\, d\omega = D \nabla \phi_s(\boldsymbol{r}) \tag{6.30}$$

Equation (6.30) is called the diffusion approximation (see Ref. [9], pp. 75–76) and D is the diffusion constant, which depends on the optical properties of the medium, but is independent of $\boldsymbol{r}$. Equation (6.29) can now be written

$$\boldsymbol{F}(\boldsymbol{r}) = -D \nabla \phi_s(\boldsymbol{r}) + (\mu_s g/\mu_{tr}) E(\boldsymbol{r}, \hat{\boldsymbol{s}}_0)\hat{\boldsymbol{s}}_0 \tag{6.31}$$

where $\mu_{tr} = \mu_t - \mu_s g = \mu_a + \mu_s(1 - g)$ is known as the transport attenuation coefficient. Equation (6.31) is identical to Eq. (6.2) and thus yields an expression for the vector source term:

$$\boldsymbol{S}_1 = (\mu_s g/\mu_{tr}) E(\boldsymbol{r}, \hat{\boldsymbol{s}}_0)\hat{\boldsymbol{s}}_0 \tag{6.32}$$

We now apply another approximation, assuming that the radiance can be written in the form

$$L_s(\boldsymbol{r},\hat{\boldsymbol{s}}) = \frac{1}{4\pi}\int\limits_{4\pi} L_s(\boldsymbol{r},\hat{\boldsymbol{s}})\,d\omega + \frac{3}{4\pi}\int\limits_{4\pi} L_s(\boldsymbol{r},\hat{\boldsymbol{s}})\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{s}}'\,d\omega = L_0(\boldsymbol{r}) + (3/4\pi)\boldsymbol{F}(\boldsymbol{r})\cdot\hat{\boldsymbol{s}} \quad (6.33)$$

Equation (6.33) is called the Eddington approximation [8] (see also Eqs. (6.10) and (6.18)) and is essentially an expansion of the radiance in terms of Legendre polynomials (see Appendix 2, Eq. (6.131)), cut off after the second term. Substituting Eq. (6.33) into Eq. (6.30) the integral becomes

$$\int\limits_{4\pi} [\hat{\boldsymbol{s}}\cdot\nabla L_0(\boldsymbol{r})]\hat{\boldsymbol{s}}\;\; d\omega + (3/4\pi)\int\limits_{4\pi} [\hat{\boldsymbol{s}}\cdot\nabla(\boldsymbol{F}(\boldsymbol{r})\cdot\hat{\boldsymbol{s}}]\hat{\boldsymbol{s}}\;\; d\omega = (1/3)\nabla\phi(\boldsymbol{r}) \quad (6.34)$$

because of the identity $\int_{4\pi}(\hat{\boldsymbol{s}}\cdot\boldsymbol{A})\hat{\boldsymbol{s}}d\omega = (4\pi/3)\boldsymbol{A}$ which holds for any vector $\boldsymbol{A}$ that is independent of $\hat{\boldsymbol{s}}$, and because of the identity $\int_{4\pi}[\hat{\boldsymbol{s}}\cdot\nabla(\boldsymbol{A}\cdot\hat{\boldsymbol{s}})]\hat{\boldsymbol{s}}d\omega = 0$. In the Eddington approximation the diffusion constant thus equals

$$D = 1/(3\;\mu_{tr}) \quad (6.35)$$

The effective attenuation coefficient (Section 6.2.4) then takes the form $\mu_{\text{eff}} = (3\mu_a\mu_{tr})^{1/2}$. Substitution of Eq. (6.31) into Eq. (6.26) and using Eq. (6.35) produces the equivalent of Eq. (6.3b)

$$\nabla^2\phi_s(\boldsymbol{r}) - 3\mu_a\mu_{tr}\phi_s(\boldsymbol{r}) + 3\mu_s\mu_{tr}E(\boldsymbol{r},\hat{\boldsymbol{s}}_0) - 3\mu_s g\nabla\cdot(E(\boldsymbol{r},\hat{\boldsymbol{s}}_0)\hat{\boldsymbol{s}}_0) = 0 \quad (6.36)$$

Thus, what transport theory adds to simple diffusion theory (Sections 6.2.1 and 6.2.2) is expressions for the source terms and, with the addition of the Eddington approximation, simple expressions for the diffusion coefficient and the effective attenuation coefficient.

6.3.5 *Validity of the Diffusion Approximation*

If the source term in the transport equation vanishes, that is far from boundaries and sources, the variables $\boldsymbol{r}$ and $\hat{\boldsymbol{s}}$ in Eq. (6.23) can be separated and the solution can be written as $L_s(\boldsymbol{r},\hat{\boldsymbol{s}}) = L_1(\boldsymbol{r})L_2(\hat{\boldsymbol{s}})$. For simplicity, without loss of generality, like in Section 6.2.4, we consider a semi infinite medium illuminated by a wide parallel light beam along the z-direction, with $z = 0$ at the surface. Equation (6.23), without the source term, then takes the form

$$\mu\,(\partial L_s(z,\mu)/\partial z) + \mu_t L_s(z,\mu) = \mu_s\int\limits_{4\pi} p(\mu')L_s(z,\mu')\,d\omega' \quad (6.37)$$

where $\mu = \hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{s}}'$ and $\hat{\boldsymbol{s}}$ points in the positive z direction. Into Eq. (6.37) we substitute $L_s(\boldsymbol{r},\hat{\boldsymbol{s}}) = L_s(z,\mu) = L_1(z)L_2(\mu)$ and $L_1(z) = L_1(0)\exp(-\mu_{\text{eff}}z)$ and obtain

$$-\mu\mu_{\text{eff}}L_2(\mu) + \mu_t L_2(\mu) = 2\pi\mu_s \int_{-1}^{+1} p(\mu')L_2(\mu')\, d\,\mu' \tag{6.38}$$

$L_2(\mu)$, the "diffusion pattern," can be solved from Eq. (6.38). Exact closed form solutions exist for isotropic and linear anisotropic scattering (Ref. [1], Section 12 and Appendix III). Using Eq. (3.11) and substituting $L_s(\boldsymbol{r}, \hat{\boldsymbol{s}}) = L_s(z, \mu) = L_1(z)L_2(\mu)$ and $L_1(z) = L_1(0)\exp(-\mu_{\text{eff}}z)$ into Eq. (6.30) produces

$$D = (1/\mu_{tr}) \int_{-1}^{+1} \mu^2 L_2(\mu)d\mu \Big/ \int_{-1}^{+1} L_2(\mu)d\mu \tag{6.39}$$

The value of D may range between $(1/3\mu_{tr})$ and $(1/\mu_{tr})$. Equation (6.39) demonstrates that the diffusion approximation holds far from boundaries and sources. The usefulness as well as the accuracy of the diffusion approximation are then determined by its application also close to boundaries and sources.

Exercise 6.2 Show that deep in a semi-infinite medium with isotropic scattering the transport equation (6.37) without sources has the solution $L(z, \mu) = L(0, 0)\exp(-\mu_{\text{eff}}z)/(1 - \mu\ \mu_{\text{eff}}/\mu_t)$. This exact solution holds for all values of μ_a and μ_s. Show that substitution into Eq. (6.37) produces an implicit formula for μ_{eff}, i.e. $\mu_{\text{eff}}/\mu_s = {}^1\!/_2\ell\text{n}((1 + \mu_{\text{eff}}/\mu_t)/(1 - \mu_{\text{eff}}/\mu_t))$. By expansion of the logarithm show that for $\mu_a/\mu_s << 1$, $\mu_{\text{eff}} \approx (3\mu_a\ \mu_t)^{1/2}$ and that the radiance approaches Eq. (6.18).

An exact solution of the source free transport equation also exists for the linear anisotropic phase function (Eq. (3.51)) and it has the form $L(z, \mu) = L(0, 0)(1 + 3\mu g\ \mu_a/\mu_{\text{eff}})\exp(-\mu_{\text{eff}}z)/(1 - \mu\ \mu_{\text{eff}}/\mu_t)$. Substitution into Eq. (6.37) again produces an implicit formula for μ_{eff} and with a similar expansion as for isotropic scattering and more tedious algebra one obtains that $\mu_{\text{eff}} \approx (3\mu_a\ \mu_{\text{tr}})^{1/2}$.

These exact solutions may give insight in the dependence of the radiance on the optical properties, but they are not useful for practical problems, because they are only valid far from boundaries and sources.

6.3.6 Validity of the Eddington Approximation

To compute a value for D one still has to solve Eq. (6.38). The Eddington approximation helps here, because by substituting Eq. (6.33) in the form of Eq. (6.18), that is $L_2(\mu) = (1 + 3(D\mu_a)^{1/2}\mu)$, into Eq. (6.39) we obtain Eq. (6.35). Note that the second term of $L_2(\mu)$ cancels in the integrals.

In order to obtain a criterion for the validity of the Eddington approximation, we need one more term in the expansion of $L(\mu)$. For convenience we take a term with μ^2 instead of the Legendre polynomial $P_2(\mu)$ and write $L_2(\mu) = (1 + \alpha\mu + \beta\mu^2)$. For the moment we assume that $|\beta| << 1$. From Eq. (6.39) we then obtain $D \approx (1/3\mu_{tr})(1 - 4\beta/15)$. The effective attenuation coefficient becomes

$\mu_{\text{eff}} = (\mu_a/D)^{1/2} \approx (3\mu_a\mu_{tr})(1 + 2\beta/15)$. The fluence rate according to Eq. (3.11), compared with Eq. (6.15) then equals $\phi(z) \approx 2(F_{z+}(z) + F_{z-}(z))(1 - \beta/6)$. Note that the term with α in $L_2(\mu)$ does not occur in the last three expressions. A value for β can be estimated using the exact solution of the source free transport equation for linear anisotropic scattering in Appendix III of Chandrasekhar's book [1]. This gives $\beta = \mu_{\text{eff}}^2/\mu_t^2 + 3\,g\mu_a/\mu_t$ where μ_{eff} is the exact value for linear anisotropic scattering. If we take $\mu_{\text{eff}} \approx (3\mu_a\mu_{tr})^{1/2}$, i.e. the Eddington approximation and assume that $\mu_a << \mu_s$ we find that $\beta \approx 3\mu_a/\mu_s$. Thus we see that the criterion for the validity of the Eddington approximation is that $\mu_a << \mu_s$. This criterion is satisfied in many biomedical optics applications. We see that D, μ_{eff} and $\phi(z)$ can then be calculated rather accurately in the Eddington approximation. Again, this does not necessarily hold near boundaries and sources.

Note that g does not occur in the criterion $\mu_a << \mu_s$. This has to do with the use of the linear anisotropic phase function for the estimates made above. With a more general phase function and a higher approximation [12], one finds a more precise criterion (see Eq. (24) of Ref. [13]) for the validity of the Eddington approximation, i.e. $\mu_a << \mu_s(1 - b_2)$ where b_2 is a coefficient in the Legendre polynomial expansion of the phase function according to Eq. (6.131), Appendix 2. Note that again $b_1 = g$ does not occur in this criterion. In tissue optics experiments, g can be measured, but determining b_2 is practically impossible. If we assume that the Henyey-Greenstein phase function can be a good approximation for the phase function of biological tissues, $b_2 = g^2$ so that the criterion for the validity of the Eddington approximation becomes $\mu_a << \mu_s(1 - g^2) \approx 2\,\mu_s(1 - g)$. Since experimentally $g \approx 0.8 - 0.9$ or even closer to one, the last criterion is much stronger than the previous one, i.e. $\mu_a << \mu_s$.

6.3.7 Phase Functions

Phase functions have been introduced in Chapter 3, Sections 3.4.3.1 and 3.4.3.2. A convenient and often used model phase function is the Henyey-Greenstein function, p_{HG}, Eq. (3.48). Although this function may not exactly represent the true phase function of biological tissue [14], it has been shown experimentally to be a reasonably good approximation [15]. It is also quite practical to work with, particularly in model calculations, because it has only one parameter, g. Furthermore, the detailed behavior of the phase function may not be important for many practical applications such as calculations of the fluence rate. That is, the calculated or measured properties are only sensitive to the first few terms in an expansion of the phase function in terms of Legendre polynomials (Eq. (6.132)), so that it may not be possible to distinguish the "real" phase function from p_{HG} (Eq. (3.48)). See also p. 324 of Ref. [9].

6.3.7.1 The "Delta-Eddington" Approximation

Biological tissues scatter light predominantly in the forward direction. As a result, diffusion theory is not always a good approximation near boundaries or sources

[12, 16]. An improvement is possible by including a delta function in the phase function as follows:

$$p_{\delta-\mathrm{E}}(\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{s}}') = (1-f)p^*(\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{s}}') + f\delta(1-\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{s}}')/(2\pi) \tag{6.40}$$

where $0 < f < 1$. The delta function causes $p_{\delta-\mathrm{E}}(\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{s}}')$ to be strongly forward peaked. Our plan is to select a simple function for $p^*(\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{s}}')$, e.g. isotropic or linear anisotropic, Eq. (3.51), to be compatible with the diffusion + Eddington approximations, and choose f so that $p_{\delta-\mathrm{E}}(\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{s}}')$ approximates p_{HG}. Since p_{HG} is considered a good model phase function for tissue, this will then also be the case for $p^*(\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{s}}')$. If $p^*(\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{s}}')$ is isotropic or linear anisotropic, Eq. (3.61), it is normalized to 1; and since the delta function is also normalized to 1 (Appendix 1), so is $p_{\delta-\mathrm{E}}$.

The procedure sketched in the previous paragraph has been called the "delta-Eddington approximation" by Joseph et al. [17]. Meador and Weaver [18] use the notation δ-E(n) where n is the number of terms in the expansion of $p^*(\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{s}}')$ in Legendre polynomials (see Eq. (6.132), Appendix 2). To see the effect of the delta function, we substitute $p_{\delta-\mathrm{E}}(\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{s}}')$, Eq. (6.40) for $p(\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{s}}')$ in the *source free* transport equation (Eq. (6.19)), where the radiance includes both scattered and primary light. The result is

$$\hat{\boldsymbol{s}}\cdot\nabla L(\boldsymbol{r},\hat{\boldsymbol{s}}) + \mu_t^*(\boldsymbol{r})L(\boldsymbol{r},\hat{\boldsymbol{s}}) = \mu_s^*\int_{4\pi} p^*(\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{s}}')L(\boldsymbol{r},\hat{\boldsymbol{s}}')d\omega' \tag{6.41}$$

This is the original transport equation with p replaced by p^* and with modified parameters $\mu_s^* = \mu_s(1-f)$ and $\mu_t^* = \mu_a + \mu_s^*$. Note that μ_s' as defined in Chapter 1, Table 1.1, the reduced scattering coefficient, is a special case of μ_s^* defined here, for $f = g$ (see also Section 6.3.7.2). Physically, the substitution of Eq. (6.40) is equivalent to shifting part of the strongly forward scattered radiance to the primary, non-scattered radiance. Therefore, Eq. (6.40) should not be substituted into Eq. (6.23), because the latter equation was obtained by separating scattered and primary light. We now apply this separation to Eq. (6.41), using Eq. (6.20):

$$\begin{aligned}\hat{\boldsymbol{s}}\cdot\nabla\left[L_s(\boldsymbol{r},\hat{\boldsymbol{s}})\right] + L_p(\boldsymbol{r},\hat{\boldsymbol{s}})\right] + \mu_t^*(\boldsymbol{r})[L_s(\boldsymbol{r},\hat{\boldsymbol{s}}) + L_p(\boldsymbol{r},\hat{\boldsymbol{s}})] \\ = \int_{4\pi}\mu_s^* p^*(\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{s}}')\left[L_s(\boldsymbol{r},\hat{\boldsymbol{s}}') + L_p(\boldsymbol{r},\hat{\boldsymbol{s}}')\right]d\omega'\end{aligned} \tag{6.42}$$

If $\hat{\boldsymbol{s}}_0$ is the direction of primary photons at $\boldsymbol{r}$ (Fig. 6.4: $\hat{\boldsymbol{s}}_A$, $\hat{\boldsymbol{s}}_B$), the primary radiance is non-zero only for $\hat{\boldsymbol{s}} = \hat{\boldsymbol{s}}_0$ so that the attenuation of primary photons in Eq. (6.42) can be expressed as

$$\hat{\boldsymbol{s}}_0\cdot\nabla L_p(\boldsymbol{r},\hat{\boldsymbol{s}}_0) = \frac{dL_p(\boldsymbol{r},\hat{\boldsymbol{s}}_0)}{ds} = -\mu_t^* L_p(\boldsymbol{r},\hat{\boldsymbol{s}}_0) \tag{6.43}$$

where s is an element of path length of primary photons in the direction $\hat{\boldsymbol{s}}_0$. Note that the attenuation coefficient of primary photons is now μ_t^* (see above) so that L_p

cancels from the left hand side of Eq.(6.42). The term containing L_p on the right hand side can be written as

$$\int_{4\pi} \mu_s^* p^*(\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{s}}') L_p(\boldsymbol{r}, \hat{\boldsymbol{s}}')\, d\omega' = (\mu_s^* E(\boldsymbol{r}, \hat{\boldsymbol{s}}_0)/2\pi) \int_{4\pi} p^*(\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{s}}')\, \delta(1 - \hat{\boldsymbol{s}}' \cdot \hat{\boldsymbol{s}}_0)\, d\omega'$$
$$= \mu_s^* E(\boldsymbol{r}, \hat{\boldsymbol{s}}_0) p^*(\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{s}}_0) \tag{6.44}$$

Here we have substituted Eq. (3.37): $L_p(\boldsymbol{r}, \hat{\boldsymbol{s}}) = E(\boldsymbol{r}, \hat{\boldsymbol{s}}_0)\delta(1 - \hat{\boldsymbol{s}}' \cdot \hat{\boldsymbol{s}}_0)/(2\pi)$ and used Eq. (6.129), Appendix 1. After the delta-Eddington approximation, the transport equation for L_s now reads

$$\hat{\boldsymbol{s}} \cdot \nabla L_s(\boldsymbol{r}, \hat{\boldsymbol{s}}) + \mu_t^*(\boldsymbol{r}) L_s(\boldsymbol{r}, \hat{\boldsymbol{s}}) = \mu_s^* \int_{4\pi} p^*(\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{s}}') L_s(\boldsymbol{r}, \hat{\boldsymbol{s}}')\, d\omega' + \mu_s^* E(\boldsymbol{r}, \hat{\boldsymbol{s}}_0) p^*(\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{s}}_0) \tag{6.45}$$

Equation (6.45) is identical to Eq. (6.23), except for the stars. The source term was obtained in a way different from Section 6.3.3. Note that if the path length of primary photons in the tissue is ℓ, we now have for a collimated light beam $E(\boldsymbol{r}, \hat{\boldsymbol{s}}_0) = E_0(\boldsymbol{r}, \hat{\boldsymbol{s}}_0) \exp(-\mu_t^* \ell)$ (Section 6.3.3) with μ_t^* again replacing μ_t.

6.3.7.2 How Delta-Eddington Approximates the Henyey-Greenstein Function

When the phase function is written as a series in Legendre polynomials according to Eq. (6.132), Appendix 2, Eq. (6.40) yields the following relationship between the coefficients b_n and b_n^* (b_n refers to $p_{\delta-E}(\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{s}}')$ and b_n^* to $p^*(\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{s}}')$):

$$b_n = (1 - f) b_n^* + f \tag{6.46}$$

Here we used the fact that the delta function is represented by setting $b_i = 1$ for all i in Eq. (6.132), which follows from the completeness relation for Legendre polynomials (Appendix 2). The idea of the delta-Eddington approximation is that with a simple phase function p^* a rather good approximation for the Henyey-Greenstein phase function (Eq. (3.48)) can be obtained. For the latter the coefficients b_i in Eq. (6.132) are $b_i = g^i$. In the combination of diffusion + Eddington approximations as discussed in this chapter the natural choices for p^* are isotropic and linear anisotropic (Eq. (3.51)) phase functions. For these two cases the procedure is illustrated in Table 6.1.

Lines 3 and 4 in Table 6.1 represent the right hand side of Eq. (6.40). If we want p^* to be isotropic ($b_0^* = 1, b_i^* = 0, i \geq 1$), since we need $b_1 = g$ for $p_{\delta\text{–E}}$, Eq. (6.46) for $n = 1$ yields $f = g$.

Lines 6 and 7 in Table 6.1 again represent the right hand side of Eq. (6.40). If we now want p^* to be linear anisotropic ($b_0^* = 1, b_1^* = g^*, b_i^* = 0, i \geq 2$) and we need $b_1 = g, b_2 = g^2$ for the approximation by $p_{\delta\text{–E}}$ of the Henyey-Greenstein function, Eq. (6.46) for $n = 2$ gives $f = g^2$ and for $n = 1$ then gives $g^* = g/(1 + g)$.

One can see in Table 6.1 that, judging by the coefficients b_i, the linear combination of p^* and delta function approximates the Henyey-/Greenstein function better as the number of terms in p^* increases (compare lines 5 and 8 with line 2) [12, 16].

Table 6.1 Approximating the Henyey-Greenstein phase function by the delta-Eddington approximation, Eq. (6.40)

Coefficients b_i in Eq. (6.132)	b_0	b_1	b_2	b_3	...	$\sum_{i=0}^{\infty}[(2i+1)/2]b_iP_i(\mu)$
Henyey-Greenstein function	1	g	g^2	g^3	...	$\sum_{i=0}^{\infty}[(2i+1)/2]g^iP_i(\mu)$
p^* isotropic ($g^*=0$)	1	0	0	0	...	multiply with $1-f=1-g$, result $=A$
delta function	1	1	1	1	...	multiply with $f=g$, result $=B$
result for $p_{\delta\text{-E}}$, Eq. (6.40)	1	g	g	g	...	sum $=A+B$
p^* linear anisotropic, Eq. (3.51) ($g^*=g/(1+g)$)	1	g^*	0	0	...	multiply with $1-f=1-g^2$, result $=C$
delta function	1	1	1	1	...	multiply with $f=g^2$, result $=D$
result for $p_{\delta\text{-E}}$, Eq. (6.40)	1	g	g^2	g^2	...	sum $=C+D$

Each line represents a function with its coefficients in the Legendre polynomial expansion (Appendix 2)

In Section 6.3.7.1 we saw that $\mu_s^* = \mu_s(1-f)$. With the help of Eq. (6.46) this can be rewritten in the form $\mu_s^*(1-b_n^*) = \mu_s(1-b_n)$. For the second example of Table 6.1 this becomes $\mu_s^*(1-g^*) = \mu_s(1-g)$. Thus, the delta-Eddington approximation leaves the reduced scattering coefficient unchanged and consequently also the effective attenuation coefficient. Furthermore, the condition for the validity of the Eddington approximation in Section 6.3.6, i.e. $\mu_a << \mu_s(1-b_2) = \mu_s(1-g^2)$ is also unchanged.

Exercise 6.3 If the Henyey-Greenstein phase function (Eq. (3.48)) is written as a series in Legendre polynomials (Eq. (6.131)), show that $b_i = g^i$.
Hint: According to Eq. (6.133) $b_n = \int_{-1}^{+1} p_{\text{HG}}(\mu)P_n(\mu)\text{d}\mu$. Write down the equations for b_{n+1} and b_{n-1}, perform partial integrations so that $\text{d}P_{n+1}(\mu)/\text{d}\mu$ and $\text{d}P_{n-1}(\mu)/\text{d}\mu$ appear in the integrals, respectively, and subtract. Apply Eq. (6.135) and subsequently Eq. (6.136) to obtain the recurrence formula $(1+g^2)b_n = g(b_{n-1} + b_{n+1})$. Since by definition $b_0 = 1$ and $b_1 = g$, it follows by induction that $b_n = g^n$.

6.3.8 Boundary Conditions Between Two Media, One Without and One with Scattering

For refractive index matched boundaries, the boundary conditions can be written in a general form (see below, Eq. (6.47). Index mismatched boundary conditions are more complicated because of reflection and refraction at the boundary. Although the present discussion will be restricted to the slab geometry, this is not a severe limitation. Results for the slab can be used in other geometries with symmetry such as spherical and cylindrical geometries. In each case, energy conservation will be used as the criterion to determine boundary conditions.

6.3.8.1 Boundary Conditions for the Diffusion Equation

Refractive Index Matched Boundaries

When no external source of diffuse (scattered) light is present, then $F_{n+}(\boldsymbol{r}) = 0$ at each point $\boldsymbol{r}$ of the boundary (Section 6.2.5, Fig. 6.3). Using Eq. (6.13a) this can be written as $\phi_s(\boldsymbol{r}) = -2\boldsymbol{F}(\boldsymbol{r}) \cdot \hat{\boldsymbol{n}}$, where $\hat{\boldsymbol{n}}$ is a unit vector pointing into the medium at point $\boldsymbol{r}$ of the boundary. Combining this with Eq. (6.2) yields

$$\phi_s(\boldsymbol{r}) - 2D\nabla\phi_s(\boldsymbol{r}) \cdot \hat{\boldsymbol{n}} + 2\boldsymbol{S}_1(\boldsymbol{r}) \cdot \hat{\boldsymbol{n}} = 0 \tag{6.47}$$

Equation (6.47) is equivalent to Eq. (9-24) of Ishimaru's book [19]. The diffusion equation (6.36) can be solved with this boundary condition. According to Eq. (6.32), $\boldsymbol{S}_1 = (\mu_s g/\mu_{tr})E(\boldsymbol{r})\hat{\boldsymbol{s}}_0$, and according to Eq. (6.35), $D = 1/(3\mu_{tr})$.

For an infinite half space with isotropic scattering, no absorption, no incident flux and sources only at infinity (the Milne problem, which can be solved exactly), the factor $2D = 2/(3\mu_{tr})$ in Eq. (6.47) changes to $0.71/\mu_s$ (Eq. (9-25) of Ref. [19]). This does not mean that energy is not conserved, because Eq. (6.47) accounts for energy conservation only *within* the diffusion approximation and the factor $(0.71/\mu_s)$ arises from an exact solution.

Refractive Index Mismatched Boundaries

In Section 6.2.5 it has been assumed that the hemispherical fluxes F_{n+} and F_{n-} at the boundary can be related by a reflection factor. Energy conservation then leads to the boundary condition $F_{n+}(r) = r_{id}F_{n-}(\boldsymbol{r})$, where $\boldsymbol{r}$ is on the boundary and r_{id} is defined in Chapter 3, Section 3.3.3. Substituting this boundary condition into Eqs. (6.13a) and (6.13b) and subtracting these equations one obtains

$$[(1 - r_{id})/(1 + r_{id})]\,\phi_s(\boldsymbol{r}) - 2D\nabla\phi_s(\boldsymbol{r}) \cdot \hat{\boldsymbol{n}} + 2\boldsymbol{S}_1(\boldsymbol{r}) \cdot \hat{\boldsymbol{n}} = 0 \tag{6.48}$$

If the z axis is parallel to $\hat{\boldsymbol{n}}$, the origin is at $\boldsymbol{r}$ and there is no source at $\boldsymbol{r}$, Eq. (6.48) can be written as $\phi_s(0) = h[\partial\phi_s(\boldsymbol{r})/\partial z]_{z=0}$ where $h = 2D(1 + r_{id})/(1 - r_{id})$; h is called the extrapolation length, because upon linear extrapolation along z, $\phi_s(-h) = \phi_s(0) - h[\partial\phi_s(\boldsymbol{r})/\partial z]_{z=0} = 0$.

The existence of a reflection factor is not obvious. One requirement is that light scattering occurs at only one side of the boundary. If $\boldsymbol{F}(\boldsymbol{r})$ is parallel to $\hat{\boldsymbol{n}}$, a reflection factor can be defined in many cases. When $\boldsymbol{F}(\boldsymbol{r})$ is not parallel to $\hat{\boldsymbol{n}}$, this reflection factor applies to the components of $\boldsymbol{F}(\boldsymbol{r})$ along $\hat{\boldsymbol{n}}$; that is, $F_{n+}(\boldsymbol{r})$ and $F_{n-}(\boldsymbol{r})$ (Eqs. (3.12) and (6.13)). This will be further discussed in Section 6.3.8.2.

6.3.8.2 Boundary Conditions as Reflection Factors for Hemispherical Fluxes

Assuming Isotropic Radiance and a Step Function for Fresnel Reflection

A simple formula for r_{id} is obtained as follows. Approximate the Fresnel reflection function for unpolarized light (Eq. (3.15a)) by a step function. That is, it equals

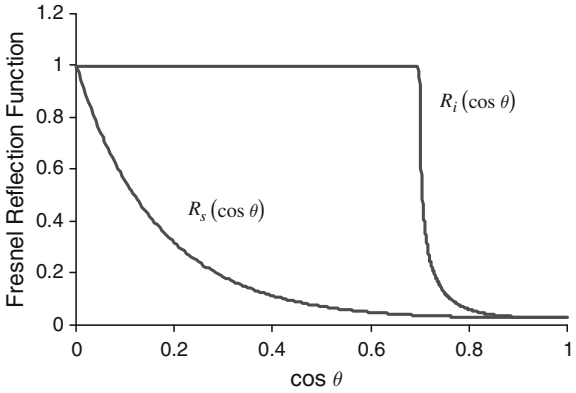


Fig. 6.5 Fresnel reflection as a function of the cosine of the angle of incidence. $R_s(\cos\theta)$ is the specular reflection coefficient according to Eq. (3.15a) for light in air incident on the boundary with a medium of $n = 1.4$ (for tissue, according to Ref. [20]) so that $\cos\theta_c = 0.71$. $R_i(\cos\theta)$ ("i" for "internal") is the same reflection function for light in the medium incident on the boundary with air

one if $\theta \geq \theta_c$ and zero if $\theta < \theta_c$ where θ_c is the critical angle (the actual form of the Fresnel reflection function is shown in Fig. 6.5). We further assume that the radiance corresponding to the scattered light incident on the boundary is isotropic. With these assumptions, $R_s(\theta_2)$ in Chapter 3, Eq. (3.29) is zero for $0 < \theta < \theta_c$ so that Eq. (3.29) becomes

$$r_{21} = \frac{F_{n+}(r)}{F_{n-}(r)} = [\cos(2\theta_c) + 1]/2 = 1 - (\sin\theta_c)^2 = 1 - n^{-2} \tag{6.49}$$

Note that we use r_{21} (clear medium 1, scattering medium 2, $n_2 > n_1$) instead of r_{id} here because the radiance in the medium is not always completely diffuse. Given the limited validity of diffusion theory, Eq. (6.49) is a rather good approximation. Physically better justified approximations will be considered below and their results are compared in Table 6.2.

Table 6.2 Internal diffuse flux reflection factor r_{21} for unpolarized light calculated using different approximations[a]

Reflection function	"Step approximation"	Fresnel "exact", Eq. (3.15a)
Isotropic radiance in medium	0.49 (Eq. (6.49))	0.53
Radiance according to Eq. (6.50)	0.45 (Eq. (6.53))	0.49

[a]See text. The scattering medium 2 (tissue) has refractive index $n_2 = 1.4$ [20] and the non-scattering medium 1 has refractive index $n_1 = 1$

Using the Radiance from Diffusion Theory and a Step Function for Fresnel Reflection

The assumption of isotropic radiance (Chapter 3, Section 3.3.3), used for the derivation of Eq. (6.49), may be approximately correct far from boundaries and sources, but tends to become incorrect at the boundary. For a slab of tissue, infinitely wide with thickness d ("one-dimensional"), a better approximation of L_s is given by Eq. (6.10) which, combined with Eqs. (6.14) and (6.15), can be written as

$$L_s(z,\mu) = [F_{n+}(z) + F_{n-}(z) + (3\mu/2)(F_{n+}(z) - F_{n-}(z))]/(2\pi) \tag{6.50}$$

where $\mu = \cos\theta$ and θ is the angle between $\hat{\boldsymbol{s}}$ and the inward normal $\hat{\boldsymbol{n}}$ which is along the z axis. For an index mismatched boundary at $z = 0$, the exact boundary condition is $L_s(0, \cos\theta) = R_s(\theta)L_s(0, \cos(\pi - \theta))$ for $0 < \theta < \pi/2$, where $R_s(\theta)$ is the Fresnel reflection factor for unpolarized light. In words: the forward radiance is equal to the reflected backward radiance. However, this condition cannot be satisfied exactly for any θ in the diffusion model, because of the limited number of adjustable parameters. Suitable boundary conditions are obtained by requiring only that the hemispherical fluxes associated with these radiances are equal. Analogous to Eqs. (6.11a, b) we thus write

$$\int_0^1 L_s(0,\mu)\mu\, d\mu = \int_{-1}^0 -R_i(\mu)L_s(0,\mu)\mu\, d\mu \quad \text{at } z = 0 \tag{6.51a}$$

and similarly

$$\int_0^1 R_i(\mu)L_s(d,\mu)\mu\, d\mu = \int_{-1}^0 -L_s(d,\mu)\mu\, d\mu \quad \text{at } z = d \tag{6.51b}$$

where d is the slab thickness and $R_i(\mu)$ is again Fresnel's specular reflection coefficient (R_s has been replaced by R_i here, to be consistent with Fig. 6.5). Approximating $R_i(\mu)$ again by a step function, that is $R_i(\mu) = 1$ for $0 < \mu < \cos\theta_c$ and $R_s(\mu) = 0$ for $\cos\theta_c < \mu < 1$, Eq. (6.51a) becomes

$$\int_0^1 L_s(0,\mu)\mu\, d\mu = \int_{-\cos\theta_c}^0 -L_s(0,\mu)\mu\, d\mu \quad \text{at } z = 0 \tag{6.52}$$

An analogous equation is obtained from Eq. (6.51b) for $z = d$. By substituting the radiance from Eq. (6.50) into Eq. (6.52) and obtaining the ratio $F_{n+}(0)/F_{n-}(0)$, the reflectance r_{21} is

$$r_{21} = F_{n+}(0)/F_{n-}(0) = [(\cos\theta_c)^2 + (\cos\theta_c)^3]/[2 - (\cos\theta_c)^2 + (\cos\theta_c)^3)] \tag{6.53}$$

where r_{21} is the reflection factor for the scattered light flux in medium 2 striking the boundary with non-scattering medium 1. Similarly, at $z = d$ we find that $r_{21} = F_{n-}(d)/F_{n+}(d)$ gives the same result as Eq. (6.53). For practical values of the refractive indices, the difference between Eq. (6.49) and Eq. (6.53) is small (Table 6.2). Note that the radiance according to Eq. (6.50) may become negative at the boundary when μ approaches the value 1 ($\cos\theta = 1$). The radiance is positive for all θ only when $r_{21} > 1/5$. This indicates that the radiance, which is a microscopic quantity, cannot always be physically meaningful in the diffusion approximation, which is essentially a macroscopic theory. Far from boundaries and sources the radiance is always positive.

Exercise 6.4 Show that for $z \to 0$ Eq. (6.50) can be written in the form $L_s(z, \mu) = \phi_s(z)[1 - (3\mu/2)(1 - r_{21})/(1 + r_{21})]/(4\pi)$. Observe that the anisotropic part becomes smaller as r_{21} approaches 1. If $r_{21} = 0$ the radiance is negative for $\mu > 2/3$ or $\theta > 48$ degrees. In tissue optics, the largest value is $r_{21} = 0.49$ (Table 6.1) and the radiance remains positive for all μ. Even though the radiance may not be important in diffusion theory, the latter fact may be favorable for the accuracy of the diffusion + Eddington approximation.

Using the Exact Fresnel Reflection Function for Unpolarized Light

Further improvement of the boundary conditions may be possible by using the exact Fresnel function (Eq. (3.15a) in Eq. (6.51a) and Eq. (6.51b)) and evaluating the integrals numerically. If one is interested in closed form solutions, a simple formula that approximates the Fresnel function rather accurately can also be used [12, 16]. This permits evaluation of r_{21} on a programmable pocket calculator. The results are all the same of course. Table 6.2 shows r_{21} for two approximations of the radiance. The differences in Table 6.2 are rather small relative to our knowledge of the exact values for tissue optical properties. Note that $r_{21} = 0.49$ for $n = 1.4$ (lower right corner) should be the most accurate value. In fact, r_{21} is dominated by the region where the angle of incidence is larger than the critical angle, so that approximations of the radiance and the Fresnel function have a relatively small effect.

So far we have considered reflection of scattered ("diffuse") light in medium 2 (tissue) at the boundary with medium 1 ($n_2 > n_1 = 1$), which is the most common situation. Similar calculations can be performed for "diffuse" light in medium 1 reflected at the boundary with medium 2 ($n_2 > n_1$). We have encountered this problem when calculating the response of isotropic light dosimetry probes as a function of the tissue refractive index [21]. Here the reflection function R_s (Fig. 6.5) must be used. Since there is not a critical angle for total reflection, the Fresnel function cannot be approximated by a step function. The results are therefore more sensitive to the actual value of the Fresnel reflection function. Nevertheless, calculations using an approximation of R_s still produce quite good results [12, 16].

Boundary Conditions When the Net Photon Flux is Not Perpendicularly Incident on the Boundary

This subject has been briefly discussed in Section 6.2.5. In Section 6.3.8.2 we have up till now considered only the slab geometry, where the net flux is perpendicular

to the boundary. When the net flux has an arbitrary direction, it is not immediately obvious that the boundary conditions can still be described by a reflection factor. For a geometry as in Fig. 6.3, energy conservation requires that the net inward-directed flux is equal to the reflected part of the outward-directed flux, when scattered light does not enter the medium. In terms of the radiance this condition can be written as

$$\int\limits_{\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{n}}>0} L(\boldsymbol{r},\hat{\boldsymbol{s}})(\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{n}})\,d\omega = -\int\limits_{\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{n}}<0} L(\boldsymbol{r},\hat{\boldsymbol{s}})R(\left|\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{n}}\right|)(\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{n}})\,d\omega \tag{6.54}$$

where R is the Fresnel specular reflection function. The radiance is now written according to Eq. (6.10), that is $L(\boldsymbol{r},\hat{\boldsymbol{s}}) = L_0(\boldsymbol{r}) + (3/4\pi)\boldsymbol{F}(\boldsymbol{r})\cdot\hat{\boldsymbol{s}} = L_0(\boldsymbol{r}) + 3(\boldsymbol{F}_n(\boldsymbol{r}) + \boldsymbol{F}_\perp(\boldsymbol{r}))\cdot\hat{\boldsymbol{s}}/4\pi$, where $\boldsymbol{F}_n(\boldsymbol{r})$ and $\boldsymbol{F}_\perp(\boldsymbol{r})$ are the components of $\boldsymbol{F}(\boldsymbol{r})$ parallel and perpendicular to $\hat{\boldsymbol{n}}$. For the integrations in Eq. (6.54) the z-axis is chosen parallel to $\hat{\boldsymbol{n}}$. The terms with $\boldsymbol{F}_\perp(\boldsymbol{r}))\cdot\hat{\boldsymbol{s}}$ include a factor $\cos\varphi$ or $\sin\varphi$ where φ is the azimuthal angle, and these terms become zero because $R(\left|\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{n}}\right|)$ does not depend on φ. The portion of $L(\boldsymbol{r},\hat{\boldsymbol{s}})$ that is left in Eq. (6.54) comprises the terms $L_0(\boldsymbol{r}) + (3/4\pi)\boldsymbol{F}_n(\boldsymbol{r})\cdot\hat{\boldsymbol{s}}$ which are identical to the right hand side of Eq. (6.50), because of Eq. (6.15). Therefore, when the net flux is obliquely incident on the boundary, the boundary condition can be written as a reflection factor for the component of the flux parallel to the normal vector $\hat{\boldsymbol{n}}$. That is, $F_{n+}(\boldsymbol{r}) = r_{21}F_{n-}(\boldsymbol{r})$.

6.3.8.3 A Useful Transformation When Boundary Conditions are Expressed in Terms of Reflection Factors

For boundaries with mismatched refractive indices, obtaining a closed-form solution of Eqs.(6.26) and (6.31) is much more tedious than for matched refractive indices. When the equations are written in terms of hemispherical fluxes, Eqs. (6.11a, b) and Eqs. (6.70a, b), a transformation is possible yielding new fluxes to which index matched boundary conditions apply.

The flowing fluxes are defined for the slab geometry:

$$F'_{n+}(z) = F_{n+}(z) - r_{21}F_{n-}(z) \tag{6.55a}$$

$$F'_{n-}(z) = F_{n-}(z) - r_{21}F_{n+}(z) \tag{6.55b}$$

and conversely

$$F_{n+}(z) = [F'_{n+}(z) + r_{21}F'_{n-}(z)]/(1 - r_{21}^2) \tag{6.56a}$$

$$F_{n-}(z) = [F'_{n-}(z) + r_{21}F'_{n+}(z)]/(1 - r_{21}^2) \tag{6.56b}$$

Substitution in Eqs. (6.26) and (6.31) or Eqs. (6.70a, b) yields equations for $F'_{n+}(z)$ and $F'_{n-}(z)$ that have the same form, but with coefficients multiplied or divided by factors $(1 - r_{21})$ and/or $(1 + r_{21})$. This is so because

$$F'_{n+}(z) - F'_{n-}(z) = (1 + r_{21})[F_{n+}(z) - F_{n-}(z)] \tag{6.57a}$$

$$F'_{n+}(z) + F'_{n-}(z) = (1 - r_{21})[F_{n+}(z) + F_{n-}(z)] \tag{6.57b}$$

The new equations can be solved as the original equations, with the advantage that the boundary conditions are $F'_{n+}(0) = 0$ and $F'_{n-}(d) = 0$, that is, the boundary conditions for index matching. The scattered part of the fluence rate becomes

$$\phi_s = \phi'_s/(1 - r_{21}), \quad \text{where } \phi'_s = 2[F'_{n+}(z) + F'_{n-}(z)] \tag{6.58}$$

Diffusion theory can be expressed as a multiple flux model by substituting Eqs. (6.11a,b) into Eqs. (6.26) and (6.31) for each component $[x, y, z]$ of the net flux $\boldsymbol{F}(\boldsymbol{r})$ and using Eq. (6.15). The transformation can therefore also be applied in three dimensions, not only in orthogonal coordinates, but also in spherical and cylindrical coordinates (Sections 6.4.2 and 6.4.3.2).

6.3.9 *Boundary Conditions Between Two Media, Both with Scattering*

First, the boundary conditions will be presented in a simple, "non-exact" way, using reflection factors. Then, the boundary conditions will be derived from the transport equation, using the criterion of energy conservation and the results of the two treatments will be compared. Let the forward and backward fluxes at the boundary be F_{1+} and F_{1-} in medium 1 and F_{2+} and F_{2-} in medium 2, respectively (Fig. 6.6). Intuitively, it seems reasonable to assume that the fluxes in medium 1 are reflected with factor r_{12} and in medium 2 with r_{21} so that, on the basis of energy conservation, the boundary conditions might be written

$$F_{2+} = (1 - r_{21})F_{1+} + r_{21}F_{2-} \tag{6.59a}$$

$$F_{1-} = r_{12}F_{1+} + (1 - r_{21})F_{2-} \tag{6.59b}$$

It turns out that Eqs. (6.59a, b) are only approximately correct. The difference with the proper treatment is relatively small however. The exact boundary condition requires that the radiance is equal to the sum of the transmitted and reflected

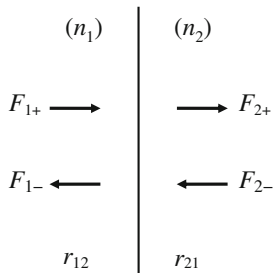


Fig. 6.6 Illustration of an intuitive but non-exact treatment of boundary conditions between two scattering media, using fluxes with reflection factors, leading to Eqs. (6.59a, b)

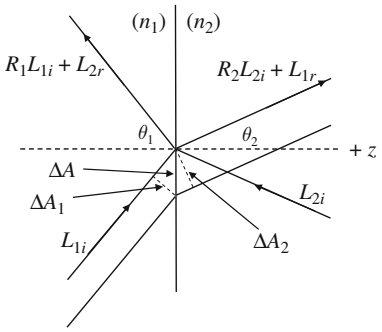


Fig. 6.7 The radiance in two media at a boundary, resulting from both reflection and refraction (see text). Note that $R_1 = R_1(\theta_1)$ and $R_2 = R_2(\theta_2)$. R_1 is the Fresnel reflection function (Eq. (3.15a)) for light in medium 1, incident on the boundary with medium 2. R_2 is the same function for light in medium 2, incident on the boundary with medium 1

radiances. This leads to two equations, one for each side of the boundary. Let $L_1(\theta)$ and $L_2(\theta)$ be the radiances in medium 1 and 2, respectively, where θ is the angle between the directional unit vector $\hat{\boldsymbol{s}}$ and the positive z-axis, which in the direction from medium 1 to medium 2. For clarity of the illustration (Fig. 6.7) we write L_{1i} for the radiance in medium 1, incident on the boundary, R_1L_{1i} for the reflected part in medium 1 and L_{1r} for the *refracted* part of that same radiance in medium 2. Quantities L_{2i} and L_{2r} are defined analogously. The exact boundary conditions for the total radiances L_1 in medium 1 and L_2 in medium 2 are, respectively

$$L_1(\pi - \theta_1) = R_1(\theta_1)L_{1i}(\theta_1) + L_{2r}(\pi - \theta_1) \tag{6.60a}$$

$$L_2(\theta_2) = R_2(\theta_2)L_{2i}(\pi - \theta_2) + L_{1r}(\theta_2) \tag{6.60b}$$

These equations must hold for all corresponding angles θ_1 and θ_2 which are related by Snell's law: $n_1 \sin\theta_1 = n_2 \sin\theta_2$. This local condition cannot be satisfied for any combination of θ_1 and θ_2 in the diffusion model. Analogous to Eqs. (6.51a, b), we therefore require that the net fluxes associated with the radiances in Eqs. (6.60a, b) satisfy these equations. Energy conservation during refraction yields relationships between L_{1i}, L_{1r} and L_{2i}, L_{2r} respectively, using the definition of radiance (Section 3.2.2.2, see also Fig. 6.7):

$$\Delta A_1 L_{1i}(\theta_1)(1 - R_1(\theta_1))[2\pi \sin\theta_1 d\theta_1] = \Delta A_2 L_{1r}(\theta_2)[2\pi \sin\theta_2 d\theta_2] \tag{6.61a}$$

$$\Delta A_2 L_{2i}(\pi - \theta_2)(1 - R_2(\theta_2))[2\pi \sin\theta_2 d\theta_2] = \Delta A_1 L_{2r}(\pi - \theta_1)[2\pi \sin\theta_1\, d\theta_1] \tag{6.61b}$$

The left hand side of Eqs.(6.61a, b) is the power of the photons incident on an area ΔA of the boundary within a given solid angle (between brackets) that will be refracted. The right hand side is the power of the refracted photons on the other side of the boundary within a solid angle that is related to the solid angle in the medium of incidence by Snell's law. Note that $\Delta A_1 = \Delta A \cos(\theta_1)$ and $\Delta A_2 = \Delta A \cos(\theta_2)$.

Obtain L_{1r} from Eq. (6.60b) and substitute into Eq. (6.61a), obtain L_{2r} from Eq. (6.60a), substitute into Eq. (6.61b) and drop the subscript i from L_{1i} and L_{2i}. This yields the following boundary conditions

$$\int_0^{\pi/2} L_2(\theta_2)\cos\theta_2 \sin\theta_2\, d\theta_2 = \int_0^{\pi/2} R_2(\theta_2)L_2(\pi-\theta_2)\cos\theta_2\sin\theta_2\, d\theta_2 + \\ + \int_0^{\pi/2} (1-R_1(\theta_1))L_1(\theta_1)\cos\theta_1\sin\theta_1\, d\theta_1 \tag{6.62a}$$

and

$$\int_0^{\pi/2} L_1(\pi-\theta_1)\cos\theta_1 \sin\theta_1\, d\theta_1 = \int_0^{\pi/2} R_1(\theta_1)L_1(\theta_1)\cos\theta_1\sin\theta_1\, d\theta_1 + \\ + \int_0^{\pi/2} (1-R_2(\theta_2))L_2(\pi-\theta_2)\cos\theta_2\sin\theta_2\, d\theta_2 \tag{6.62b}$$

When $\mu = \cos\theta$ is substituted into Eqs. (6.62a, b) the resulting equations are similar to Eqs. (6.51a, b), with terms added for refracted light. Substituting Eq. (6.50) for the radiance, the boundary conditions (Eqs. (6.62a, b)) can be written

$$F_{2+} = [(1-A_{12})F_{1+} + A_{21}F_{2-} - B_{12}F_{1-}]/(1-B_{21}) \tag{6.63a}$$

$$F_{1-} = [A_{12}F_{1+} + (1-A_{21})F_{2-} - B_{21}F_{2+}]/(1-B_{12}) \tag{6.63b}$$

Here,

$$A_{12} = \int_0^1 (\mu + (3/2)\,\mu^2)R_1(\mu)\, d\mu \tag{6.64a}$$

$$A_{21} = \int_0^1 (\mu + (3/2)\,\mu^2)R_2(\mu)\, d\mu \tag{6.64b}$$

$$B_{12} = \int_0^1 (\mu - (3/2)\,\mu^2)R_1(\mu)\, d\mu \tag{6.65a}$$

$$B_{21} = \int_0^1 (\mu - (3/2)\,\mu^2)R_2(\mu)\, d\mu \tag{6.65b}$$

R_1 and R_2 have been defined in Fig. 6.7. If $n_2 > n_1$, total reflection occurs in medium 2, but not in medium 1, and then $R_1 \equiv R_s$ and $R_2 \equiv R_i$ (see Fig. 6.5). As an example, take $n_1 = 1$ and $n_2 = 1.4$ (tissue). Then, $A_{12} = 0.062$, $B_{12} = 0.014$, $A_{21} = 0.46$ and $B_{21} = 0.070$. Note that according to previous definitions,

$r_{12} = A_{12}/(1 - B_{12})$ and $r_{21} = A_{21}/(1 - B_{21})$. Using the approximation $B_{12} = 0$, Eqs. (6.63a, b) become

$$F_{2+} = (1 - r_{12})F_{1+}/(1 - B_{21}) + r_{21}F_{2-} \tag{6.66a}$$

$$F_{1-} = r_{12}F_{1+} + [1 - (1 - B_{21})r_{21}]F_{2-} - B_{21}F_{2+} \tag{6.66b}$$

One can see that the difference between Eqs. (6.66a, b) and Eqs. (6.63a, b) or Eqs. (6.59a, b) is rather small, depending of course on the difference between n_1 and n_2. When higher order approximations of the transport equation are used, such as the P_3 approximation [12, 16], a more precise treatment of boundary conditions is required and the use of reflection factors for net fluxes is no longer possible.

6.4 Diffusion Theory in Simple Geometries

6.4.1 Plane Geometry

In this geometry, both the net flux vector $\boldsymbol{F}(\boldsymbol{r})$ and the normal vector $\hat{\boldsymbol{n}}$ in Eq. (6.14) are parallel to the positive z axis. For convenience the subscript n will be dropped so that $F_{n+}(z) = F_+(z)$ and $F_{n-}(z) = F_-(z)$.

6.4.1.1 Measurable Quantities: Fluence Rate, Absorbed Energy, Reflection and Transmission

The most common situation is surface irradiation of a slab with thickness d by a perpendicularly incident collimated light beam with uniform irradiance E_0. The total radiant energy fluence rate ϕ_t is the sum of the scattered and nonscattered fluence rates, that is

$$\phi_t(z) = \phi_s(z) + \phi_p(z) = 2[F_+(z) + F_-(z)] + (1 - r_{ce})E_0 \exp(-\mu_t z) \tag{6.67}$$

Here, $F_+(z)$ and $F_-(z)$ are the hemispherical fluxes. The scattered part of the fluence rate $\phi_s(z)$ is given by Eq. (6.15) and r_{ce} is the specular reflection coefficient for collimated light. If a delta-Eddington transformation is used to represent the phase function (Section 6.3.7.1), then $\phi_s(z)$ and $\phi_p(z)$ are mathematical quantities that cannot be measured separately. The total fluence rate can be measured in situ using an isotropic probe [21].

The absorbed optical power density is equal to $\mu_a\phi_t(z)$. When the optical power is converted to heat, a temperature rise is produced that can be measured and thermal tissue damage may occur, the level of which depends on the temperature-time history. If μ_a is the absorption coefficient of a photosensitizer, light absorption leads to photochemical reactions and to subsequent photodynamic damage. The photochemical reaction may dominate over thermal effects when the light fluences are

low. Both thermal and photodynamic damage may lead to a change in the optical properties of the tissue.

The reflected flux, which is identical to the net flux in the negative z direction, is given by

$$R = (1 - r_{21})F_-(0) + r_{ce}E_0 \tag{6.68}$$

If a delta-Eddington transformation has been applied, the scattered and collimated parts of the reflection are again mathematical quantities that cannot be measured separately. Only their sum can be measured. The same holds for the transmission. Note that F_- is proportional to E_0. If the boundary conditions cannot be written in terms of diffuse reflection factors, the reflected radiance (see Section 6.3.9) must be used to calculate the reflected flux.

The transmitted flux is given by

$$T = (1 - r_{21})F_+(d) + (1 - r_{ce})^2 E_0 \exp(-\mu_t d) \tag{6.69}$$

When tissue sections are held between glass slides, additional reflecting surfaces must be considered. Expressions for $F_+(z)$ and $F_-(z)$ follow from the solution given in the next section.

Instead of a collimated light beam, it is also possible to have diffuse incident light. The expressions for fluence rate, reflection and transmission are obtained by deleting the collimated terms from Eqs. (6.67)–(6.69). The solution of the slab problem in this case is given in Section 6.4.1.4 below.

6.4.1.2 Infinitely Wide Slab of Finite Thickness with Wide-Beam Collimated Irradiance and Refractive Index Matched Boundaries

Assume a slab of thickness d with $z = 0$ at the index matched environment-medium boundary. In this geometry the collimated irradiance of the incident beam has the form $E(\boldsymbol{r}) = (1 - r_{ce})E_0 \exp(-\mu_t z)$. Using Eq. (6.15), Eqs. (6.26) and (6.31) can be written, respectively, as

$$\partial[F_+(z) - F_-(z)]\partial z = -2\mu_a[F_+(z) + F_-(z)] + \mu_s(1 - r_{ce})E_0 \exp(-\mu_t z) \tag{6.70a}$$

$$\partial[F_+(z) + F_-(z)]\partial z = -(3/2)\,\mu_{tr}[F_+(z) - F_-(z)] + (3/2)g\mu_s(1 - r_{ce})E_0 \exp(-\mu_t z) \tag{6.70b}$$

Solving for $F_+(z)$ and $F_-(z)$ produces

$$\partial^2 F_\pm(z)/\partial z^2 - 3\mu_a\mu_{tr}F_\pm(z) + S_\pm(1 - r_{ce})E_0 \exp(-\mu_t z) = 0 \tag{6.71a, b}$$

where S_+ and S_- are given by

$$S_+ = (\mu_s/4)[(5 + 9\,g)\,\mu_a + 5\mu_s] \tag{6.72a}$$

$$S_- = (\mu_s/4)[(1-3\,g)\,\mu_a + \mu_s] \tag{6.72b}$$

Adding Eqs. (6.71a) and (6.71b) and using Eq. (6.15) yields the equivalent of Eq. (6.36) for the slab geometry with wide-beam collimated irradiance. Mathematically it makes no difference which equations are solved. Since the boundary conditions for the hemispherical fluxes $F_\pm$ are physically more transparent than those for the fluence rate $\phi_s(\boldsymbol{r})$, Eqs. (6.47) and (6.48), we use the equations for $F_\pm$.

Since all fluxes are proportional to $(1-r_{ce})E_0$ we set $(1-r_{ce})E_0 = 1$. Equations (6.71a, b) have two types of solutions. The first is the particular solution, of the form $A_\pm \exp(-\mu_t z)$. The coefficients A_+ and A_- are obtained by substitution into Eqs. (6.71a, b) and the result is

$$A_\pm = -S_\pm/(\mu_t^2 - \mu_{\text{eff}}^2) \tag{6.73}$$

where $\mu_{\text{eff}}^2 = 3\mu_a\mu_{tr} = 3\mu_a(\mu_a + (1-g)\mu_s)$. The second type of solution is any solution of the source free equations (= zero irradiance), also called the homogeneous equations. If $A_\pm \exp(-\mu_t z)$ satisfies Eqs. (6.71a, b) then $A_\pm \exp(-\mu_t z)$ plus any linear combination of the functions $\exp(\pm\mu_{\text{eff}} z)$ also satisfies these equations. Therefore, the general solution is

$$F_+(z) = A_{++}\exp(\mu_{\text{eff}} z) + A_{+-}\exp(-\mu_{\text{eff}} z) + A_+\exp(-\mu_t z) \tag{6.74a}$$

$$F_-(z) = A_{-+}\exp(\mu_{\text{eff}} z) + A_{--}\exp(-\mu_{\text{eff}} z) + A_-\exp(-\mu_t z) \tag{6.74b}$$

The four unknown coefficients are determined by the boundary conditions, which provide two relationships, one at $z = 0$ and one at $z = d$, plus the two equations (6.70a, b). Straightforward algebra gives the following results. Substitute Eqs. (6.74a, b) into Eqs. (6.70a, b). Since the coefficients of each exponential must be zero we obtain

$$A_{++} = qA_{-+} \tag{6.75}$$

$$A_{--} = qA_{+-} \tag{6.76}$$

where

$$q = (\mu_{\text{eff}} - 2\mu_a)/(\mu_{\text{eff}} + 2\mu_a) \tag{6.77}$$

The boundary conditions for refractive index matching are $F_+(0) = 0$ and $F_-(d) = 0$ where d is the slab thickness. Substitution into Eqs. (6.74a, b) yields

$$A_{-+} = \frac{-A_+\exp(-\mu_{\text{eff}} d) + (A_-/q)]\exp(-\mu_t d)}{q\exp(-\mu_{\text{eff}} d) - (1/q)\exp(\mu_{\text{eff}} d)} \tag{6.78}$$

$$A_{+-} = \frac{(A_+/q)\exp(\mu_{\text{eff}} d) - A_-\exp(-\mu_t d)}{q\exp(-\mu_{\text{eff}} d) - (1/q)\exp(\mu_{\text{eff}} d)} \tag{6.79}$$

For large values of d, $A_{-+} \to 0$ and thus $A_{++} \to 0$. The other coefficients become $A_{+-} = -A_+$ and $A_{--} = -qA_+$. The fluence rate in a slab of large thickness, including nonscattered light, can then be written as

$$\begin{aligned}\phi_t(z) &= \phi_p(z) + \phi_s(z) = (1 - r_{ce})E_0 \exp(-\mu_t z) + 2[F_+(z) + F_-(z)] \\ &= \exp(-\mu_t z) + 2[S_+(1+q)\exp(-\mu_{\text{eff}} z) - (S_+ + S_-)\exp(-\mu_t z)]/(\mu_t^2 - \mu_{eff}^2)\end{aligned} \tag{6.80}$$

In the limit of $\mu_a << \mu_s(1-g)$ this reduces to

$$\phi_t(z) = [5/(1 + 2\mu_a/\mu_{\text{eff}})]\exp(-\mu_{\text{eff}} z) - 2\exp(-\mu_t z) \tag{6.81}$$

In this approximation, terms of order (μ_a/μ_s) and higher have been neglected. Equation (6.81) shows that in the diffusion approximation, for an index matched boundary, the largest possible value of the fluence rate in the medium at $z = 0$ is three times the incident irradiance. For an index mismatched boundary this value can be even larger (see below, Section 6.4.1.3).

The explicit solution of the slab problem with index mismatched boundaries becomes rather tedious. It may then be convenient to use the transformation described in Section 6.3.8.3.

6.4.1.3 Semi-Infinite Medium with Wide Beam Collimated Irradiance and Refractive Index Mismatched Boundaries

A semi-infinite medium is a "slab" with infinite thickness d. Many laser irradiated tissues may be considered semi-infinite. The expressions for the fluxes are sufficiently simple to give the complete diffusion theory solution for a refractive index mismatched boundary.

In the previous section we have seen that if $d \to \infty$, $A_{-+} = A_{++} = 0$ and $A_{--} = qA_{+-} = -qA_+$. Inserting these values in Eqs. (6.74a, b) and applying the boundary condition $F_+(0) = r_{21}F_-(0)$ yields $A_{+-} = (S_+ - r_{21}S_-)/[(1 - r_{21}q)(\mu_t^2 - \mu_{\text{eff}}^2)]$. Equations (6.74a, b) then reduce to

$$F_+(z) = \frac{S_+ - r_{21}S_-}{(\mu_t^2 - \mu_{\text{eff}}^2)(1 - r_{21}q)}\exp(-\mu_{\text{eff}} z) - \frac{S_+}{(\mu_t^2 - \mu_{\text{eff}}^2)}\exp(-\mu_t z) \tag{6.82a}$$

$$F_-(z) = \frac{q(S_+ - r_{21}S_-)}{(\mu_t^2 - \mu_{\text{eff}}^2)(1 - r_{21}q)}\exp(-\mu_{\text{eff}} z) - \frac{S_-}{(\mu_t^2 - \mu_{\text{eff}}^2)}\exp(-\mu_t z) \tag{6.82b}$$

where $S_\pm$ and q have been defined in Section 6.4.1.2. The total fluence rate is given by Eq. (6.67) and with $(1 - r_{ce})E_0 = 1$, for $\mu_a << \mu_s$ reduces to

$$\phi_t(z) = \frac{5 - r_{21}}{1 - r_{21}}\frac{\exp(-\mu_{\text{eff}} z)}{1 + h\mu_{\text{eff}}} - 2\exp(-\mu_t z) \tag{6.83}$$

where $h = 2D(1 + r_{21})/(1 - r_{21})$ is the extrapolation length (Section 6.3.8.1) and $D = (1/3\mu_{tr})$.

It is instructive to consider "conservative scattering," that is $\mu_a = 0$, $\mu_{\text{eff}} = 0$ and $\mu_a/\mu_{\text{eff}} = 0$. In this limit $\phi_t(0)$ from Eq. (6.83) can also be derived from physical arguments, rather than from a complete solution of the equations of diffusion theory. Assume a collimated light beam with irradiance $(1 - r_{ce})E_0 = 1$ perpendicularly incident on a semi infinite, scattering but non-absorbing medium. There is no net flux $\boldsymbol{F}(z)$ deep inside the medium, since $\mu_{\text{eff}} = 0$. This is confirmed by Eqs. (6.82a, b) where $F_+(z) = F_-(z)$ for large z and $\mu_a = 0$ $(q = 1)$. Because absorption is zero, there is no energy loss in the semi infinite volume. Therefore, to conserve energy, the flux reflected from the medium must equal the incident irradiance. That is, $F(z) = -\hat{\boldsymbol{s}}_z$ ($\hat{\boldsymbol{s}}_z$ is a unit vector in the positive z direction), or $F_-(0) = 1$ for refractive index matching. The total fluence rate at the surface is the sum of the collimated irradiance (=1) and the scattered fluence rate, that is $2[F_+(0)+F_-(0)] = 2$, so that $\phi_t(0) = 3$, in agreement with Eq. (6.83). For a refractive index mismatched boundary the reflected flux must still equal 1, but this now means that $(1 - r_{21})F_-(0) = 1$ and $F_+(0) = r_{21}F_-(0)$. For the total fluence rate at the surface in the medium, this yields $\phi_t(0) = 1 + 2[F_+(0) + F_-(0)] = (3 + r_{21})/(1 - r_{21})$, again in agreement with Eq. (6.83).

For conservative scattering $\phi_t(z)$ can also be calculated by integrating Eq. (6.26) with $\mu_a = 0$. This yields $\boldsymbol{F}(z) = -\exp(-\mu_s z)\hat{\boldsymbol{s}}_z$. $\boldsymbol{F}(z)$ is then substituted into Eq. (6.31), which is subsequently integrated to produce $\phi_t(z) = \exp(-\mu_s z) + \phi_s(0) + 3[1-\exp(-\mu_s z)]$. This equation is independent of the boundary conditions. The first exponential term is the contribution of the incident beam and $\phi_s(0)$ is the scattered fluence rate at $z = 0$. Substituting the proper value of $\phi_s(0)$ for refractive index matching or mismatching (see above), respectively, again produces equations that are identical to Eq. (6.83) in the appropriate approximation. For index matching the fluence rate has a surface value $\phi_s(0) = 3$ and increases with depth to a plateau value of 5. The effect of refractive index mismatch is a constant shift of the fluence rate by $4r_{21}/(1 - r_{21})$ for all z. See the illustration in Fig. 6.8. When μ_a increases

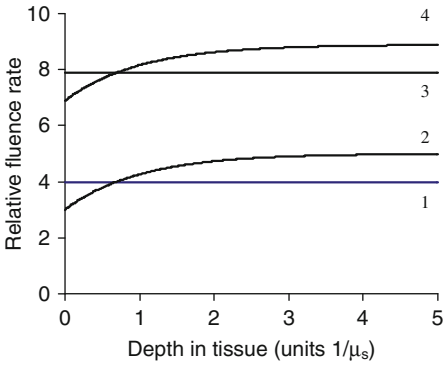


Fig. 6.8 Fluence rate as a function of depth in tissue for conservative scattering in a semi-infinite medium, relative to the irradiance of the incident beam. 1 and 2: Refractive index matching; 1: diffuse irradiance; 2: collimated irradiance; 3 and 4; refractive index mismatch, $n = 1.4$, $r_{21} = 0.49$; 3: diffuse irradiance; 4: collimated irradiance

to a non-zero value, the plateau of the refractive index mismatched case turns into a subsurface maximum in the fluence rate.

Exercise 6.5 It is very easy to see that Eqs. (6.82a, b) satisfy the boundary condition $F_+(0) = r_{21}F_-(0)$. Now show that the solution given by Eq. (6.83) satisfies the equivalent boundary condition given by Eq. (6.48). This may seem trivial, but there is a snag.

Exercise 6.6 Consider Eq. (6.83) without and with the delta-Eddington approximations presented in Table 6.1. The fluence rate $\phi_t(0)$ is the same in all cases. Why is that so? Note that if Eqs. (6.82a,b) are used to calculate $\phi_t(0)$ without the approximation $\mu_a << \mu_s$, there is a small difference between the three cases.

Use the optical properties $\mu_a = 0.27\ \text{mm}^{-1}$, $\mu_s = 18.73\ \text{mm}^{-1}$, $g = 0.81$ [17] and $r_{21} = 0.49$ to calculate the maximum fluence rate and the depth z_m at which it occurs, for a refractive index matched as well as mismatched boundary at $z = 0$. Do this for the "regular" diffusion approximation as well as for the δ-E approximations of Table 6.1 ($g^* = 0$ and $g^* = g/(1 + g)$ respectively). Note that the subsurface maximum becomes less pronounced going from "regular diffusion" to $g^* = g/(1 + g)$ to $g^* = 0$ and that it is also less pronounced for the index mismatched compared to the index matched boundary. For an index mismatched boundary the maximum is absent when $g^* = 0$. It may also be instructive to calculate and plot the diffuse fluxes separately, together with the total fluence rate.

Comparing Eq. (6.83) with Monte Carlo simulations using the Henyey-Greenstein phase function and with the P_3 – approximation [12, 16], it appears that δ-E with $g^* = g/(1 + g)$ is probably the best choice in the diffusion approximation.

6.4.1.4 Infinitely Wide Slab of Finite Thickness with Wide-Beam Diffuse Irradiance and Refractive Index Matched Boundaries

In medical applications, diffuse incident light is less common than collimated light. Ambient light is partly diffuse, which plays a role in photosynthesis following illumination of plant leaves. Diffuse incident light occurs in the measurement of optical properties using the double integrating sphere method [22, 23]. For illumination of skin, for example in photodynamic therapy, diffuse light may be created by using a small integrating sphere, fed by an optical fiber connected to a laser. In such a setup, no light is lost by reflection from the skin surface. It is important to note that for the same irradiance, the energy fluence rate in tissue associated with either collimated or diffuse light may be different, depending on the tissue optical properties.

With diffuse incident light, the source terms are absent from Eqs. (6.70a, b) so that the equations for F_+ and F_- become

$$\partial[F_+(z) - F_-(z)]\partial z = -2\mu_a[F_+(z) + F_-(z)] \tag{6.84a}$$

$$\partial[F_+(z) + F_-(z)]\partial z = -(3/2)\,\mu_{tr}[F_+(z) - F_-(z)] \tag{6.84b}$$

The solution of Eqs. (6.84a, b) has the same form as Eqs. (6.74a, b) but without the terms proportional to $\exp(-\mu_t z)$. When the refractive indices on both sides of the boundaries match, the boundary conditions are very simple, i.e., $F_+(0) = 1$ and $F_-(d) = 0$, where the incident flux has been normalized to one. The solution for the fluxes can then be written in a relatively simple form, using the functions $\sinh(z)$ and $\cosh(z)$:

$$F_+(z) = \frac{2\cosh[\mu_{\text{eff}}(d-z)] + [(\mu_{\text{eff}}/2\mu_a) + (2\mu_a/\mu_{\text{eff}})]\sinh[\mu_{\text{eff}}(d-z)]}{2\cosh(\mu_{\text{eff}}d) + [(\mu_{\text{eff}}/2\mu_a) + (2\mu_a/\mu_{\text{eff}})]\sinh(\mu_{\text{eff}}d)]} \tag{6.85a}$$

$$F_-(z) = \frac{[(\mu_{\text{eff}}/2\mu_a) - (2\mu_a/\mu_{\text{eff}})]\sinh[\mu_{\text{eff}}(d-z)]}{2\cosh(\mu_{\text{eff}}d) + [(\mu_{\text{eff}}/2\mu_a) + (2\mu_a/\mu_{\text{eff}})]\sinh(\mu_{\text{eff}}d)]} \tag{6.85b}$$

The total fluence rate is given by Eq. (6.67) without the collimated term.

6.4.1.5 Kubelka-Munk

The Kubelka-Munk model (KM) has been used mainly in the paint industry, but also in the early days of tissue optics [24]. Relatively recently, elaborate extensions of KM have been published [25], so that this model deserves a few words. In our notation, the KM equations can be written:

$$\partial[F_+(z) - F_-(z)]\partial z = -K[F_+(z) + F_-(z)] \tag{6.86a}$$

$$\partial[F_+(z) + F_-(z)]\partial z = -(2S + K)[F_+(z) - F_-(z)] \tag{6.86b}$$

Equations (6.87a, b) are quite similar to Eqs. (6.84a, b). Kubelka [24] argues that for completely diffuse light $K = 2\mu_a$ and $S = 2\mu_s$ so that Eqs. (6.84a) and (6.86a) are identical. Furthermore, KM considers only backward scattering, so that $g = -1$ and therefore $\mu_{tr} = \mu_a + \mu_s(1-g) = \mu_a + 2\mu_s = (K+2S)/2$. Equations (6.84b) and (6.86b) are then identical except for a factor 3/4. The latter is explained by the fact that KM lets the scattering process act upon the flux, essentially a one dimensional process, rather than on the three dimensional radiance as in transport theory. KM does not consider the radiance.

Since ambient light is often diffuse, the choice of diffuse incident light by KM is reasonable when studying layers of paint. KM only consider refractive index matched boundaries. Finally, only reflection and transmission are considered, not fluence rate which is important in tissue optics.

6.4.1.6 Semi-Infinite Medium with Wide-Beam Diffuse Irradiance and Refractive Index Mismatched Boundaries

For a semi-infinite medium the expressions for the fluxes in a refractive index mismatched situation are sufficiently simple to present the complete solution. If the

incident diffuse flux is F_i a fraction r_{sd} is reflected, where r_{sd} is given by Eq. (3.28) [the subscript *sd* means "specular diffuse"]. The incident flux at the boundary in the medium equals $(1 - r_{sd})F_i$ which is normalized to 1 for this discussion. The boundary condition at $z = 0$ is $F_+(0) = 1 + r_{21}F_-(0)$. The solution for the fluxes then becomes

$$F_+(z) = [1/(1 - r_{21}q)] \exp(-\mu_{\text{eff}}z) \tag{6.87a}$$

$$F_-(z) = [q/(1 - r_{21}q)] \exp(-\mu_{\text{eff}}z) \tag{6.87b}$$

where r_{21} has been defined in Section 6.3.8.2 and q is given by Eq. (6.77). The total fluence rate is given by Eq. (6.67) without the collimated term and can be written as

$$\phi(z) = 2[(1 + q)/(1 - r_{21}q)] \exp(-\mu_{\text{eff}}z) = \frac{4}{1 - r_{21}} \frac{\exp(-\mu_{\text{eff}}z)}{1 + h\mu_{\text{eff}}} \tag{6.88}$$

where $h = 2D(1 + r_{21})/(1 - r_{21})$ and $D = (1/3\mu_{tr})$.

Again it is instructive to consider conservative scattering ($\mu_a = 0$). For a semi-infinite medium the fluence rate can again be derived from physical arguments and compared to Eq. (6.88). Within the medium there is no net flux, $\boldsymbol{F}(z) = 0$. Since there is no collimated term and $\mu_{\text{eff}} = 0$, the fluence rate in the medium is independent of depth. The incident diffuse flux (which was normalized to 1) must be totally reflected, that is $(1 - r_{21})F_-(0) = 1$. The refractive index mismatched boundary condition is $F_+(0) = 1 + r_{21}F_-(0)$. The fluence rate then equals $\phi(z) = 2[F_0(0) + F_-(0)] = 4/(1 - r_{21})$, in agreement with Eq. (6.88). See also Fig. 6.8. Note that when μ_a increases to a non-zero value, the maximum fluence rate occurs at $z = 0$, rather than below the surface for collimated irradiance.

For conservative scattering, the fluxes in a slab of finite thickness d can be derived easily from Eqs. (6.26) and (6.31). Since $\mu_a = 0$, the fluxes must be linear functions of z. Using the boundary conditions $F_+(0) = 1 + r_{21}F_-(0)$ and $F_-(d) = r_{21}F_+(d)$ one obtains:

$$F_+(z) = [4D/(1 - r_{21}) + (d - z)]/[4d(1 + r_{21}) + d(1 - r_{21})] \tag{6.89a}$$

$$F_-(z) = [4r_{21}D/(1 - r_{21}) + (d - z)]/[4D(1 + r_{21}) + d(1 - r_{21})] \tag{6.89b}$$

Here, $D = (1/\mu_{tr})$ is the diffusion constant. Note that the sum of reflection and transmission equals one in this case, that is $(1 - r_{21})F_-(0) + (1 - r_{21})F_+(d) = 1$. The fluence rate follows from Eqs. (6.89a, b) in the usual way.

6.4.2 Spherical Geometry with Isotropic Point Source

6.4.2.1 Equations and Solution

Equations

Assume a "spherical slab" with radii r_1 and r_2 (Fig. 6.9). Equations (6.26) and (6.31) are expressed in spherical polar coordinates r, θ and φ, using the well known expressions for $\nabla \cdot \boldsymbol{F}(\boldsymbol{r})$ and $\nabla\phi$ in these coordinates. Because of the symmetry, only the radial dependence remains as shown in the following equations:

$$\begin{aligned}(1/r^2)\partial\{r^2[F_{r+}(r) - F_{r-}(r)]\}/\partial r = &-2\mu_a[F_{r+}(r) + F_{r-}(r)] \\ &+\mu_s(P/4\pi r^2)\exp[-\mu_t(r - r_1)]\end{aligned} \tag{6.90a}$$

$$\begin{aligned}\partial[F_{r+}(r) + F_{r-}(r)]/\partial r = &-(3/2)\mu_{tr}[F_{r+}(r) - F_{r-}(r)] \\ &+(3/2)g\mu_s(P/4\pi r^2)\exp[-\mu_t(r - r_1)]\end{aligned} \tag{6.90b}$$

which are the equivalent forms of Eqs. (6.70a) and (6.70b). Here we have used the relationships $\phi_s(r) = 2(F_{r+}(r) + F_{r-}(r))$ for scattered light, Eq. (6.15), $F_r = F_{r+} - F_{r-}$, Eq. (6.14) and have assumed an isotropic point source at the origin with power P so that the irradiance is $E(r) = P/(4\pi r^2)$.

Solution

Multiply Eqs. (6.90a, b) by r^2. Solve $\partial\{r^2[F_{r+}(r) - F_{r-}(r)]\}/\partial r$ from Eq. (6.90a) and substitute into the derivative of Eq. (6.90b). Using the identity $(\partial/\partial r)\{r^2\partial f(r)/\partial r]\} = r\partial^2[rf(r)]\partial r^2$ one obtains

$$\begin{aligned}\partial^2\{r[F_{r+}(r) + F_{r-}(r)]\}/\partial r^2 - 3\mu_a\mu_{tr}r[F_{r+}(r) + F_{r-}(r)] = \\ = -[P(S_+ + S_-)/(4\pi r)]\exp[-\mu_t(r - r_1)]\end{aligned} \tag{6.91}$$

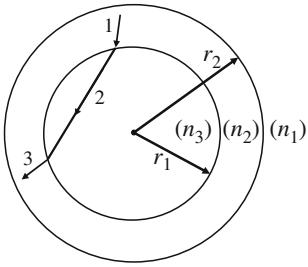


Fig. 6.9 "Spherical slab", that is, an isotropic point source at the center of two spheres with radii r_1 and r_2 with a scattering and absorbing medium for $r_1 < r < r_2$ (refractive index n_2) and non-scattering and non-absorbing media elsewhere (refractive indices $n_1 < n_2$ and $n_3 < n_2$). The light ray (1, 2, 3) illustrates the boundary condition $F_{r+}(r_1) = F_{r-}(r_1)$, resulting from the complete symmetry in the trajectories of photons entering and leaving medium 2 at $r = r_1$

where S_+ and S_- are given by Eqs. (6.72a, b). From Eq. (6.91), $r[F_{r+}(r) + F_{r-}(r)]$ can be solved and substituted into Eq. (6.90b) to obtain $F_{r+}(r) - F_{r-}(r)$. The solution of Eq. (6.91) is the sum of the solution of the full equation (i.e. a particular solution) plus the solutions of the source free equation multiplied by constants $B_\pm$ that are determined by the boundary conditions. The solutions of Eq. (6.91) without sources are

$$F_{r+}(r) + F_{r-}(r) = B_\pm (1/r) \exp(\pm \mu_{\text{eff}} r) \tag{6.92}$$

Equation (6.92) is substituted into Eq. (6.90b) without the source term, which produces

$$F_{r+}(r) - F_{r-}(r) = B_\pm [2(1 \mp \mu_{\text{eff}} r)/(3\mu_{tr} r^2)] \exp(\pm \mu_{\text{eff}} r) \tag{6.93}$$

The inhomogeneous equation, that is Eq. (6.91) with sources, can be written in the form

$$\partial^2 H_B(r)/\partial r^2 - \mu_{\text{eff}}^2 H_B(r) = -[P(S_+ + S_-)/(4\pi r)] \exp[-\mu_t (r - r_1)] = -h_B(r) \tag{6.94}$$

where $H_B(r) = r[F_{r+}(r) + F_{r-}(r)]$ and $h_B(r)$ is a shorthand notation for the right hand side of the equation. The particular solution of Eq. (6.94) is

$$H_B(r) = \frac{1}{2\mu_{eff}} \int_r^{r_0} [\exp(\mu_{\text{eff}} r) \exp(-\mu_{\text{eff}} x) - \exp(-\mu_{\text{eff}} r) \exp(\mu_{\text{eff}} x)] h_B(x) dx \tag{6.95}$$

The integration constant r_0 can be chosen arbitrarily, because if the upper bound is replaced by $r_0{}'$ this amounts to adding an integral similar to Eq. (6.95), between r_0 and $r_0{}'$. Since the variable r does not occur in the bounds of the added integral, this is simply a linear combination of $\exp(\mu_{\text{eff}} r)$ and $\exp(-\mu_{\text{eff}} r)$, that is, a solution of the homogeneous equation. For computational reasons we choose $r_0 = \infty$. The particular solution can then be written

$$\begin{aligned} H_B(r) = P \frac{3\mu_s \exp(\mu_t r_1)}{16\pi \mu_{\text{eff}}} [(1+g)\mu_a + \mu_s] [\exp(\mu_{\text{eff}} r) B_0((\mu_t + \mu_{\text{eff}}) r, \infty) \\ - \exp(-\mu_{\text{eff}} r) B_0((\mu_t - \mu_{\text{eff}}) r, \infty)] \end{aligned} \tag{6.96}$$

where

$$B_0(a, b) = \int_a^b \frac{\exp(-u)}{u} du \tag{6.97}$$

For this integral very good polynomial approximations exist [26], which facilitate numerical calculations. The complete solution for $F_{r+}(r) + F_{r-}(r)$ can now be written

$$F_{r+}(r) + F_{r-}(r) = (B_+/r)\exp(\mu_{\text{eff}} r) + (B_-/r)\exp(-\mu_{\text{eff}} r) + H_B(r)/r \quad (6.98)$$

The solution for $F_{r+}(r) - F_{r-}(r)$ then follows from Eq. (6.90b). The total fluence rate is

$$\begin{aligned}\phi_t(r) &= 2[F_{r+}(r) + F_{r-}(r)] + P\exp[-\mu_t(r - r_1)]/(4\pi r^2) = \\ &= 2(B_+/r)\exp(\mu_{\text{eff}} r) + 2(B_-/r)\exp(-\mu_{\text{eff}} r) + \\ &+ P\frac{3\mu_s \exp(\mu_t r_1)}{8\pi\,\mu_{\text{eff}} r}[(1+g)\mu_a + \mu_s][\exp(\mu_{\text{eff}} r)B_0((\mu_t + \mu_{\text{eff}})r, \infty) \\ &- \exp(-\mu_{\text{eff}} r)B_0((\mu_t - \mu_{\text{eff}})r, \infty)] + \\ &+ P\exp[-\mu_t(r - r_1)]/(4\pi r^2)\end{aligned} \quad (6.99)$$

for $r_1 < r < r_2$. The parameters B_+ and B_- are integration constants, determined by the boundary conditions (Section 6.4.2.2).

Equation (6.99) looks rather complicated, but the geometry and optical properties in practical applications allow considerable simplifications, as will be shown in Section 6.4.2.3.

6.4.2.2 Boundary Conditions for the Spherical Geometry

The discussion of boundary conditions expressed in terms of fluxes in Sections 6.3.8 and 6.3.9 applies also to the spherical geometry. At $r = r_1$ the boundary condition is exactly satisfied. That is, the radiance at the boundary in the medium is identical for the outward and inward directions, even for refractive index mismatch between medium and cavity, because all light scattered into the cavity re-enters the medium elsewhere at a geometrically identical point (Fig. 6.9). If $n_2 > n_3$ and the reflection factor for scattered light is r_{23}, then at each point of the surface $F_{r+}(r_1) = r_{23}F_{r-}(r_1) + (1 - r_{23})F_{r-}(r_1) = F_{r-}(r_1)$. At $r = r_2$ the boundary condition is $F_{r-}(r_2) = r_{21}F_{r+}(r_2)$, similar to the flat slab problem (Section 6.3.8.2). This can also be written as $F_{r+}(r_2) + F_{r-}(r_2) = A(F_{r+}(r_1) - F_{r-}(r_1))$ where $A = (1 + r_{21})/(1 - r_{21})$ The two boundary conditions yield two equations from which the constants B_+ and B_- can be calculated.

If scattered ("diffuse") light occurs at both sides of a boundary, the treatment in Section 6.3.9 must be applied. An example is the calculation of the response of isotropic light dosimetry probes as a function of the tissue refractive index [21]. This type of boundary condition will not be further discussed in this chapter.

6.4.2.3 Approximate Solutions for $r_2 \to \infty$

$r_2 \to \infty$ and $r_1 >> 1/\mu_t$

A practical application is photodynamic therapy for superficial bladder cancer by whole bladder wall illumination [27, 29], where an isotropic point source is placed at the center of the bladder cavity.

If $r_2 \rightarrow \infty$, B_+ must be zero. If $r_1 \neq 0$, tissue optical properties and common geometries make sure that $\mu_t r \geq \mu_t r_1 >> 1$. In that case, a in $B_0(a, \infty)$ is large and the integral (Eq. (6.97)) approaches $\exp(-a)/a$ [26]. Equation (6.98) then reduces to

$$F_{r+}(r) + F_{r-}(r) = (B_-/r)\exp(-\mu_{\text{eff}} r) - 3P\exp[-\mu_t(r - r_1)]/(8\pi r^2) \quad (6.100)$$

where we have also used $\mu_a << \mu_s$ (Section 6.3.6). Note that the last term of Eq. (6.100), representing the particular solution (Section 6.4.2.1), has the same form as the last term in Eq. (6.99), representing the source.

Using Eqs. (6.100) and (6.90b), the boundary condition $F_{r+}(r_1) = F_{r-}(r_1)$ yields $B_- = (3P\mu_{tr})\exp(\mu_{\text{eff}} r_1)/[8\pi(1 + \mu_{\text{eff}} r_1)]$, if $\mu_a << \mu_s$. The total fluence rate in this approximation then becomes

$$\phi_t(r) = (3P/4\pi)\{\mu_{tr}\exp[-\mu_{\text{eff}}(r-r_1)]/[r(1+\mu_{\text{eff}} r_1)] - (2/3)\exp[-\mu_t(r-r_1)]/r^2\} \quad (6.101)$$

Note the similarity between Eq. (6.101) and Eq. (6.83), in particular the last term.

Exercise 6.7 Use Eq. (6.101) to calculate the fluence rate $\phi_t(r_1)$ at the surface $r = r_1 = 33$ mm of a spherical cavity (volume 150 ml), in a medium with optical properties $\mu_a = 0.11\ \text{mm}^{-1}$, $\mu_s = 16.9\ \text{mm}^{-1}$, $g = 0.89$ [29] and $r_2 = \infty$. Show that if $\mu_{\text{eff}} r_1 >> 1$ the ratio between $\phi_t(r_1)$ and the irradiance from the source, i.e. $P/(4\pi {r_1}^2)$, can be written as $\mu_{\text{eff}}/\mu_a - 2 = 5.3$, independent of r_1 and of the refractive indices.

$r_2 \rightarrow \infty, r_1 \rightarrow 0$ and $r >> 1/\mu_t$

This is the classical situation of an isotropic point source in an infinite medium, which is of practical importance in laser thermotherapy. Again, B_+ must be zero. To determine B_- we first consider $r \rightarrow 0$ $(r > r_1)$. Then, in Eq. (6.97) $a \rightarrow 0$ and $b \rightarrow \infty$ (Section 6.4.2.1) so that $B_0(a, \infty) \approx -\ell n(a)$ [26]. The last factor in Eq. (6.96) then becomes

$$-\exp(\mu_{\text{eff}} r)\ell n[(\mu_t+\mu_{\text{eff}})r] + \exp(-\mu_{\text{eff}} r)\ell_n[(\mu_t-\mu_{\text{eff}})r] \approx \ell n\frac{\mu_t - \mu_{\text{eff}}}{\mu_t + \mu_{\text{eff}}} \approx \frac{-2\mu_{\text{eff}}}{\mu_t} \quad (6.102)$$

where terms of order $(\mu_{\text{eff}}/\mu_t)^3$ and higher have been neglected, since $\mu_a << \mu_s$ (Section 6.3.6). Equation (6.102), substituted into Eq. (6.96) yields a simple form for the particular solution ($r_1 = 0$), which leads to the following result for Eq. (6.98)

$$F_{r+}(r) + F_{r-}(r) = (B_-/r)\exp(-\mu_{\text{eff}} r) - 3P\mu_s/(8\pi r) \quad (6.103)$$

The exponential in Eq. (6.103) could have been omitted (since $r \rightarrow 0$) but was kept for the sake of clarity. Note that Eq. (6.103) differs from Eq. (6.100)

because the second term on the right no longer has the functional form of the non-scattered irradiance, but has the form of the solution of Eq. (6.91) without the source term. Substituting Eq. (6.103) into Eq. (6.90b) and applying the boundary condition $F_{r+}(r_1) = F_{r-}(r_1)$ gives $B_- = 3P\mu_{tr}/(8\pi)$ (for $r_1 \to 0$), the same form as in the beginning of this Section. Having determined B_- we now restrict ourselves to $r >> 1/\mu_t$, so that Eq. (6.100) holds again with $r_1 = 0$, and the total scattered plus non-scattered fluence rate is

$$\phi_t(r) = [(3P\mu_{tr}/(4\pi r)]\exp(-\mu_{\text{eff}} r) - [P/(2\pi r^2)]\exp(-\mu_t r) \tag{6.104}$$

Equation (6.104) is identical to Eq. (6.101) with $r_1 = 0$. Since $r >> 1/\mu_t$, which holds for most tissues and wavelengths within one mm from the source, we can neglect the last term.

Conservative scattering, $0 < r_1 \le r \le r_2 < \infty$

As in Sections 6.4.1.3 and 6.4.1.6, it is instructive to consider conservative scattering ($\mu_a = 0$). Energy conservation requires that the total net outward flux of scattered light, plus the flux of non-attenuated light from the source, is equal to $P/(4\pi r^2)$ for all r, or $P/(4\pi r^2) = F_{r+}(r) - F_{r-}(r) + P\exp[-\mu_t(r - r_1)]/(4\pi r^2)$. When this equation for $F_{r+}(r) - F_{r-}(r)$, which satisfies the boundary condition at $r = r_1$, is substituted into Eq. (6.90b) and the resulting equation for $F_{r+}(r) + F_{r-}(r)$ is integrated, one obtains

$$\phi_s(r) = \frac{3P\mu_s(1-g)}{4\pi r} - \frac{3P\mu_s^2}{4\pi}\exp(\mu_s r_1)\int_{\mu_s r}^{\infty}\frac{\exp(-u)}{u^2}du + C \tag{6.105}$$

The integration constant C is determined by the boundary condition at $r = r_2$. For most tissues at common wavelengths $r\mu_t >> 1$ within one mm from the source. In that approximation the integral approaches $\exp(\mu_s r)/(\mu_s r)^2$. Equation (6.105) then yields for the total fluence rate

$$\phi_t(r) = \frac{3P\mu_s(1-g)}{4\pi}\left[\frac{1}{r} - \frac{1}{r_2}\right] + \frac{P}{2\pi r_2^2}\frac{1+r_{21}}{1-r_{21}} + (-3+1)\frac{P\exp[-\mu_s(r-r_1)]}{4\pi r^2} \tag{6.106}$$

where terms with $\exp(-\mu_s(r_2 - r_1))$ have been neglected. The last term includes scattered ("−3") and non-scattered ("+1") fluence rate. It appears that Eq. (6.106) is also valid for $r_1 = 0$. If $r_2 \to \infty$, Eq. (6.106) reduces to Eq. (6.101) and, if in addition $r_1 \to 0$, Eq. (6.106) reduces to Eq. (6.104) ($\mu_a = \mu_{\text{eff}} = 0$).

If $r_{21} \to 1$, $\phi_t(r)$ is dominated by the second term in Eq. (6.106). For $r < r_1$ the scattered fluence rate $\phi_s(r) = \phi_s(r_1)$ is independent of r, like in an integrating sphere [28] and the total fluence rate $\phi_t(r) = [P/(2\pi r_2^2)][(1 + r_{21})/(1 - r_{21})] + P/(4\pi r^2)$.

6.4.3 Cylindrical Geometry

One example of a practical problem with axial symmetry is a "cylindrical diffuser" or "linear light source" in a tissue. This is an optical fiber covered over a certain length at the tip with a light scattering material and connected to a laser, so that light is emitted uniformly in a cylindrical pattern at a constant rate per unit length. This is applied in interstitial photodynamic therapy and interstitial thermotherapy. Another example is an optical fiber bundle touching a tissue with, for example, a central fiber emitting light and one or more fibers at some distance from the central fiber detecting backscattered light, in order to characterize the tissue by the measurement of optical properties (Section 6.4.4).

6.4.3.1 Linear Light Source as a Sum of Isotropic Point Sources

The easiest way to obtain the fluence rate at a point some distance from a linear light source is to approximate this source by a chain of n discrete isotropic point sources with power P, use Eq. (6.104) (neglecting the last term) and sum the individual contributions numerically. Let the linear source have length ℓ and be placed on the z axis with the center at $z = 0$ (Fig. 6.10). For the mathematics in the present section we describe the linear source as a continuous line of isotropic point sources with power $P^* = nP/\ell$ [W/m]. We use cylindrical polar coordinates r, θ and z, where θ cancels because of the symmetry. Let R be the distance between a point $\boldsymbol{r}(z_0, r)$ in the tissue and a point with coordinate z on the linear light source. According to Eq. (6.104), omitting the last term, the fluence rate at point $\boldsymbol{r}$ is then

$$\phi(\boldsymbol{r}) = \sum_{k=1}^{n} \frac{3P\mu_{tr}}{4\pi} \frac{\exp(-\mu_{\text{eff}} R_k)}{R_k} = \frac{3P^*\mu_{tr}}{4\pi} \int_{-\ell/2}^{+\ell/2} \frac{\exp(-\mu_{\text{eff}} R)}{R} dz \qquad (6.107)$$

where the subscript k indicates discrete sources and $R^2 = (z_0 - z)^2 + r^2$. For finite ℓ the integral cannot be obtained in closed form and the line of point sources is the best alternative.

We will now derive from Eq. (6.107) a simple formula for a linear light source of infinite length ℓ. In that case we can take $z_0 = 0$. Let $\mu_{\text{eff}} z = \zeta$ and $\mu_{\text{eff}} r = \rho > 1$.

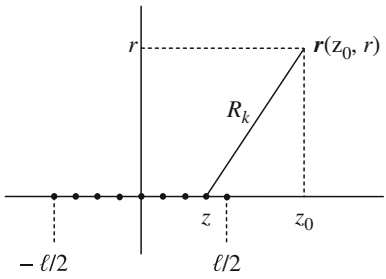


Fig. 6.10 The geometry for the calculation of the fluence rate in tissue due to a linear light source simulated by point sources (*black dots*). *Symbols* are explained in the text

Change from variable ζ to ξ as follows: $(\rho^2+\zeta^2)^{1/2} = \rho+\xi$ so that $\zeta = (\xi^2 + 2\rho\xi)^{1/2} = (2\rho\xi)^{1/2}[1+\xi/(2\rho)]^{1/2} = (2\rho\xi)^{1/2}[1+\xi/(4\rho) - \xi^2/(32\rho^2)+\ldots] \approx (2\rho\xi)^{1/2}[1+\xi/(4\rho)]$. Equation (6.107) can now be written

$$\begin{aligned}\phi(r) = \phi(\boldsymbol{r}) &= \frac{3P^*\mu_{tr}\exp(-\rho)}{2\pi\sqrt{2\rho}}\int_0^\infty \frac{\exp(-\xi)}{\sqrt{\xi}}[1-\xi/(4\rho)]d\xi \\ &= \frac{3P^*\mu_{tr}\exp(-\mu_{\text{eff}}r)}{\sqrt{8\pi}\sqrt{\mu_{\text{eff}}r}}[1-1/(8\mu_{\text{eff}}r)]\end{aligned} \tag{6.108}$$

The variable ξ is changed to y^2 and then the argument of the integral becomes the error function, for which the integrals are well known. In the last formula we may neglect the factor $[1-1/(8\mu_{\text{eff}}r)] \approx 1$, because of the assumption that $\mu_{\text{eff}}r = \rho > 1$.

6.4.3.2 Fluence Rate About a Linear Light Source from the Diffusion Equations

In this section we calculate the fluence rate as a function of the distance from an infinitely long linear light source, to show an alternative route to and validate Eq. (6.108). Because of the symmetry, of the cylindrical polar coordinates r, θ and z only r occurs in the equations. Equations (6.26) and (6.31) take the form

$$(1/r)\partial\{r[F_{r+}(r)-F_{r-}(r)]\}/\partial r = -2\mu_a[F_{r+}(r)+F_{r-}(r)] + S_{c1}(r) \tag{6.109a}$$

$$\partial[F_{r+}(r)+F_{r-}(r)]/\partial r = -(3/2)\,\mu_{tr}[F_{r+}(r)-F_{r-}(r)] + S_{c2}(r) \tag{6.109b}$$

where $S_{c1}(r)$ and $S_{c2}(r)$ are the source terms for this geometry, which follow from Eqs. (6.26) and (6.31) respectively. Since the source is a diffuser, like in Eq. (6.104) we neglect non-scattered light and assume that the light is "diffuse" for all r larger than the radius of the source. We thus omit the source terms. From Eqs. (6.109a, b) we obtain

$$r^2(\partial^2/\partial r^2)[F_{r+}(r)+F_{r-}(r)]+r(\partial/\partial r)[F_{r+}(r)+F_{r-}(r)]-\mu_{\text{eff}}^2 r^2[F_{r+}(r)+F_{r-}(r)] = 0 \tag{6.110}$$

The solutions of this equation are the Bessel functions $I_0(\mu_{\text{eff}}r)$ and $K_0(\mu_{\text{eff}}r)$, but only $K_0(\mu_{\text{eff}}r)$ applies because $I_0(\mu_{\text{eff}}r)$ diverges when $r \to \infty$. We thus find that $F_{r+}(r) + F_{r-}(r) = CK_0(\mu_{\text{eff}}r)$ and from Eq. (6.109b): $F_{r+}(r) - F_{r-}(r) = -[2/(3\mu_{tr})]C\partial K_0(\mu_{\text{eff}}r)/\partial r$. Consider a value r_0 of r such that $\mu_{\text{eff}}r_0 << 1$ and that within a cylinder of unit length and radius r_0 only a small fraction of the source power P^* is absorbed. Then, the net outward flux $F_{r+}(r_0) - F_{r-}(r_0) = P^*/(2\pi r_0)$, so that $C = -3\mu_{tr}P^*/[(4\pi r_0)(\partial K_0(\mu_{\text{eff}}r)/\partial r)]$ for $r = r_0$. If $\mu_{\text{eff}}r_0 << 1$, $K_0(\mu_{\text{eff}}r)$ can be approximated by $-\ell n(\mu_{\text{eff}}r)$, so that $C = 3\mu_{tr}P^*/(4\pi)$.

We now consider values of r for which $\mu_{\text{eff}} r > 1$. In that case $K_0(\mu_{\text{eff}} r) \approx [\pi/(2r)]^{1/2} \exp(-\mu_{\text{eff}} r)[1 - 1/(8\mu_{\text{eff}} r)]$ so that the fluence rate becomes

$$\phi(r) = 2[F_{r+}(r) + F_{r-}(r)] = \frac{3\mu_{tr} P^*}{2\pi} \sqrt{\frac{\pi}{2r}} \exp(-\mu_{\text{eff}} r)[1 - 1/(8\mu_{\text{eff}} r)] \quad (6.111)$$

Equation (6.111) is identical to Eq. (6.108), which validates the description of a line source by a chain of point sources. The latter is more convenient in numerical calculations for sources of finite length. Like in Section 6.4.3.1 we may approximate $1 - 1/(8\mu_{\text{eff}} r) \approx 1$, because $\mu_{\text{eff}} r > 1$.

Conservative scattering was discussed for the plane (Sections 6.4.1.3 and 6.4.1.6) and for the spherical geometry (Section 6.4.2.3). For a line source of infinite length, Eq. (6.111) cannot be used for that purpose because it was assumed that $\mu_{\text{eff}} r > 1$, so that μ_{eff} cannot be set to zero in Eq. (6.111). In fact, for an infinite linear light source and conservative scattering, the fluence rate diverges if the tissue volume is not finite.

Consider a cylinder of tissue in air with radius r_c and infinite length, with a line source on the axis. For conservative scattering, at all r the outward flux must be equal to the flux from the source, i.e. $F_{r+}(r) - F_{r-}(r) = P^*/(2\pi r) = -D\partial\phi(r)/\partial r$, where $D = (1/3\mu_{tr})$ is the diffusion coefficient. Integration of this equation yields $\phi(r) = -[P^*/(2\pi D)]\ell n(r) + \text{constant}$. We write the boundary condition at $r = r_c$ in the form $\phi(r_c) - 2DA[\partial\phi(r)/\partial r]_{r=rc} = 0$ (Section 6.3.8.1) where $A = (1 + r_{21})/(1 - r_{21})$. The fluence rate then becomes

$$\phi(r) = -[P^*/(2\pi D)]\ell n(r/r_c) + [P^*/(\pi r_c)](1 + r_{21})/(1 - r_{21}) \quad (6.112)$$

Equation (6.112) diverges when $r_c \to \infty$. In an infinite non-absorbing medium transport by diffusion is apparently not fast enough to carry away all photons produced by a linear light source of infinite length.

6.4.4 *Spatially Resolved Reflection and Fluence Rate Using the Diffusion Dipole Model*

We are interested in the fluence rate in and the reflection from a semi-infinite medium illuminated by a collimated light beam or by an optical fiber touching the medium. The diffusion theory of this problem has been developed by Reynolds et al. [30] and adapted to refractive index mismatched situations by Groenhuis et al. [31]. The analytical solutions involve integrals and infinite sums of Bessel functions [30, 31]. In order to extract optical properties from experimental data [32] or calculate the fluence rate in the medium [33], it would be more convenient to have simple closed-form solutions. This can be achieved with the "diffusion dipole model."

A pencil beam or an optical fiber as light source can be simplified by first noting that anisotropic scattering can be transformed to isotropic scattering using the delta-Eddington approximation (Section 6.3.9, Table 6.1). The last term in the

diffusion equation (Eq. (6.36)) then becomes zero ($g^* = 0$) and $\mu_t' = \mu_{tr}' = \mu_a + \mu_s' = \mu_a + \mu_s(1-g)$. When a semi-infinite medium is illuminated by a pencil beam along the z-axis (positive into the medium), the third term in Eq. (6.36) is $3\mu_s'\,\mu_{tr}'\,E_0 \exp(-\mu_t' z)\,\delta(x-x_0)\,\delta(y-y_0)$ where E_0 is the irradiance of the incident beam [Wm^{-2}] after specular reflection at the surface and x_0, y_0 are the coordinates of the spot where the beam enters the medium. This source term may be seen as a line of isotropic sources along the z-axis in the medium, with decreasing power according to $\exp(-\mu_t' z)$. Alternatively, since almost 90% of the scattering occurs within two optical mean free paths from the surface, one could approximate the pencil beam or optical fiber by one isotropic source at a depth $z \approx 1/\mu_t'$, i.e. the optical mean free path [32]. This makes it possible to apply the diffusion dipole model.

6.4.4.1 Diffusion Dipole Model

Let an isotropic point source (1) (Fig. 6.11a) with power P be located at depth z_0 in an infinite scattering an absorbing medium. The fluence rate at B equals $\phi_1(r,z) = [3P\mu_{tr}'/(4\pi)]\exp(-\mu_{\mathrm{eff}}\rho_1)/\rho_1$ (Eq. (6.104)). Everywhere in the medium, except at the source position, $\phi_1(r,z)$ satisfies the source free diffusion equation

$$\nabla^2\,\phi(r,z) - \mu_{\mathrm{eff}}^2\,\phi(r,z) = 0 \tag{6.113}$$

Another isotropic point source (2) located at depth $-z_0$ contributes at point B the fluence rate $\phi_2(r,z) = [3P\mu_{tr}'/(4\pi)]\exp(-\mu_{\mathrm{eff}}\rho_2)/\rho_2$ and $\phi_2(z,r)$ is also a solution of Eq. (6.113). Since $\rho_1^2 = (z-z_0)^2+r^2$ and $\rho_2^2 = (z+z_0)^2+r^2$ the function $\phi_d(r,z) = \phi_1(r,z)-\phi_2(r,z) = 0$ for $z=0$. Subtracting ϕ_2 can be seen as adding the fluence rate

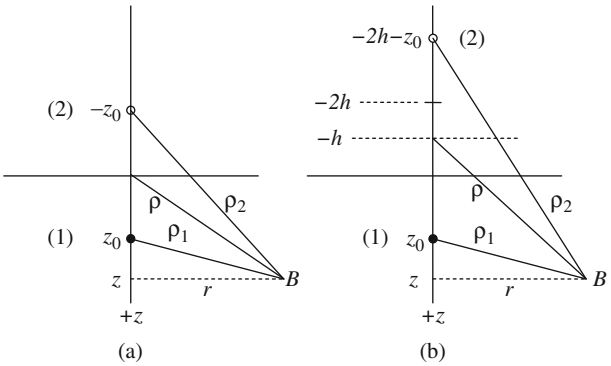


Fig. 6.11 (**a**) Diffusion dipole according to Fretterd and Longini [34]; $\rho_1^2 = (z-z_0)^2 + r^2$, $\rho_2^2 = (z+z_0)^2 + r^2$, $\rho^2 = z^2 + r^2$. $\phi_d(0,r) = 0$ so that the boundary condition for diffusion theory, Eq. (6.116), is not satisfied. (**b**) Diffusion dipole after shifting of the mirror plane, to approximately satisfy the boundary condition; $\rho_1^2 = (z-z_0)^2+r^2$, $\rho_2^2 = (z+2\,h+z_0)^2+r^2$, and $\rho^2 = (z+h)^2+r^2$

of a negative source ("image source"). The plane $z = 0$ is called the "mirror plane." This procedure is identical to the calculation of the electric potential about a point charge in the vicinity of a grounded conducting plane. The combination of the point sources (1) and (2) was therefore called a diffusion dipole by Fretterd and Longini [34] who introduced it. A negative light source (the "image source"), is nonexistent; it is just a mathematical construct. We are interested in the fluence rate due to an isotropic point source just below the surface in a semi-infinite medium ($z > 0$). The diffusion dipole can replace this point source and thus solve the problem, if it can be modified in such a way that the incorrect boundary condition, i.e. $\phi_d(r, 0) = 0$, is replaced by the boundary condition required by diffusion theory.

6.4.4.2 Boundary Condition and Solution

Rather than $\phi_d(r, 0) = 0$ (see previous section) we must apply the boundary condition of Eq. (6.48) with $\boldsymbol{S}_1(\boldsymbol{r}) = 0$, that is

$$\phi_d(r, z) - h\partial\phi_d(r, z)/\partial z = 0 \quad \text{for all } r \text{ and } z = 0 \tag{6.114}$$

where $h = 2D(1 + r_{21})/(1 - r_{21})$. We approximate $\phi_d(r, 0) - h[\partial\phi_d(r, z)/\partial z]_{z=0} \approx \phi_d(r, -h)$, i.e. we cut off Taylor's series after the second term. We see that the boundary condition is approximately satisfied on the plane $z = -h$ (the extrapolated boundary).

Allen and McKenzie [33] pointed out that the plane on which Eq. (6.114) holds is curved. Since $\phi_d(r, z) = \phi_1(r, z) - \phi_2(r, z)$ is known (previous section) a formula can be derived for the coordinate $z_m(r, z)$ where Eq. (6.114) holds exactly (note that the minus sign in Eq. (6.114) makes z_m positive, even though the extrapolated boundary is on the negative side of the z-axis). First define $f(r, z) = [3P\mu'_{tr}/(4\pi)]\exp(-\mu_{\text{eff}}\rho)/\rho$, where $\rho^2 = r^2 + z^2$. Express $\phi_1(r, z)$ and $\phi_2(r, z)$ as Taylor's series in $f(r, z)$ and its derivatives: $\phi_1(r, z) = f(r, z) - z_0\partial f(r, z)\partial z + (z_0^2/2)\partial^2 f(r, z)\partial z^2 - (z_0^3/6)\partial^3 f(r, z)\partial z^3 + \ldots$ and $\phi_2(r, z) = f(r, z) - z_0\partial f(r, z)\partial z + (z_0^2/2)\partial^2 f(r, z)\partial z^2 + (z_0^3/6)\partial^3 f(r, z)\partial z^3 + \ldots$ so that $\phi_d(r, z) = \phi_1(r, z) - \phi_2(r, z) = -2z_0\partial f(r, z)\partial z - 2(z_0^3/6)\partial^3 f(r, z)\partial z^3 + \ldots$. After calculating the derivatives one can see that the second term can be neglected relative to the first if $(z_0/\rho)^2 << 1$, or $(\rho\mu'_t)^2 >> 1$, and in addition $\mu_a << \mu'_t$ is required, the condition for the Eddington approximation.

We now have $\phi_d(r, z) \approx -2z_0\partial f(r, z)\partial z = [3P\mu'_{tr}/(2\pi)](z_0 z)(\mu_{\text{eff}} + 1/\rho)\exp(-\mu_{\text{eff}}\rho)/\rho^2$. This is the formula for the dipole model that was used by Fretterd and Longini [34] and by Allen and McKenzie [33]. Note that the "strength" q of the dipole equals $q = 2z_0 P$. Using the definition of $f(r, z)$, substitution into Eq. (6.114) gives

$$z_m(r, z) = h\left[1 - \frac{z_m^2}{\rho^2}\frac{(\mu_{eff}^2\rho^2 + 3\mu_{eff}\rho + 3)}{1 + \mu_{eff}\rho}\right] \tag{6.115}$$

When is $z_m{=}h$? If $\mu_{\text{eff}}\rho{>>}1$, Eq. (6.115) reduces to $z_m{\approx}h[1{-}(z_m\mu_{\text{eff}})^2/(\mu_{\text{eff}}\rho)]$. Since $z_m{\approx}h, (h\mu_{\text{eff}})^2/\mu_{\text{eff}}\rho{=}(4/3)[(1+r_{21})^2/(1-r_{21})^2](\mu_a/\mu'_{tr})/(\mu_{\text{eff}}\rho) = (11.4)(\mu_a/\mu'_{tr})/(\mu_{\text{eff}}\rho)$ for $r_{21}{=}0.49$ (Table 6.2). If $\mu_a/\mu'_{tr} \leq 0.1$, the boundary condition ($z_m{=}h$) is well satisfied for $\mu_{\text{eff}}\rho >> 1$. For very small absorption ($\mu_{\text{eff}} \to 0$), the latter condition requires very large ρ. However, in that limit Eq. (6.115) reduces to $z_m{=}h[1-3(z_m/\rho)^2]$. Substituting $z_m{\approx}h$ between the brackets we obtain $(\rho\mu'_{tr}){>>}3.4$ as requirement for the validity of the boundary condition $z_m{=}h$. For a connection between the two requirements the reader is referred to Exercise 6.8.

The new boundary condition requires the mirror plane to be placed at $z = -h$. Since the (real) source remains at $z = z_0$, the image source should be at $z = -2\,h{-}z_0$ (Fig. 6.11b) and $\rho_1^2 = (z{-}z_0)^2{+}r^2$ is unchanged whereas $\rho_2^2 = (z{+}2\,h{+}z_0)^2{+}r^2$. The solution of the diffusion dipole problem describing a single point source at $z = z_0$ can now be written

$$\phi_d(r,z) = \frac{3P\mu'_{tr}}{4\pi}\left[\frac{\exp(-\mu_{eff}((z-z_0)^2+r^2)^{1/2})}{((z-z_0)^2+r^2)^{1/2}} - \frac{\exp(-\mu_{eff}((z+2\,h+z_0)^2+r^2)^{1/2})}{((z+2\,h+z_0)^2+r^2)^{1/2}}\right] \tag{6.116}$$

Equation (6.116) satisfies the diffusion equation (Eq. (6.113)) and also the boundary condition (6.114). Furthermore, close to the real source the relative contribution of the image source can be neglected and the exponential approaches 1 for small ρ_1. The total outward flux from the real source then equals $4\pi\rho_1^2\,|\boldsymbol{F}(\rho_1)| = 4\pi\rho_1^2 D\,|\nabla\phi_d| = P$ which is essentially also a boundary condition.

Since $h = 2D(1+r_{21})/(1-r_{21}) = [2/(3\mu'_t)](1+r_{21})/(1-r_{21})$ and $z_0 \approx 1/\mu'_t$ (assumption by Farrell et al. [32]) we have for refractive index matching $h \approx 2/(3\mu'_t)$ and for index mismatch ($n = 1.4$ for tissue) $r_{21} = 0.49$ (Table 6.2) so that $h \approx 2/\mu'_t \approx 2z_0$.

Both Fretterd and Longini [34] and Allen and McKenzie [33] used $\phi_d(z,r) \approx -2z_0\partial f(r,z)\partial z$ (see above) rather than Eq. (6.116). In addition, Allen and McKenzie applied Eq. (6.114) as boundary condition, shifting the mirror plane by h, so that

$$\phi_d(r,\,z) \approx [3P\mu'_{tr}/(2\pi)](z_0+h)(z+h)(\mu_{\text{eff}}+1/\rho)\exp(-\mu_{\text{eff}}\rho)/\rho^2 \tag{6.117}$$

where $\rho^2 = (z+h)^2 + r^2$ (Fig. (6.11b)). Finally, z_0 was set to zero [33], which is not entirely consistent with the physics of the model.

Equation (6.117) is nearly equal to Eq. (6.116) if $\rho >> (z_0{+}h)$, which is a much stronger criterion than before the shift of the mirror plane (i.e. $\rho >> z_0$, see above). We therefore prefer Eq. (6.116).

Exercise 6.8 Condition for $z_m = h$. In biological tissues μ_a varies much more than μ_s between tissues and as a function of wavelength. It is therefore useful to write $\mu_a = \eta\mu'_{tr}$, where $\eta = 1 - a' << 1$ (condition for the Eddington approximation).

Express the requirements for $z_m = h$ from Section 6.4.4.2 in terms of η and μ'_{tr} and show that they are equal when $\eta = 0.26$ ($\eta = 2.3$ for index matching). For larger values of η the condition for the Eddington approximation is hardly satisfied. Therefore, the general condition is that $\rho\mu'_{tr} >> 3.4$ (>> 1 for index matching). Using the optical properties of Exercise 6.6, giving $\mu'_{tr} = 3.6\ \text{mm}^{-1}$, this means that $\rho >> 1$ mm (>> 0.3 mm for index matching).

6.4.4.3 Wide Beam Collimated Irradiance and the Condition for $z_0 = 1/\mu'_t$

From Eq. (6.116) we derive the fluence rate due to a wide beam of collimated light. According to the assumption at the end of Section 6.4.4, this beam can be represented by a sheet of isotropic point sources at depth z_0 in tissue. Let P' [Wm^{-2}] be the surface density of point sources with power P. The fluence rate produced by the point sources is calculated using $\phi_d(z, r)$, Eq. (6.116), which is integrated over all dipole pairs:

$$\phi_d(z, z_0) = \int_{r=0}^{\infty} \phi_d(r, z)\, 2\pi r\, dr \tag{6.118}$$

where P in Eq. (6.116) has been replaced by P'. In performing the integration it is necessary to distinguish $0 \le z \le z_0$ and $z \ge z_0$. The results are

$$0 \le z \le z_0: \quad \phi_d(z, z_0) = \frac{3P'\mu'_{tr}}{\mu_{\text{eff}}} \exp(-\mu_{\text{eff}}(z_0 + h)) \sinh(\mu_{\text{eff}}(z + h)) \quad [\text{Wm}^{-2}] \tag{6.119a}$$

$$z \ge z_0: \quad \phi_d(z, z_0) = \frac{3P'\mu'_{tr}}{\mu_{\text{eff}}} \exp(-\mu_{\text{eff}}(z + h)) \sinh(\mu_{\text{eff}}(z_0 + h)) \quad [\text{Wm}^{-2}] \tag{6.119b}$$

For $z > z_0$ we want the result of the integration to be equal to Eq. (6.83), i.e. the fluence rate in a semi-infinite medium irradiated by an infinitely wide collimated light beam, that is

$$\frac{3P'\mu'_{tr}}{\mu_{\text{eff}}} \exp(-\mu_{\text{eff}}(z + h)) \sinh(\mu_{\text{eff}}(z_0 + h)) = \frac{(5 - r_{21})E_0}{1 - r_{21}} \frac{\exp(-\mu_{\text{eff}} z)}{1 + h\mu_{\text{eff}}} \tag{6.120}$$

Here $\mu'_{tr} = \mu_a + \mu'_s$. We have omitted the term with $\exp(-\mu_t z)$ from Eq. (6.83) because it decreases rapidly for $z > z_0$. Since the light of each pencil beam was assumed to be converted to an isotropic point source, $P' = (\mu'_s/(\mu_a + \mu'_s))E_0$ [Wm^{-2}] where $\mu'_s/(\mu_a + \mu'_s) = a'$, the reduced albedo, accounts for the fraction of attenuated light actually scattered. To derive an expression for z_0 from Eq. (6.120), drop $\exp(-\mu_{\text{eff}} z)$ from both sides. Consider very small absorption, that is $\mu_{\text{eff}} \to 0$ and $a' \to 1$. Equation (6.120) then reduces to $3\mu'_{tr}(z_0 + h) = (5 - r_{21})/(1 - r_{21})$. Since $h = 2D(1 + r_{21})/(1 - r_{21})$ we obtain from Eq. (6.120) that $z_0 = 1/\mu'_{tr} = 1/\mu'_s$. This result is not limited to very small absorption, however.

For non-zero absorption, if $\mu_a/\mu'_{tr} \leq 0.1$ and $z_0 \approx 1/\mu'_{tr}$, $\sinh(\mu_{\text{eff}}(z_0 + h))\exp(-\mu_{\text{eff}}h)$ can be approximated by $\mu_{\text{eff}}(z_0 + h)/(1 + h\mu_{\text{eff}})$ with an error of less than 10% (see Exercise 6.9). Eq. (6.120) now reduces to $3a'\mu'_{tr}(z_0 + h) = (5 - r_{21})/(1 - r_{21})$ and

$$z_0 \approx 1/\mu'_{tr} = 1/\mu'_t \tag{6.121}$$

if $\mu_a/\mu'_{tr} \leq 0.1$ (so that $\mu'_s \approx \mu'_{tr}$ or $a' \approx 1$), the requirement for the Eddington approximation. Equation (6.121) is just the assumption of Farrell et al. [32]. It should be noted that if we do not set $a' \approx 1$ we find that $z_0 = (1 + 2.9\mu_a/\mu'_{tr})/\mu'_{tr}$ (see Exercise 6.9).

Exercise 6.9 Like in Exercise 6.8, write $\mu_a = \eta\mu'_{tr}$ and show numerically that if $z_0 = 1/\mu'_{tr}$ the approximation $\sinh(\mu_{\text{eff}}(z_0 + h))\exp(-\mu_{\text{eff}}h) \approx \mu_{\text{eff}}(z_0 + h)/(1 + h\mu_{\text{eff}})$ holds well for $\eta < 0.15$ with an error of 10% or less. Note that for the largest values of η the arguments of sinh and exp are of order 1, so that expansion in terms of η does not seem to help. If $a' = 1 - \eta$ (and not $a' = 1$) is substituted in the equation $3a'\mu'_{tr}(z_0 + h) = (5 - r_{21})/(1 - r_{21})$, show that $z_0 \approx (1 + (1 + 2A/3)\eta)/\mu'_{tr} = (1 + 2.9\eta)/\mu'_{tr}$ (index mismatch). This restricts the validity of the approximation six lines up. It is then better to substitute $\mu_a = \eta\mu'_{tr}$ directly in Eq. (6.120) without approximations and, after dropping $\exp(-\mu_{\text{eff}}z)$, to show numerically that up to $\eta = 0.3$, $z_0 = 1/\mu'_{tr}$ with an error of less than 10%. Note that Eq. (6.83), i.e. the right hand side of Eq. (6.120), is also an approximation.

The approximation in the second line above seems to appear out of the blue, arising from the end result to be achieved. If you prefer analytical over numerical methods, note the following. Series expansion of $\sinh(x)$ yields $x(1 + x^2/6 + x^4/120 + \ldots)$, which converges rapidly, even if $x \approx 1$. For $y < 1$, $\exp(y)/(1+ y)$ can be written as the product of the series expansion of $\exp(y)$ and the arithmetic series $1/(1+ y) = 1 - y + y^2 - y^3 + \ldots$ and the result is $\exp(y)/(1+ y) \approx 1 + \alpha y^2$ where $\alpha = \sum_{n=2}^{\infty} \frac{(-1)^n}{n!} = 0.368$. This approximation appears to hold within 1–2% up to $y = 2$. Note that $x = (1+2A/3)(3\eta)^{1/2}$ and $y = (2A/3)(3\eta)^{1/2}$, where $A = (1+r_{21})/(1-r_{21})$. Substitution for x and y shows that $(1 + x^2/6 + \ldots)$ and $(1 + 0.37y^2)$ increase at approximately the same rate as η increases and that the ratio of the two differs from 1 by not more than 10% up to $\eta = 0.15$. This proves the approximation in line two of this exercise.

Now consider $0 \leq z \leq z_0$. Here Eq. (6.119a) is in poor agreement with Eq. (6.83) because $\exp(-\mu'_{tr}z)$ is not small and for the diffusion dipole model the wide collimated beam has been replaced by a sheet of point sources. At $z = 0$, for very small μ_{eff}, $\exp(-\mu_{\text{eff}}(z_0 + h)) \approx 1$ and $\sinh(h\mu_{\text{eff}}) \approx h\mu_{\text{eff}}$. Since $P' = (\mu'_s/(\mu_a + \mu'_s))E_0$, $\phi_d(0, z_0) \approx 3E_0\mu'_s h \approx 2E_0(1 + r_{21})/(1 - r_{21})$. This is equal to the scattered part of the fluence rate at $z = 0$ in Eq. (6.83). Since Eq. (6.83) includes the collimated irradiance, for $E_0 = 1$ the scattered part equals $(5 - r_{21})(1 - r_{21}) - 3 = 2(1 + r_{21})/(1 - r_{21}) = 2A$. For very small absorption, the single point source approximation thus only reproduces the scattered part of the

fluence rate at $z = 0$ (like for $z > z_0$, see above). The total diffuse reflection involves only scattered light and is still correctly predicted (see also Sections 6.4.4.4 and 6.4.4.5). Derived from Eq. (6.119a), it is equal to $(1 - r_{21}F_-(0)) = \phi_d(0, z_0)/2A$, which can be shown numerically to differ by less than 10% from the value derived from Eq. (6.82b) or Eq. (6.83), if $\mu_a/\mu'_{tr} < \approx 0.1$ and $g > 0.8$.

The fluence rate $\phi_d(z, z_0)$ represented by Eq. (6.119a, b) shows a maximum at $z = z_0$. Such a maximum also follows from Eq. (6.83) (depending on the value of g) and is caused by the build-up of scattered fluence upon irradiation by a collimated light beam. However, $\partial\phi_d(z)/\partial z$ from Eqs. (6.119a, b) shows a discontinuity at $z = z_0$, which is an artifact of the model.

6.4.4.4 Reflectance in the Single Point Source Model

Combined with Eq. (6.2), for a semi-infinite geometry Eqs. (6.13a, b) can be rewritten as $F_+(z, r) = \phi(r, z)/4 - (D/2)\partial\phi(r, z)/\partial z$ and $F_-(r, z) = \phi(r, z)/4 + (D/2)\partial\phi(r, z)/\partial z$, where F_+ and F_- are the hemispherical fluxes (Section 6.2.3) and $F_+ - F_-$ is the component of the flux vector $\boldsymbol{F}(r, z)$ along the z-axis. When no sources are present at the boundary, the boundary condition is $F_+(r, 0) = r_{21}F_-(r, 0)$ (Section 6.2.5), which is identical to Eq. (6.114) (substitute the foregoing expressions for F_+ and F_- and the expression for h). Since in Eq. (6.116) $P = (\mu'_s/\mu'_t)E_0$ [W], because we consider a pencil beam and $D = (1/3\mu_{tr}{}')$ the reflectance can be written as $R(r, z_0) = (1 - r_{21})F_- = D\partial\phi_d(r, z)/\partial z$ for $z = 0$ or

$$R(r, z_0) = \frac{\mu'_s E_0}{4\pi\mu'_t}\left[z_0\left(\mu_{\text{eff}} + \frac{1}{\rho_1}\right)\frac{\exp(-\mu_{\text{eff}}\rho_1)}{\rho_1^2} + (z_0 + 2h)\left(\mu_{\text{eff}} + \frac{1}{\rho_2}\right)\frac{\exp(-\mu_{\text{eff}}\rho_2)}{\rho_2^2}\right] \tag{6.122a}$$

[Wm^{-2}], where $\rho_1^2 = z_0^2 + r^2$ and $\rho_2^2 = (z_0 + 2h)^2 + r^2$. Equation (6.122a) was used by Farrell et al. [32] with $z_0 = 1/\mu'_t$. An alternative formula for the reflectance can be obtained from Eq. (6.114), i.e. $R(r, z_0) = D[\partial\phi_d(r, z)/\partial z]_{z=0} = [(D/h)\phi_d(r, z)]_{z=0} = \phi_d(r, 0)/(2A)$ so that

$$R(r,\, z_0) = \frac{\mu'_s E_0}{4\pi h\mu'_t}\left[\frac{\exp(-\mu_{\text{eff}}\rho_1)}{\rho_1} - \frac{\exp(-\mu_{\text{eff}}\rho_2)}{\rho_2}\right] \quad [\text{Wm}^{-2}] \tag{6.122b}$$

Eq. (6.122b) is equivalent to Eq. (6.122a) but not identical, because the two are related by the boundary condition, Eq. (6.114), which is satisfied if $\rho\mu'_{tr} >> 1$ (Section 6.4.4.2) but not for small r ($\rho^2 = (z+h)^2 + r^2$, Fig. (6.11b)). There is, however, no theoretical reason to prefer Eq. (6.122a) over Eq. (6.122b). The difference merely shows a limitation of the model.

The total diffuse reflectance $R_d(z_0)$ is obtained by integrating Eq. (6.122a) as follows

$$R_d(z_0) = \int_{r=0}^{\infty} R(r,\, z_0)2\pi r dr = \frac{\mu'_s E_0}{2\mu'_t}[\exp(-\mu_{\text{eff}}(z_0 + 2h))] \quad [\text{W}] \tag{6.123a}$$

For conservative scattering $\mu_{\text{eff}} = 0$ and $\mu_s' = \mu_t'$ so that $R_d = E_0$ expressing the conservation of energy. Alternatively we calculate the total reflectance using Eq. (6.122b) in the integration and obtain

$$R_d(z_0) = \frac{\mu_s' E_0}{h\mu_t' \mu_{\text{eff}}} \exp(-\mu_{\text{eff}}(z_0 + h)) \sinh(\mu_{\text{eff}} h) \quad [\text{W}] \tag{6.123b}$$

Equation (6.123b) has been put in a form facilitating comparison with Eq. (6.119a). If we apply the formula $R_d = [(D/h)\phi_d(r,z)]_{z=0}$ to Eq. (6.119a) we obtain Eq. (6.123b). If $\mu_s'/\mu_t' \leq 0.01$ both Eq. (6.123a) and Eq. (6.123b) approach $R_d(z_0) \approx E_0(\mu_s'/\mu_t') \exp(-\mu_{\text{eff}} z_0)/(1 + \mu_{\text{eff}} h)$ with an error of less than 3% (see Exercise 6.10).

Exercise 6.10 If $\eta = \mu_a/\mu_t' \leq 0.01$ and $x = \mu_{\text{eff}} h = (2A/3)\sqrt{3\eta}$ then $x \leq 0.34$ if $r_{21} \leq 0.49$. Now first show that $\exp(2x) \approx (1 + x)/(1 - x)$ is a good approximation if $x < 0.4$. Apply this approximation to $\exp(-2\mu_{\text{eff}} h)$ to show that both Eqs. (6.123a) and (6.123b) approach $E_0(\mu_s'/\mu_t') \exp(-\mu_{\text{eff}} z_0)/(1 + \mu_{\text{eff}} h)$. Subsequently show that Eq. (6.123) can be written in the form $R_d(z_0) = \frac{a' E_0}{2} \exp\left(-\sqrt{3(1-a')}\right)\left[1 + \exp\left((-4A/3)\sqrt{3(1-a')}\right)\right]$ which is Eq. (22) of Farrell et al. [32] if $E_0 = 1$, $z_0 = 1/\mu_t'$ and a' is the reduced albedo. Show also that Eq. (6.123b) can be written in the form $R_d(z_0) = \frac{a' E_0}{(4A/3)\sqrt{3(1-a')}} \exp\left(-\sqrt{3(1-a')}\right)\left[1 - \exp\left((-4A/3)\sqrt{3(1-a')}\right)\right]$ and that if $\mu_a/\mu_t' \leq 0.01$ both expressions approach $R_d \approx a' E_0 \exp\left(-\sqrt{3(1-a')}\right) / \left(1 + (2A/3)\sqrt{3(1-a')}\right)$.

6.4.4.5 Pencil Beam as a Continuous Line of Point Sources

A pencil beam would be better described by a line of isotropic point sources along the positive z-axis, rather than by a single point source at $z = z_0$, as we did above. The point source power over a depth dz_0 is then $\mu_s' E_0 \exp(-\mu_t' z_0)\, dz_0$ (E_0 in [W]) and this expression replaces $\mu_s' E_0/\mu_t'$ in Eq. (6.122a) and Eq. (6.122b). The reflectance is

$$R_d(r) = \int_{z_0=0}^{\infty} R_d(z_0)\, \mu_s' \exp(-\mu_t' z_0)\, dz_0 \quad [\text{Wm}^{-2}] \tag{6.124}$$

Substituting Eq. (6.122a) yields Eq. (19) of Farrell et al. [32]. Alternatively, Eq. (6.122b) can be used. In both cases the integral cannot be evaluated in closed form, but numerical calculation is a relatively easy alternative. If Eq. (6.117) is used to calculate $R(r, z_0) = D[\partial\phi_d(r,z)/\partial z]_{z=0}$ or $R(r, z_0) = [(D/h)\phi_d(r,z)]_{z=0}$ and this is subsequently substituted into Eq. (6.124), the integral can be evaluated and the only factor that matters is $\int_{z_0=0}^{\infty} (z_0 + h) \exp(-\mu_t z_0) dz_0 = (1/\mu_t' + h)/\mu_t'$. The end result for $R_d(r)$ is equal to the single point source case using Eq. (6.117). This is not surprising since more than 95% of all photons are scattered or absorbed between

$z = 0$ and $z = 3/\mu_t' = 3z_0$ and Eq. (6.117) is valid for $\rho >> (z_0 + h)$ (Section 6.4.4.2).

The fluence rate for an infinitely wide incident collimated light beam can be calculated in closed form which yields an interesting result. Consider sheets of point sources with decreasing power as z_0 increases. The power density of the sources is $P' = \mu_s' E_0 \exp(\mu_t' z_0)$, ([$\mathrm{Wm}^{-3}$], E_0 in [Wm^{-2}]). We integrate $\phi_d(z, z_0)$ as follows

$$\phi_s(z) = \left[\int_{z_0=0}^{z_0=z} + \int_{z_0=z}^{z_0=\infty} \right] \phi_d(z, z_0) \mu_s' \exp(-\mu_t' z_0)\, dz_0 \tag{6.125}$$

where P' in Eqs. (6.119a, b) should be replaced by E_0. Substituting Eqs. (6.119a, b) for the respective intervals we obtain

$$\begin{aligned} \phi_s(z) = \frac{3E_0 \mu_s' \mu_{tr}'}{2\mu_{eff}} \Bigg[\exp(-\mu_{\mathrm{eff}}(z+h)) \left[\frac{\exp(\mu_{eff} h)}{\mu_t' - \mu_{\mathrm{eff}}} - \frac{\exp(-\mu_{\mathrm{eff}} h)}{\mu_t' + \mu_{\mathrm{eff}}} \right] \\ + \left[\frac{\exp(-\mu_t' z)}{\mu_t' + \mu_{\mathrm{eff}}} - \frac{\exp(-\mu_t' z)}{\mu_t' - \mu_{\mathrm{eff}}} \right] \Bigg] \end{aligned} \tag{6.126}$$

To make Eq. (6.126) more transparent we apply the following approximations. The last two terms between brackets can be rewritten as $\exp(-\mu_t' z)[1/(1+\mu_{\mathrm{eff}}/\mu_t') - 1/(1-\mu_{\mathrm{eff}}/\mu_t')]/\mu_t' \approx -(2\mu_{\mathrm{eff}}/\mu_t'^2) \exp(-\mu_t' z)$, after expansion in terms of $\mu_{\mathrm{eff}}/\mu_t'$ and neglecting higher order terms. For an error of less than 3% this requires $\mu_a/\mu_t' \leq 0.01$. With the same condition the term $\exp(-\mu_{\mathrm{eff}}(z+h))[\ldots]$ can be approximated by $\exp(-\mu_{\mathrm{eff}} z)[2\mu_{\mathrm{eff}}/\mu_t'((1/\mu_t') + h)/(1 + \mu_{\mathrm{eff}} h)]$ with a 2% error (see Exercise 6.11). Since $h = 2AD$, we find that $(1/\mu_t') + h = (1+2A/3)/\mu_t' = (5 - r_{21})/(3\mu_t'(1 - r_{21}))$. Equation (6.126) now simplifies to

$$\phi_s(z) = \frac{5 - r_{21}}{1 - r_{21}} \frac{E_0}{1 + h\mu_{\mathrm{eff}}} \exp(-\mu_{\mathrm{eff}} z) - 3E_0 \exp(-\mu_t' z) \tag{6.127}$$

Equation (6.127) equals the scattered part of Eq. (6.83). Indeed, the point sources only represent scattered light. The total fluence rate is obtained upon adding the fluence rate of the incident collimated light beam, i.e. $E_0 \exp(-\mu_t' z)$, yielding an expression identical to Eq. (6.83). Thus, the fluence rate for wide beam collimated irradiation of a semi infinite medium, calculated with the diffusion dipole model is consistent with the "conventional" diffusion theory calculations of Section 6.4.1.3.

Exercise 6.11 See also Exercise 6.10 for notation and approximation. Write $\exp(-\mu_{\mathrm{eff}}(z + h))[\ldots]$ from Eq. (6.126) in the form $\exp(-\mu_{\mathrm{eff}} z)\left[(1-\lambda)^{-1} - \exp(-2x)(1+\lambda)^{-1}\right]/\mu_t'$ where $\lambda = (3\eta)^{1/2}$. Approximate $\exp(-2x)$ and obtain $2\exp(-\mu_{\mathrm{eff}} z)(\lambda + x)/\left[(1-\lambda^2)(1+x)\mu_t'\right] \approx (2\mu_{\mathrm{eff}}'/\mu_t') \exp(-\mu_{\mathrm{eff}} z)\left[(1/\mu_t' + h)/(1 + \mu_{\mathrm{eff}} h)\right]$ with an error of less than 2% if $\eta \leq 0.01$.

Now calculate the total reflectance R_d for a wide collimated incident light beam represented by a uniform distribution of continuous lines of point sources in the medium. To do so, replace in Eq. (6.123a) $\mu_s' E_0/\mu_t'$ (the point source power) by $\mu_s' E_0 \exp(-\mu_t' z_0)\, dz_0$ and integrate over z_0 from 0 to ∞. Show that $R_d = \frac{a'}{2} E_0 \left[1 + \exp\left(-(4A/3)\sqrt{3(1-a')}\right)\right] \Big/ \left(1 + \sqrt{3(1-a')}\right)$. This is Eq. (23) of Farrell et al. [32]. An equivalent but not identical result is obtained when Eq. (6.123b) is used. Show that the integration then yields $R_d = \frac{a'}{(4A/3)\sqrt{3(1-a')}} E_0 \left[1 - \exp\left(-(4A/3)\sqrt{3(1-a')}\right)\right] \Big/ \left(1 + \sqrt{3(1-a')}\right)$. Using the approximation of $\exp(-2x)$ in Exercise 6.10 show that both formulas approach $R_d = a' E_0 \left[\left(1 + (2A/3)\sqrt{3(1-a')}\right)\left(1 + \sqrt{3(1-a')}\right)\right]^{-1}$. The latter result is also obtained when the integration as above is applied to the expression at the end of section 6.4.4.4, i.e. $R_d(z_0) \approx E_0 (\mu_s'/\mu_t') \exp(-\mu_{\text{eff}} z_0)/(1 + \mu_{\text{eff}} h)$.

Finally, show that from Eq. (6.127) one obtains the reflectance $R_d = \phi_s(0)/(2A) = (1 - \lambda)/(1 + \mu_{\text{eff}} h)$ which for $\eta \leq 0.01$ is nearly equal to the expression four lines up. This completes the consistency of the various treatments.

6.4.4.6 Discussion

The predictions of the diffusion dipole model for the fluence rate appeared to agree reasonably well with measurements of "optical isodoses" in phantoms [33]. Note that Allen and McKenzie used Eq. (6.117) with $z_0 = 0$, which differs from Eq. (6.116) at short distance. The measured fluence rate was also an average factor $k = 1.33$ larger than their theory. Since the "dipole moment" was obtained by comparison with Eq. (6.88) for diffuse incident light rather than Eq. (6.83) for collimated irradiance, the factor k should be reduced by $(5 - r_{21})/4$ so that $k \approx 1.18$.

The spatially resolved reflectance according to Eq. (6.122b) agreed better with Monte Carlo simulations [36] than Eq. (6.122a). There is another ("multipole") model which satisfies the boundary condition Eq. (6.114) exactly ("partial current boundary condition"), without a shift of the "mirror plane." This model [35, 36] involves one point source of power P at $z = z_0 = 1/\mu_{tr}'$, another point source of power P at $z = -z_0$ and a continuous line of negative point sources from $z = -z_0$ to $z = \infty$, with total power $2P$, decaying as $\exp(-l/h)$, where l is the distance from the source at $z = -z_0$. Even though the boundary condition is exactly satisfied, the agreement between this model and Monte Carlo simulations [36] was not as good as Eq. (6.122b). The model is also mathematically more complicated.

It should be kept in mind that the diffusion dipole model depends on the transformation of g to zero, using the delta-Eddington approximation (Table 6.1, p^* isotropic). For wide beam collimate irradiance, just below the tissue surface, the difference in predicted fluence rate between this approximation and other models or Monte Carlo simulations can become quite large [12]. However, if the diffusion dipole model is only used to analyze spatially resolved reflectance measurements [32] the difference between models may be less.

Even though the diffusion dipole model works well with optical phantoms [32, 37, 38], its use to non-invasively determine tissue optical properties is not without

problems. Since the model fails close to the source, one must interrogate a relatively large volume, yielding a volume average of optical properties. This problem may be alleviated by the P_3-approximation [37], which is valid much closer to the source. But at short distance from the source, tissue inhomogeneities (e.g. capillaries) may lead to large variations in locally measured optical properties [38].

6.5 Summary and Conclusion

The diffusion equation follows from Fick's law and energy conservation. In principle, these equations alone permit calculation of the radiant energy fluence rate. For quantitative results, boundary conditions are required. These are obtained from energy conservation across boundaries, expressed in terms of hemispherical fluxes. A relationship is then needed between the hemispherical fluxes and the fluence rate. This is achieved by the Eddington approximation, which assumes that the radiance $L(\boldsymbol{r},\hat{\boldsymbol{s}})$ is only slightly anisotropic (like in Fig. 6.1) as a function of the directional unit vector $\hat{\boldsymbol{s}}$. If written as a series in Legendre polynomials, the radiance is assumed to contain only two terms, a constant term $L_0(\boldsymbol{r})$ and a small term proportional to $\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{u}} = P_1(\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{u}})$ where P_1 is the Legendre polynomial of order 1, which is equal to the cosine of the angle between $\hat{\boldsymbol{s}}$ and the direction $\hat{\boldsymbol{u}}$ of the flux vector $\boldsymbol{F}(\boldsymbol{r})$. In tissue optics, the Eddington approximation is usually regarded as an integral part of the diffusion approximation.

The diffusion equation is a second order differential equation for the fluence rate. Fick's law and energy conservation are expressed as first order differential equations for the hemispherical fluxes. Since the boundary conditions are also expressed in terms of hemispherical fluxes, formulation of diffusion theory in terms of these fluxes appears to be practically useful.

The theory includes the absorption coefficient μ_a and the diffusion constant D plus unknown source terms. Expressions for D and for the source terms are obtained from the transport equation or equation of transfer, which is an integro-differential equation for the radiance. Diffusion theory is derived as an approximate solution of this equation. Integration of the transport equation over all directions $\hat{\boldsymbol{s}}$ produces the law of energy conservation (Eqs. (6.3a) and (6.26)). Multiplication of the transport equation by $\hat{\boldsymbol{s}}$ and again integrating over all $\hat{\boldsymbol{s}}$ produces Fick's law (Eqs. (6.2) and (6.31)), including the source term, provided that the Eddington approximation is valid. At the same time the relationship $D = 1/(3\mu_{tr})$ is obtained, where $\mu_{tr} = \mu_a + \mu_s(1-g)$.

The diffusion approximation holds far away from boundaries and sources (Section 6.3.5), and is not subject to the condition that $\mu_a << \mu_s(1-g)$. The Eddington approximation does require (Section 6.3.6) that $\mu_a << \mu_s(1-g)$ so that the diffusion pattern, i.e. the radiance as a function of direction $\hat{\boldsymbol{s}}$ is only slightly anisotropic.

The constants μ_a, μ_s and g can be measured, but since biological tissues scatter light predominantly in the forward direction ($g <\approx 1$), direct measurement of g is not easy. In fact, experiments often only yield μ_s and g in the combination

$\mu_s(1-g)$. When the scattering or phase function is written as a series in Legendre polynomials (Eq. (6.132)), $b_0 = 1$ and $b_1 = g$ are the only coefficients that occur in diffusion theory, and this is not without effect, in particular close to boundaries. Measurement of $b_i\,(i > 1)$ is practically impossible, but a complete phase function may be needed, e.g. in Monte Carlo simulations. The Henyey-Greenstein function (Eq. (3.48) is a useful model phase function, which involves only one parameter g ($b_i = g^i$ in Eq. (6.132)). For $g = 0$, H-G describes isotropic scattering and for $g = 1$ it represents a delta function, describing purely forward scattering. The delta-Eddington approximation introduces a delta function into the phase function in such a way that an approximation of the Henyey-Greenstein function is obtained which improves as $g \to 1$ (Table 6.1). The delta-Eddington approximation seems to be a good method to include as much as possible of a complete, strongly forward scattering model phase function in the diffusion equations. Note that in these equations, g occurs separately as well as in the form $\mu_s(1-g)$, so that the choice of the type of delta-Eddington approximation (Table 6.1, g^*) may affect the accuracy of the model.

A semi-infinite medium, a wide slab, or a spherical geometry (Sections 6.4.1, 6.4.2) are examples involving only one variable co-ordinate, which permit simple exact solutions of the diffusion theory equations, exactly satisfying the boundary conditions. A pencil beam perpendicularly incident on a semi infinite medium involves two variable coordinates. This problem can also be solved exactly, but elaborate mathematics hampers analysis of experimental data. Here, the diffusion dipole model is a useful further approximation. This model involves relatively easy calculations of fluence rate and spatially resolved reflectance, but this has a price. First, the boundary conditions are only approximately satisfied. Second, the diffusion dipole model requires $g = 0$, which is achieved with the delta Eddington approximation (lines 3–5 of Table 6.2, $g^* = 0$). Both limitations cause reduced accuracy close to the light source.

The diffusion approximation has also been called P_1 - approximation, because in a geometry with one variable co-ordinate (e.g. Section 6.4.1.3) the integrations discussed in the third paragraph of this section and in Section 6.3.4 amount to multiplication of the transport equation by the Legendre polynomials $P_0(\cos\theta) = 1$ or $P_1(\cos\theta) = \cos\theta = \hat{s}_z$ and integration over all directions $\hat{\boldsymbol{s}}$. The diffusion approximation can be improved when the transport equation is further multiplied by the Legendre polynomials $P_2(\cos\theta)$ and $P_3(\cos\theta)$, respectively, and again integrated. This yields four equations for the radiance, and the latter now has four terms, proportional to P_0, P_1, P_2 and P_3. This procedure is called the P_3 approximation (and in general P_n approximation [10]). Its solution, which requires additional, "higher order" boundary conditions [12, 16] is straightforward in simple geometries. The delta Eddington approximation in this case (shorthand δ–E(4) [18]) requires $f = g^4$ (Section 6.3.7.2) so that $b_n^* = (b_n - g^4)/(1-g^4)\,(n = 0\ldots3)$ (Eq. (6.46)) and $b_n^* = 0$ $(n > 3)$; $b_n = g^n$ $(n = 0\ldots3)$ and $b_n = g^4$ $(n > 3)$, according to the procedure of Table 6.1. δ–E(4) Appears to agree very well with Monte Carlo simulations, which employ the full Henyey-Greenstein phase function [12]. The P_3 approximation was also shown to agree better than diffusion theory with Monte

Carlo simulations of reflectance from a semi infinite medium perpendicularly irradiated by a pencil beam and also better predicted phantom measurements [37]. Since Monte Carlo simulations are time consuming, these might be replaced by P_3 calculations in some cases. For real tissues, because of inhomogeneities and inaccurate optical properties, the P_3 approximation may be too good and diffusion theory is probably adequate, in particular if the geometry is not simple.

Appendix 1: Delta Functions

A delta function $\delta(x - x_0)$ is defined in an integral as follows. If the function $f(x)$ is continuous at $x = x_0$,

$$\int_{-\infty}^{+\infty} f(x)\ \delta(x - x_0)\ dx = f(x_0) \tag{6.128}$$

The delta function can be pictured as a function that peaks at $x = x_0$ and is zero everywhere else. If $f(x) \equiv 1$, the delta function appears to be normalized to 1. In three dimensions (two mathematical dimensions) we show that the following relationship holds (see also Chapter 3, Section 3.4.3):

$$\delta(1 - \hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{s}}_0)/(2\pi) = \delta(\mu - \mu_0)\delta(\varphi - \varphi_0) \tag{6.129}$$

where $\mu = \cos\theta$ (Fig. 3.5). Equation (6.129) can be validated when it is multiplied by a continuous function $f(\hat{\boldsymbol{s}}) = f(\varphi, \mu)$ and integrated over all directions of the unit vector $\hat{\boldsymbol{s}}$. The right hand side then yields $f(\hat{\boldsymbol{s}}_0)$. On the left hand side, let the z-axis be in the direction of $\hat{\boldsymbol{s}}_0$, so that $\mu_0 = 1$ and $\delta(1 - \hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{s}}_0) = \delta(1 - \mu)$. Integration over μ yields $f(\mu_0) = f(\hat{\boldsymbol{s}}_0)$. Since $f(\hat{\boldsymbol{s}})$ is continuous, $f(\hat{\boldsymbol{s}}_0)$ is independent of φ. Integration over φ therefore yields 2π, which is cancelled by $1/(2\pi)$ in the integrand. This proves Eq. (6.129).

Appendix 2: Legendre Polynomials

The diffusion equation contains the Laplace operator ∇^2 (see Eqs. (6.3b) and (6.36)). When this operator is separated in spherical coordinates, the Legendre differential equation arises [39, 40]. The solutions of this equation are Legendre functions of which associated Legendre polynomials are an important subset, denoted by $P_\ell^m(\mu)$. For $m = 0$, $P_\ell^m(\mu) = P_\ell(\mu)$ are called Legendre polynomials. We note that $P_0 = 1$ and $P_1 = \mu$.

Since the radiance is a function of the unit vector $\hat{\boldsymbol{s}}$, it may be practical to expand the radiance in terms of spherical harmonics $Y_{\ell m}(\mu, \varphi) = A_{\ell m} P_\ell^m(\mu) \exp(im\varphi)$ which form a complete orthonormal set of functions on the unit sphere ($A_{\ell m}$ are constants involving ℓ and m). In tissue optics, symmetries occur, such as a tissue irradiated by a collimated light beam with circular cross section, or the fact that the

phase function $p(\hat{\boldsymbol{s}}, \hat{\boldsymbol{s}}')$ depends only on $\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{s}}'$. Expansion of $L(\boldsymbol{r}, \hat{\boldsymbol{s}})$ and $p(\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{s}}')$ in terms of Legendre polynomials therefore seems a natural choice. Legendre polynomials form a complete orthogonal (not orthonormal) set of functions. The orthogonality is expressed as

$$\int_{-1}^{+1} P_i(\mu) P_j(\mu) d\mu = [2/(2i+1)]\delta_{ij} \tag{6.130}$$

where δ_{ij} is the Kronecker delta. We assume that a function $f(\mu)$ can be expressed as a series in Legendre polynomials

$$f(\mu) = \sum_{i=0}^{\infty} [(2i+1)/2] b_i P_i(\mu) \tag{6.131}$$

The phase function, for example, can then be written as

$$p(\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{s}}') = (1/2\pi) \sum_{i=0}^{\infty} [(2i+1)/2] b_i P_i(\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{s}}') \tag{6.132}$$

where the factor 2π is a normalization factor for the variable φ. To obtain the coefficients b_i in Eq.(6.131), both sides are multiplied by $P_j(\mu)$ and integrated over μ:

$$\int_{-1}^{+1} f(\mu) P_j(\mu) d\mu = \sum_{i=0}^{\infty} \frac{2i+1}{2} \int_{-1}^{+1} P_i(\mu) P_j(\mu) b_i d\mu = \sum_{i=0}^{\infty} \delta_{ij} b_i = b_j \tag{6.133}$$

where Eq. (6.130) has been used. If Eq. (6.133) is correct, substitution of b_j into Eq. (6.131) should yield an identity:

$$f(\mu) \equiv \int_{-1}^{+1} [\sum_{i=0}^{\infty} ((2i+1)/2) P_i(\mu) P_i(\mu')] f(\mu')\, d\mu' \tag{6.134}$$

The expression between square brackets must therefore be a representation of the delta function $\delta(\mu - \mu')$. This is called the completeness relation for Legendre polynomials. It is finally useful to mention two recurrence formulas for Legendre polynomials:

$$dP_{n+1}(\mu)/d\mu - dP_{n-1}(\mu)/d\mu = (2n+1) P_n(\mu) \tag{6.135}$$

$$(2n+1)\, \mu P_n(\mu) = (n+1) P_{n+1} + n P_{n-1}(\mu) \tag{6.136}$$

For further details about spherical harmonics etc. the reader is referred to one of the many books on mathematics for physics [39, 40].

References

1. Chandrasekhar S. *Radiative transfer.* Dover, New York (1960).
2. Svaasand LO and Gomer CJ. Optics of tissue. In: GJ Müller and DH Sliney (eds) *Dosimetry of laser radiation in medicine and biology*, Vol. IS5. SPIE Optical Engineering, Bellingham, Washington, pp. 114–132 (1989).
3. Mourant JR, Bocklage TJ, Powers TM, Greene HM, Bullock KL, Marr-Lyon LR, Dorin MH, Waxman AG, Zsemlye MM, and Smith HO. *In vivo* light scattering measurements for detection of precancerous conditions of the cervix. *Gynecol. Oncol*., 105:439–445 (2007).
4. Keijzer M, Star WM, and Storchi, PRM. Optical diffusion in layered media. *Appl. Opt.*, 27:1820–1824 (1988).
5. Pogue BW, Davis SC, Song X, Brooksby BA, Dehghani H, and Paulsen KD. Image analysis methods for diffuse optical tomography. *J. Biomed. Opt.*, 11:33001 (2007).
6. Gibson AP, Hebden JC, and Arridge SR. Recent advances in diffuse optical tomography. *Phys. Med. Biol*., 50:R1–R43 (2005).
7. Glasstone S and Edlund MC. *The elements of nuclear reactor theory.* Van Nostrand, New York (1952).
8. Shettle EP and Weinman JA. The transfer of solar irradiance through inhomogeneous turbid atmospheres evaluated by Eddington's approximation. *J. Atmos. Sci.*, 27: 1048–1055 (1970).
9. Hulst HC van de. *Multiple light scattering. Tables, formulas and applications,* Vol. I and II. Academic, New York (1980).
10. Davison B. *Neutron transport theory.* Clarendon, Oxford (1957).
11. Case KM and Zweifel PF. *Linear transport theory*, Addison-Wesley, Reading, MA (1967).
12. Star WM. Diffusion theory of light transport. In: AJ Welch and MJC Van Gemert (eds) *Optical-thermal response of laser-irradiated tissue*, Plenum, New York and London, 1st Edition, pp. 131–206 (1995).
13. Furutsu K. Diffusion equation derived from space-time transport equation. *Appl. Opt*. 70: 360–366 (1980).
14. Hulst HC van de. *Light scattering by small particles.* Dover, New York (1981).
15. Jacques SL, Alter CA, and Prahl SA. Angular dependence of He-Ne light scattering by human dermis. *Lasers Life Sci.*, 1:309–333 (1987).
16. Star WM. Comparing the P3-approximation with diffusion theory and with Monte Carlo calculations of light propagation in a slab geometry. In: GJ Müller and DH Sliney (eds) *Dosimetry of laser radiation in medicine and biology*, Vol. IS5. SPIE Optical Engineering, Bellingham, Washington, pp. 146–154 (1989).
17. Joseph JH, Wiscombe WJ, and Weinman JH. The delta-Eddington approximation for radiative flux transfer. *J. Atmos. Sci.,* 33:2452–2459 (1976).
18. Meador WE and Weaver WR. Diffusion approximation for large absorption in radiative transfer. *Appl. Opt*., 18:1204–1208 (1979).
19. Ishimaru A. *Wave propagation and scattering in random media, Vol. 1: Single scattering and transport theory.* Academic, New York (1978).
20. Bolin FP, Preuss LE, Taylor RC, and Ference RJ. Refractive index of some mammalian tissues using a fiber optic cladding method. *Appl. Opt.,* 28:2297–2303 (1989).
21. Star WM and Marijnissen JPA. Calculating the response of isotropic light dosimetry probes as a function of the tissue refractive index. *Appl. Opt*., 28:2288–2291 (1989).
22. Pickering JW, Moes CJM, Sterenborg HJCM, Prahl SA, and van Gemert MJC. Two integrating spheres with an intervening scattering sample. *J. Opt. Soc. Am.*, A9:621–631 (1992).

23. Pickering JW, Prahl SA, van Wieringen N, Beek JF, Sterenborg HJCM, and van Gemert MJC. Two integrating spheres with an intervening scattering sample. *J. Opt. Soc. Am.*, A9:621–631 (1992).
24. Kubelka P. New contributions to the optics of intensely light-scattering materials. Part I. *J. Opt. Soc. Am.,* 38:448–457 (1948).
25. Fukshansky-Kazarinova N, Lork W, Schafer E, and Fukshansky L. Photon flux gradients in layered turbid media: Application to biological tissues. *Appl. Opt.*, 25:780–788 (1986).
26. Abramowitz M and Stegun I. *Handbook of mathematical functions.* Dover, New York (1970).
27. Marijnissen JPA, Jansen H, and Star WM. Treatment system for whole bladder wall photodynamic therapy with in vivo monitoring and control of light dose rate and dose. *J. Urol.*, 142:1351–1355 (1989).
28. Star WM. The relationship between integrating sphere and diffusion theory calculations of fluence rate at the wall of a spherical cavity. *Phys. Med. Biol.*, 40:1–8 (1995).
29. Van Staveren HJ, Beek JF, Ramaekers JWH, Keijzer M, and Star WM. Integrating sphere effect in whole bladder wall photodynamic therapy: I. 532 versus 630 nm optical irradiation. *Phys Med. Biol.*, 39:947–959 (1994).
30. Reynolds L, Johnson C, and Ishimaru A. Diffuse reflectance from a finite blood medium: Applications to the modeling of fiber optic catheters. *Appl. Opt.,* 15:2059–2076 (1976).
31. Groenhuis RAJ, Ferwerda HA, and Ten Bosch JJ. Scattering and absorption of turbid materials determined from reflection measurements. *Appl. Opt.*, 22:2456–2462 (1983).
32. Farrell TJ, Patterson MS, and Wilson BC. A diffusion theory model of spatially resolved, steady state diffuse reflectance for the noninvasive determination of tissue optical properties. *Med. Phys.*, 19:879–888 (1992).
33. Allen V and McKenzie AL. The modified diffusion dipole model. *Phys. Med. Biol.*, 36:1621–1638 (1991).
34. Fretterd RJ and Longini RL. Diffusion dipole source. *J. Opt. Soc. Am.*, 63:336–337 (1973).
35. Haskell RC, Svaasand LO, Tsay T-T, Feng T-C, McAdams MS, and Tromberg BJ. Boundary conditions for the diffusion equation in radiative transfer. *J. Opt. Soc. Am. A*, 11:2727–2741 (1994).
36. Kienle A and Patterson MS. Improved solutions of the steady-state and the time-resolved diffusion equations for reflectance from a semi-infinite turbid medium. *J. Opt. Soc. Am. A*, 14:246–254 (1997).
37. Hull EL and Foster TH. Steady-state reflectance spectroscopy in the P_3 approximation. *J. Opt. Soc. Am. A*, 18:584–599 (2001).
38. van Veen RLP, Sterenborg HJCM, Marinelli AWKS, and Menke-Pluymers M. Intraoperatively assessed optical properties of malignant and healthy breast tissue used to determine the optimum wavelength of contrast for optical mammography. *J. Biomed. Opt.*, 9:1129–1136 (2004).
39. Margenau H and Murphy GM. *The mathematics of physics and chemistry.* Van Nostrand, Princeton, NJ (1956).
40. Wylie CR and Barrett LC. *Advanced engineering mathematics.* McGraw-Hill, Singapore (1985).

Chapter 7
From Electrodynamics to Monte Carlo Simulations

Jaro Rička and Martin Frenz

7.1 Basic Scattering Concepts, Born Approximation

Biological tissues are highly heterogenous condensed media, whose optical properties are quite difficult to model. In fact, it is much easier to say what tissues *are not*: they are not dilute suspensions of scattering sphere-like particles. This statement sounds much like a truism, which it is. However, the mathematics employed for the description of light propagation in tissues are a result of propagation theory associated with incoherent stellar light in interstellar atmospheres [1], which certainly *are* such particle suspensions, even though extremely dilute. Surprisingly, these techniques are highly successful in propagating coherent laser light in tissues. The suggestive power of this success strongly influences the way we think about tissue optics: when thinking about scattering in tissues, one inevitably has Mie particles in mind. New approaches to tissue optics are now emerging. Instead of picturing the tissue as a cloud of independent particles, one seeks to characterize its random dense structure in terms of density correlation functions or spatial power spectra [2]. These developments appear to have drawn their inspiration from two sources. One is the wave propagation in turbulent atmosphere as discussed, e.g., in [3]; and the second inspiration are concepts from statistical physics that were originally developed for small angle X-ray and neutron scattering [2, 4–6] in soft condensed matter. Since soft condensed matter seems to be much closer to biological tissue than turbulent atmosphere, we shall pursue the second path. In small angle scattering one usually employs two assumptions: (i) the interaction of the radiation with the matter is weak, so that the scattering can be treated in the first Born approximation; multiple scattering is usually negligible; (ii) because of the inherently small scattering angle (typically 0.1–5°), polarization can be neglected. Fortunately, we can retain the first part of assumption (i), because the refractive index differences of the tissue constituents are relatively small. They lay between the extremes water and fat, i.e. in the range 1.33–1.45 [7]. However, multiple scattering is what we are finally

J. Rička (✉)
Institute of Applied Physics, Sidlerstr. 5, 3012 Bern, Switzerland
e-mail: jaroslav.ricka@iap.unibe.ch

A.J. Welch, M.J.C. van Gemert (eds.), *Optical-Thermal Response of Laser-Irradiated Tissue*, 2nd ed., DOI 10.1007/978-90-481-8831-4_7,

interested in. Moreover, the present task is to include in the multiple scattering the effect of polarization. Thus, in order to make use of the new approaches to tissue optics, we must learn to combine the concepts involving the correlation functions with the polarization concepts, including the Jones matrix of the scattering process.

7.1.1 Scattering Geometry

We begin with the most simple and well known case of scattering from a single dielectric particle, for example a single molecule. The particle scatters laser light: The lasers oscillating electric field $\mathbf{E}^I(\mathbf{r})\exp(-i\omega t) \equiv \left|E^I\right\rangle$ excites in the particle an oscillating current density $\mathbf{j}(\mathbf{r},t) = \mathbf{j}(\mathbf{r})\exp(-i\omega t)$, which in turn radiates electromagnetic energy through a quasi-spherical wave $\mathbf{E}^s(\mathbf{r})\exp(-i\omega t) \equiv |E^s\rangle$. (Dirac's bra-ket-notation is described in Chapter 4, Section 4.3.3.) From the scattered field that radiates in all directions we select for observation a certain component $|E^{os}\rangle$. The selection is done with a beam filter, as discussed in Chapter 4, Section 4.5 and illustrated in Fig. 4.8. Scattering by isolated particles is well known from standard texts, but we use the opportunity to prepare tools for treating more difficult aspects of light scattering and propagation in tissues. What we are aiming at is a scattering transition matrix, i.e. a Jones matrix $\mathbf{T}$ that captures the transition of a photon from the laser beam into the receiver beam, including not only polarization and amplitude of the process but also the phase. The matrix $\mathbf{T}$ operates on the Jones vectors of complex amplitudes (recall Section 4.5.2):

$$\begin{pmatrix} \mathcal{E}_x^O \\ \mathcal{E}_y^O \end{pmatrix} = \mathbf{T}_{\mathbf{lo}}^{\mathbf{L}} \begin{pmatrix} \mathcal{E}_x^L \\ \mathcal{E}_y^L \end{pmatrix} \tag{7.1}$$

The superscripts refer to the fact that the Jones vectors are defined in two different coordinate systems, L and O. Such a transition matrix is needed in order to design a model for polarized Monte Carlo simulation.

The geometry of the experiment is sketched in Fig. 7.1. The tripod $\hat{\mathbf{x}}_\mathbf{L}$, $\hat{\mathbf{y}}_\mathbf{L}$, $\hat{\mathbf{z}}_\mathbf{L}$ defines the laboratory coordinate system with the origin at $\mathbf{r}_L$. The left gray horizontal plane of Fig. 7.1 symbolizes the top surface of an optical table defined by the unit vectors $\hat{\mathbf{y}}_\mathbf{L}$, $\hat{\mathbf{z}}_\mathbf{L}$. The optical axis of the illuminating laser beam $\hat{\mathbf{s}}_\mathbf{i}$ coincides with $\hat{\mathbf{z}}_\mathbf{L}$. The $\hat{\mathbf{x}}_\mathbf{L}$-axis is perpendicular to the optical table onto which the elements needed to manipulate the polarization state of the laser are mounted. The molecule is expected to scatter in all directions; therefore, the receiver and the elements for polarization analysis are mounted on a rotational stage at observation angles ϕ and θ. The laser beam and the receiver beam are aligned so that they intersect in a common focus at $\mathbf{r}_f$. The propagation vectors $\hat{\mathbf{s}}_\mathbf{i}$ and $\hat{\mathbf{s}}_\mathbf{o}$ define the scattering plane π, as discussed in Section 4.3.5. The molecule under consideration is assumed fixed close to the focus $\mathbf{r}_f$, in the region where the two beams overlap. Details of this region are shown in Fig. 7.2. Since positioning of single molecules is rather difficult, we allow for a small misalignment so that the position of the molecule is $\mathbf{r}_m = \mathbf{r}_f + \mathbf{m}$. Note that there are three coordinate systems involved. In addition to the lab system there

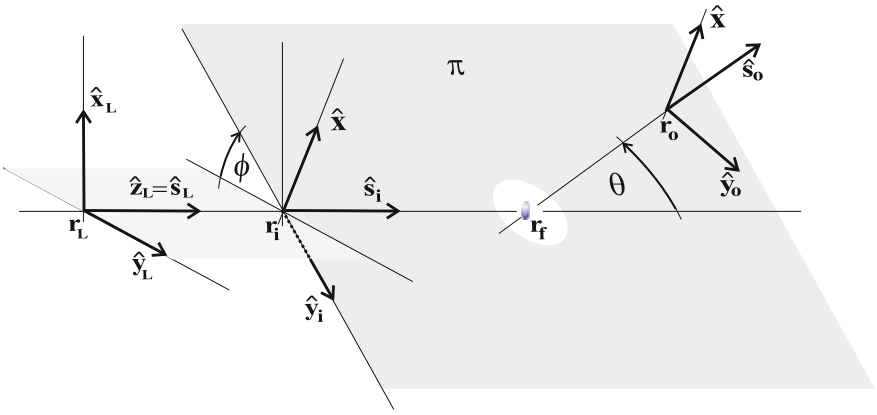


Fig. 7.1 Scattering geometry. Details of the region around beam focus $\mathbf{r}_f$ are shown in Fig. 7.2

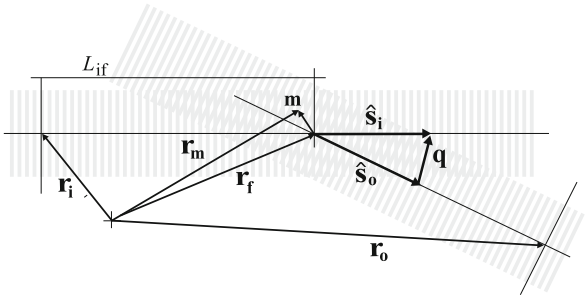


Fig. 7.2 Scattering geometry, reference planes, scattering vector **q**

are the two coordinate systems: the I-system $\hat{\mathbf{x}}, \hat{\mathbf{y}}_\mathbf{i}, \hat{\mathbf{s}}_\mathbf{i}$ with the origin at $\mathbf{r}_i$ and the O-system $\hat{\mathbf{x}}, \hat{\mathbf{y}}_\mathbf{o}, \hat{\mathbf{s}}_\mathbf{o}$ with the origin at $\mathbf{r}_o$.

7.1.2 Transition Matrix

Now we put the beam filter into operation in order to select from the scattered field $|E^s\rangle$ an observed partial beam $|E^{os}\rangle$ that propagates into the receiver. The selection is conveniently achieved using the projection technique discussed in Section 4.5. In the absence of a polarizer in the detection path, the beam filter accepts two modes with the same beam profile but with orthogonal polarizations. Thus, the selected and observed component of the scattered field is the superposition of two orthogonal components, i.e. $|E^{os}\rangle = \left|E_x^{os}\right\rangle + \left|E_y^{os}\right\rangle$, where

$$\begin{aligned} \left|E_x^{os}\right\rangle &= \left|o_x\right\rangle \mathcal{E}_x^o = \left|o_x\right\rangle\left\langle o_x \,\middle|\, E^s\right\rangle \\ \left|E_y^{os}\right\rangle &= \left|o_y\right\rangle \mathcal{E}_y^o = \left|o_y\right\rangle\left\langle o_y \,\middle|\, E^s\right\rangle \end{aligned} \tag{7.2}$$

Once $\mathcal{E}_x$ and $\mathcal{E}_y$ are known, then the actual field $\mathbf{E}^{os}(\mathbf{r})$ that is selected by the receiver from the scattered field can be obtained in any position $\mathbf{r}$ simply by combining the amplitudes with the virtual observation modes $|o_x\rangle$ and $|o_y\rangle$. Thus, the Jones vector consisting of the two amplitudes is a complete representation of the selected state:

$$\begin{pmatrix} \mathcal{E}_x^O \\ \mathcal{E}_y^O \end{pmatrix} = \begin{pmatrix} \langle o_x \mid E^s \rangle \\ \langle o_y \mid E^s \rangle \end{pmatrix} \tag{7.3}$$

Since the source of the scattered field, the current density distribution, $\mathbf{j}(\mathbf{r}, t)$, is well localized within a small particle, we calculate the scalar products $\langle o_x \mid E^s \rangle$ and $\langle o_y \mid E^s \rangle$ using Eq. (4.122), reproduced here for convenience:

$$\langle o \mid E^s \rangle = -\frac{1}{4} \int_V \hat{\mathbf{E}}^{o*} \cdot \mathbf{j} \, d^3 r \tag{7.4}$$

The integration extends over the volume V of the particle. The only assumption that we make about the distribution $\mathbf{j}(\mathbf{r}, t)$ is that it results from *linear dielectric response* of the particle material to the impinging electric field (we neglect magnetic response). Recall from macroscopic electrodynamics that $\mathbf{j}(\mathbf{r}) = \partial \mathbf{P}(\mathbf{r})/\partial t$ and $\mathbf{P}(\mathbf{r}) = \epsilon_o \boldsymbol{\chi}(\mathbf{r}) \mathbf{E}_{int}(\mathbf{r})$, where $\boldsymbol{\chi}(\mathbf{r})$ is the susceptibility tensor and $\mathbf{E}_{int}(\mathbf{r})$ is the local internal field, as it would be measured within the matter at the position $\mathbf{r}$. Here we employ a generalized version of this relation, which explicitly considers *non-local response*:

$$\mathbf{j}(\mathbf{r}, t) = -i\omega \int_V \epsilon_o \boldsymbol{\chi}(\mathbf{r}, \mathbf{r}', \omega) \cdot \mathbf{E}^l(\mathbf{r}', t) \, d^3 r' \tag{7.5}$$

Here $\boldsymbol{\chi}(\mathbf{r}, \mathbf{r}', \omega)$ is a tensor density, a non-local version of the susceptibility tensor $\boldsymbol{\chi}$. Note that $\mathbf{E}^l(\mathbf{r}', t)$ is the *impinging field* as it would exist in the medium in the absence of the scattering particle, and not the *internal field* $\mathbf{E}_{int}(\mathbf{r}, t)$, that would be measured within the dielectric particle.[1] Note also that Eq. (7.5) is exact; no approximations are yet involved. The purpose of this purely formal expression is to make the vectorial linear relationship between the current density $\mathbf{j}$ and the impinging field $\mathbf{E}^l$ explicit. Thus, the scalar product has the general form

$$\left\langle o \,\middle|\, E^s \right\rangle = \frac{i\omega}{4} \int\int_V \hat{\mathbf{E}}^{o*}(\mathbf{r}) \cdot \epsilon_o \boldsymbol{\chi}(\mathbf{r}, \mathbf{r}') \cdot \mathbf{E}^l(\mathbf{r}') \, d^3 r' \, d^3 r = \left\langle o \, |D| E^l \right\rangle \tag{7.6}$$

On the right hand side we introduce a convenient abbreviation. Recall from Chapter 4, Section 4.5.2, that the impinging laser field can be expressed as a linear superposition of two base states with two complex amplitudes $\mathcal{E}$. For compatibility reasons we choose XY-base states in the I-system and write:

[1] The *internal field* $\mathbf{E}_{int}(\mathbf{r}, t)$ is related to $\mathbf{E}^l(\mathbf{r}', t)$ through the relation $\boldsymbol{\chi}(\mathbf{r}) \cdot \mathbf{E}_{int}(\mathbf{r}, t) = \int_V \boldsymbol{\chi}(\mathbf{r}, \mathbf{r}', \omega) \cdot \mathbf{E}^l(\mathbf{r}', t) \, d^3 r'$. This is essentially a short hand version of the Ewald Oseen extinction theorem (see e.g. [8]). Internal fields are the subject of classical treatments of scattering by particles (see, e.g. Section 4.3 in [9]).

$$\left|E^l\right\rangle \equiv \mathbf{E^l}\left(\mathbf{r}\right) = \left|l_x^l\right\rangle \mathcal{E}_{lx}^I + \left|l_y^l\right\rangle \mathcal{E}_{ly}^I \tag{7.7}$$

Applying this decomposition to the scalar products $\langle o \mid E^s \rangle$ (which are linear in $\left|E^l\right\rangle$), we may re-write Eq. (7.3) as

$$\begin{aligned}\begin{pmatrix}\mathcal{E}_x^O\\ \mathcal{E}_y^O\end{pmatrix} &= \begin{pmatrix}\langle o_x|D \mid l_x^l\rangle \mathcal{E}_{lx}^I + \left\langle o_x|D \,\middle|\, l_y^l\right\rangle \mathcal{E}_{ly}^I\\ \langle o_y|D \mid l_x^l\rangle \mathcal{E}_{lx}^I + \left\langle o_y|D \,\middle|\, l_y^l\right\rangle \mathcal{E}_{ly}^I\end{pmatrix}\\ &= \begin{pmatrix}\langle o_x|D \mid l_x^l\rangle & \left\langle o_x|D \,\middle|\, l_y^l\right\rangle\\ \langle o_y|D \mid l_x^l\rangle & \left\langle o_y|D \,\middle|\, l_y^l\right\rangle\end{pmatrix}\begin{pmatrix}\mathcal{E}_{lx}^I\\ \mathcal{E}_{ly}^I\end{pmatrix}\end{aligned} \tag{7.8}$$

The right hand side of Eq. (7.8) defines the desired transition matrix in the I-system. The transition from the L-system to the I-system is achieved by a simple 2D rotation $\mathbf{R}_2\left(\phi\right)$

$$\mathbf{T_{lo}^L} = \mathbf{T_{lo}^I}\mathbf{R}_2\left(\phi\right) = \begin{pmatrix}\langle o_x|D \mid l_x^l\rangle & \left\langle o_x|D \,\middle|\, l_y^l\right\rangle\\ \langle o_y|D \mid l_x^l\rangle & \left\langle o_y|D \,\middle|\, l_y^l\right\rangle\end{pmatrix}\begin{pmatrix}\cos\phi & \sin\phi\\ -\sin\left(\phi\right) & \cos\left(\phi\right)\end{pmatrix} \tag{7.9}$$

Equation (7.9) is exact and quite general. It applies to any kind of linear scattering system, illuminated by any kind of laser field and observed with any kind of single mode receiver. The real work begins with the calculation of the matrix elements $\langle o|D \mid l\rangle$.

7.1.3 Matrix Elements

The first step is to specify the impinging laser field and the observation modes. The laser beam is a quasi-monochromatic quasi-plane wave, as introduced in Chapter 4, Section 4.2.3.2. The laser power is P_l. The beam profile close to focus is $X_l(\mathbf{r})$, the propagation is directed along $\hat{\mathbf{s}}_\mathbf{i}$ and the polarization is specified by the complex polarization vector $\hat{\mathbf{e}}_\mathbf{l}$ (with components $\left\{e_{xl}^L, e_{yl}^L, 0\right\}$ in the L-system). We also specify the phase $\Phi(i)$ in a reference plane at $\mathbf{r}_i$ and $t = 0$:

$$\mathbf{E^l}\left(\mathbf{r},\, t\right) \equiv \left|E^l\right\rangle = \sqrt{P_l}\sqrt{2/c\epsilon_o}\, X_l\left(\mathbf{r}\right) e^{ik\hat{\mathbf{s}}_\mathbf{i}\cdot\left(\mathbf{r}-\mathbf{r}_i\right)} e^{i\Phi(i)} e^{-i\omega t}\hat{\mathbf{e}}_\mathbf{l} \tag{7.10}$$

where k is the wave number.

In the matrix elements we express the laser beam in the XY-polarization bases in the I-system: $\left|E^l\right\rangle \equiv \mathbf{E^l}\left(\mathbf{r}\right) = \left|l_x^l\right\rangle \mathcal{E}_{lx}^I + \left|l_y^l\right\rangle \mathcal{E}_{ly}^I$. That base modes are

$$\begin{aligned}\left|l_x^l\right\rangle &\equiv \hat{\mathbf{E}}_\mathbf{x}^\mathbf{l}\left(\mathbf{r}\right) = \ldots\sqrt{2/c\epsilon_o}\, X_l\left(\mathbf{r}\right) e^{ik\hat{\mathbf{s}}_\mathbf{i}\cdot\left(\mathbf{r}-\mathbf{r}_i\right)}\hat{\mathbf{x}}\\ \left|l_y^l\right\rangle &\equiv \hat{\mathbf{E}}_\mathbf{y}^\mathbf{l}\left(\mathbf{r}\right) = \ldots\sqrt{2/c\epsilon_o} X_l\left(\mathbf{r}\right) e^{ik\hat{\mathbf{s}}_\mathbf{i}\cdot\left(\mathbf{r}-\mathbf{r}_i\right)}\hat{\mathbf{y}}_\mathbf{I}\end{aligned} \tag{7.11}$$

Recall from Chapter 4, Section 4.5.2 that the base modes are a kind of unit vectors, normalized such that $\langle l \mid l \rangle = 1$. Therefore the factor $\sqrt{P}$ is missing. The phase of a base mode is irrelevant, but for convenience we include the reference point $\mathbf{r}_i$. The corresponding base modes in the O-system (observation modes) are

$$\begin{aligned} |o_x\rangle &\equiv \hat{\mathbf{E}}^{\mathbf{o}}_{\mathbf{x}}(\mathbf{r}) = \ldots \sqrt{2/c\epsilon_o}\, X_o(\mathbf{r})\, e^{ik\hat{\mathbf{s}}_\mathbf{o}\cdot(\mathbf{r}-\mathbf{r}_f)}\hat{\mathbf{x}} \\ |o_y\rangle &\equiv \hat{\mathbf{E}}^{\mathbf{o}}_{\mathbf{y}}(\mathbf{r}) = \ldots \sqrt{2/c\epsilon_o} X_o(\mathbf{r})\, e^{ik\hat{\mathbf{s}}_\mathbf{o}\cdot(\mathbf{r}-\mathbf{r}_f)}\hat{\mathbf{y}}_\mathbf{o} \end{aligned} \tag{7.12}$$

The Cartesian unit vectors $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}_\mathbf{I}$, and $\hat{\mathbf{y}}_\mathbf{o}$ are defined in Fig. 7.1. We insert these base modes in place of the fields in Eq. (7.6) and calculate the elements T_{ol} of the $\mathbf{T}^{\mathbf{I}}_{\mathbf{lo}}$-matrix. Denoting with $\hat{\mathbf{e}}_\mathbf{l}$ and $\hat{\mathbf{e}}_\mathbf{o}$ the involved pairs of polarization vectors, we write

$$T_{ol} = i\frac{k}{2}\frac{1}{\epsilon_o}\iint_V X_o(\mathbf{r})\, X_l(\mathbf{r}')\, e^{ik\hat{\mathbf{s}}_\mathbf{i}\cdot(\mathbf{r}'-\mathbf{r}_i)} e^{ik\hat{\mathbf{s}}_\mathbf{o}\cdot(\mathbf{r}-\mathbf{r}_f)} \left[\hat{\mathbf{e}}_\mathbf{o}\cdot\epsilon_o\boldsymbol{\chi}(\mathbf{r}\cdot\mathbf{r}')\cdot\hat{\mathbf{e}}_\mathbf{l}\right] d^3r'\, d^3r \tag{7.13}$$

Now comes a first approximation. We assume that the size a of the scattering particle is much smaller than the beam radii a_l and a_o so that $X_o(\mathbf{r})X_1(\mathbf{r}')$ are approximately constant across the scattering particle. Thus, we take the profiles out of the integration:

$$T_{ol} = i\frac{k}{2}\frac{1}{\epsilon_o} X_o(\mathbf{r}_m)\, X_l(\mathbf{r}_m)\iint_V e^{ik\hat{\mathbf{s}}_\mathbf{i}\cdot(\mathbf{r}'-\mathbf{r}_i)} e^{-ik\hat{\mathbf{s}}_\mathbf{o}\cdot(\mathbf{r}-\mathbf{r}_f)} \left[\hat{\mathbf{e}}_\mathbf{o}\cdot\epsilon_o\boldsymbol{\chi}(\mathbf{r}\cdot\mathbf{r}')\cdot\hat{\mathbf{e}}_\mathbf{l}\right] d^3r'\, d^3r \tag{7.14}$$

Now we transform the integration variables from $\mathbf{r}$ and $\mathbf{r}'$ to $\mathbf{u} = \mathbf{r} - \mathbf{r}_m$ and $\mathbf{u}' = \mathbf{r}' - \mathbf{r}_m$ and take the part of the phasor that does not depend on $\mathbf{u}$ or $\mathbf{u}'$ out of the integration. The remaining integral defines an auxiliary quantity $\mathbf{A}_\mathbf{m}$ that we call *multipole polarizability tensor*:

$$\mathbf{A}_\mathbf{m} = \iint_V e^{ik\hat{\mathbf{s}}_\mathbf{i}\cdot\mathbf{u}'} e^{-ik\hat{\mathbf{s}}_\mathbf{o}\cdot\mathbf{u}}\epsilon_o\boldsymbol{\chi}(\mathbf{u},\mathbf{u}')\, d^3u'\, d^3u \tag{7.15}$$

This $\mathbf{A}_\mathbf{m}$ is a generalized version of the dipole polarizability tensor introduced in Eq. (4.41). We also introduce the dimensionless normalized version $\mathbf{A} = \mathbf{A}_\mathrm{m}/\alpha$, where $\alpha = \mathrm{Trace}(\mathbf{A}_\mathbf{m})/3$ is a generalized polarizability. Finally, we re-arrange the phasor $\exp\left[ik\,\hat{\mathbf{s}}_\mathbf{i}\cdot(\mathbf{r}_m - \mathbf{r}_i)\right]\exp\left[ik\,\hat{\mathbf{s}}_\mathbf{o}\cdot(\mathbf{r}_m - \mathbf{r}_f)\right]$, so that the matrix element can be written as

$$T_{ol} = X_o(\mathbf{r}_m)\, X_l(\mathbf{r}_m)\, e^{i\mathbf{q}\cdot\mathbf{r}_\mathbf{m}} e^{ikL_{if}}\frac{ik}{2}\frac{\alpha}{\epsilon_o}\left[\hat{\mathbf{e}}_\mathbf{o}\cdot\mathbf{A}\cdot\hat{\mathbf{e}}_\mathbf{l}\right] \tag{7.16}$$

In kL_{if} we included the phase contribution that does not depend on the position $\mathbf{r}_m$ of the scatterer relative to the focus $\mathbf{r}_f$. As shown in Fig. 7.2, the quantity

kL_{if} represent the optical path length from $\mathbf{r}_i$ to $\mathbf{r}_f$. One easily finds that $kL_{if} = k\hat{\mathbf{s}}_\mathbf{i} \cdot (\mathbf{r}_f - \mathbf{r}_i)$. Note that we regard the medium as isotropic: both polarizations propagate with the same phase speed and in the same beam. This is a quite serious simplifying approximation, considering the fact that biological tissues can be birefringent. The second phasor $\exp(i\mathbf{q} \cdot \mathbf{r}_m)$ measures the particle position. The vector $\mathbf{q}$ is the so called *scattering vector* that is defined as

$$\mathbf{q} = \mathbf{k}_I - \mathbf{k}_o = k\left(\hat{\mathbf{s}}_\mathbf{I} - \hat{\mathbf{s}}_\mathbf{o}\right) \tag{7.17}$$

The magnitude of the scattering vector decreases with the scattering angle θ, since $q^2 = 2k^2[1 - \cos(\theta)]$. Note in Fig. 7.2 the fringes in the illuminated and observed region (IOR); their wave fronts are perpendicular to $\mathbf{q}$ and their wavelength is $\lambda_q = 2\pi/q$. Note also that $\hbar\mathrm{q} = \Delta\hbar\mathrm{k}$ is the momentum transferred from the photon to the scattering molecule. By using in the filter $|o\rangle$ the same wave number $k = \omega/c$ as in the input state $|E^I\rangle$, we are restricting the discussion to elastic or quasi-elastic scattering: no change of photon energy $\hbar\omega$ in the scattering process, only the direction of the momentum $\hbar\mathbf{k}$ changes. Note that in $\exp(i\mathbf{q} \cdot \mathbf{r}_m)$ the position $\mathbf{r}_m$ is coded much more accurately than in the smooth profiles $X_o(\mathbf{r}_m)$, $X_I(\mathbf{r}_m)$. Therefore, the scattering vector $\mathbf{q}$ plays a central role in the analysis of the structure of the matter by scattering techniques.

7.1.4 Amplitude Scattering Matrix

Equation (7.16) is quite general, the only assumption being that the scattering particle is small, so that the illumination and observation modes can be viewed as plane waves. Unfortunately, however, direct calculation of $\mathbf{A_m}$ is a formidable task, too difficult for the present aims and present skills. Therefore we cast Eq. (7.8) in a canonical form to exploit the tremendous amount of theoretical and experimental work that has been done on arbitrary scattering particles by various techniques. Upon inserting the matrix elements from Eq. (7.16) into Eq. (7.8) we obtain after some re-arrangements:

$$\begin{pmatrix} \mathcal{E}_x^O \\ \mathcal{E}_y^O \end{pmatrix} = \mathcal{B}(\mathbf{r}_m) \left\{ \frac{ik^3}{4\pi} \frac{\alpha}{\epsilon_o} \begin{pmatrix} \hat{\mathbf{x}} \cdot \mathbf{A} \cdot \hat{\mathbf{x}} & \hat{\mathbf{x}} \cdot \mathbf{A} \cdot \hat{\mathbf{y}}_\mathbf{I} \\ \hat{\mathbf{y}}_O \cdot \mathbf{A} \cdot \hat{\mathbf{x}} & \hat{\mathbf{y}}_O \cdot \mathbf{A} \cdot \hat{\mathbf{y}}_\mathbf{I} \end{pmatrix} \right\} \begin{pmatrix} \mathcal{E}_{lx}^I \\ \mathcal{E}_{ly}^I \end{pmatrix} \tag{7.18}$$

where

$$\mathcal{B}(\mathbf{r}_m) = X_I(\mathbf{r}_m) X_o(\mathbf{r}_m) \frac{2\pi}{k^2} e^{i\,\mathbf{q}\cdot\mathbf{m}} e^{ikL_{if}} \tag{7.19}$$

The quantity in curly brackets { } is the classical *amplitude scattering matrix* **S**, see, e.g., Section 3.2 in [9]. Keep in mind that an **S**-matrix is always defined in the I-system! The present notation is somewhat unusual. For convenience, we decomposed the S-matrix into the pre-factor that is proportional to the generalized

polarizability α, and the polarization matrix $\mathbf{\Pi}$ that concerns only the polarization transfer:

$$\mathbf{S} = \frac{ik^3}{4\pi}\frac{\alpha}{\varepsilon_o}\begin{pmatrix} \hat{\mathbf{x}}\cdot\mathbf{A}\cdot\hat{\mathbf{x}} & \hat{\mathbf{x}}\cdot\mathbf{A}\cdot\hat{\mathbf{y}}_\mathbf{I} \\ \hat{\mathbf{y}}_O\cdot\mathbf{A}\cdot\hat{\mathbf{x}} & \hat{\mathbf{y}}_O\cdot\mathbf{A}\cdot\hat{\mathbf{y}}_\mathbf{I} \end{pmatrix} \quad \mathbf{\Pi} = \begin{pmatrix} \hat{\mathbf{x}}\cdot\mathbf{A}\cdot\hat{\mathbf{x}} & \hat{\mathbf{x}}\cdot\mathbf{A}\cdot\hat{\mathbf{y}}_\mathbf{I} \\ \hat{\mathbf{y}}_O\cdot\mathbf{A}\cdot\hat{\mathbf{x}} & \hat{\mathbf{y}}_O\cdot\mathbf{A}\cdot\hat{\mathbf{y}}_\mathbf{I} \end{pmatrix} \tag{7.20}$$

The standard notation is

$$\mathbf{S} = \begin{pmatrix} S_{xx} & S_{xy} \\ S_{yx} & S_{yy} \end{pmatrix} = \begin{pmatrix} S1 & S4 \\ S3 & S2 \end{pmatrix} \tag{7.21}$$

The numbering of the elements S1-4 corresponds to the notation in [10]. Working in the *xy*-linear base, the four matrix elements concern the following polarization transfers:

$$\begin{aligned} &S1:\ \hat{\mathbf{e}}_{\mathbf{i}\perp} = \hat{\mathbf{x}} \rightarrow \hat{\mathbf{x}} = \hat{\mathbf{e}}_{\mathbf{o}\perp} && S4:\ \hat{\mathbf{e}}_{\mathbf{i}\parallel} = \hat{\mathbf{y}}_\mathbf{i} \rightarrow \hat{\mathbf{x}} = \hat{\mathbf{e}}_{\mathbf{o}\perp} \\ &S3:\ \hat{\mathbf{e}}_{\mathbf{i}\perp} = \hat{\mathbf{x}} \rightarrow \hat{\mathbf{y}}_\mathbf{o} = \hat{\mathbf{e}}_{\mathbf{o}\parallel} && S2:\ \hat{\mathbf{e}}_{\mathbf{i}\parallel} = \hat{\mathbf{y}}_\mathbf{i} \rightarrow \hat{\mathbf{y}}_\mathbf{o} = \hat{\mathbf{e}}_{\mathbf{o}\parallel} \end{aligned} \tag{7.22}$$

The canonical form $|E^o\rangle = \mathcal{B}(\mathbf{r}_m)\, S\, |E^{II}\rangle$ is applicable to the scattering of a quasi-plane wave by an arbitrary particle of arbitrary size a such that $a << a_I$. The dimensionless factor $\mathcal{B}(\mathbf{r}_m)$ corresponds to the factor $\exp(ikr)/kr$ of the traditional far-field formulation of the scattering problem. The far-field result can be easily derived using the present mode projection approach, one only has to replace the beam filter that consists of a lens and fiber or pinhole, (see Fig. 4.8) with a dipole antenna far from the scatterer.

7.1.5 Scattering Signal, Cross-Sections and Phase Function

To understand the significance of the S-matrix, we consider the measurable quantity, the received power P_o. Since we removed the polarizer from the detection channel, the received power is the sum of two contributions from the two orthogonal modes:

$$P_{os} = \langle E^{os}|E^{os}\rangle = |\mathcal{E}_x^O|^2 + |\mathcal{E}_y^O|^2,$$

where

$$\begin{aligned} |\mathcal{E}_x^O|^2 &= \left|T_{xx}\mathcal{E}_{lx}^I + T_{xy}\mathcal{E}_{ly}^I\right|^2 = |\mathcal{B}(\mathbf{r}_m)|^2\left|S_{xx}\mathcal{E}_{lx}^I + S_{xy}\mathcal{E}_{ly}^I\right|^2 \\ |\mathcal{E}_y^O|^2 &= \left|T_{yx}\mathcal{E}_{lx}^I + T_{yy}\mathcal{E}_{ly}^I\right|^2 = |\mathcal{B}(\mathbf{r}_m)|^2\left|S_{yx}\mathcal{E}_{lx}^I + S_{yy}\mathcal{E}_{ly}^I\right| \end{aligned} \tag{7.23}$$

In the concise Jones vector notation this reads

$$P_{os} = \langle E_l^I|\mathbf{T}_\mathbf{lo}^\dagger\mathbf{T}_\mathbf{lo}|E_l^I\rangle = |\mathcal{B}(\mathbf{r}_m)|^2\ \langle E_l^I|\mathbf{S}^\dagger\mathbf{S}|E_l^I\rangle \tag{7.24}$$

where † represents conjugate transpose (adjoint).

Most often we do not care about the absolute power received, but the ratio of input and output, i.e. in the *probability* $P_{os}\,(\theta,\,\phi)$ that a photon traveling in the laser beam is scattered into the observation beam aligned along $\hat{\mathbf{s}}_\mathbf{o} \equiv (\theta,\,\phi)$:

$$\mathcal{P}_{os}(\theta,\phi) = \frac{P_{os}}{P_l} = |\mathcal{B}(\mathbf{r}_m)|^2 \frac{\langle E_l^I|\mathbf{S}^\dagger\mathbf{S}|E_l^I\rangle}{\langle E_l^I|E_l^I\rangle} = |\mathcal{B}(\mathbf{r}_m)|^2 \frac{\langle E_l^l|\mathbf{R}_2^{-1}(\phi)\mathbf{S}^\dagger\mathbf{S}|\mathbf{R}_2(\phi)E_l^l\rangle}{\langle E_l^l|E_l^l\rangle} \tag{7.25}$$

where, of course $\langle E^l|E^l\rangle = \langle E^I|E^I\rangle = P_l = |\mathcal{E}_{lx}^I|^2 + |\mathcal{E}_{ly}^I|^2$. This looks like a trivial re-arrangement, but one should realize that the complex amplitudes can be interpreted as $\mathcal{E}_{lx}^I = \sqrt{P_l}e^{i\Phi}e_{xl}^I$ and $\mathcal{E}_{ly}^I = \sqrt{P_l}e^{i\Phi}e_{yl}^I$. Here e_{xl}^I and e_{yl}^I are the components of a complex unit polarization vector in the I-system, such that $|e_{xl}^I|^2 + |e_{yl}^I|^2 = 1$. Thus, the fraction $\langle E_l^I|\mathbf{S}^\dagger\mathbf{S}|E_l^I\rangle/\langle E^l|E^l\rangle$ on the right hand side of Eq. (7.25) depends only on the polarization and not on the power and phase of the laser beam. In order to establish a link with the common radiometric language, we note that the pre-factor $|\mathcal{B}(\mathbf{r}_m)|^2$ (recall B from Eq. (7.19)) can be expressed as follows

$$|\mathcal{B}(\mathbf{r}_m)|^2 = X_l(\mathbf{r}_m)Y_o(\mathbf{r}_m)\Omega_o\frac{1}{k^2} \tag{7.26}$$

The quantities Ω_o and Y_o have been defined in Chapter 4, Section 4.2.3.2 in the context of Eq. (4.40): $\Omega_o = 4\pi/k^2a_o^2$ is the effective solid angle accepted by the receiver, a_o is the waist radius of the observation mode. The dimensionless profile $Y_o(\mathbf{r}_m) = \pi a_o^2 X_o^2(\mathbf{r}_m)$ quantifies the alignment of the receiver. One can set $Y(\mathbf{r}_m) = 1$ if the scatterer is exactly on the beam axis. Further, we recall from Section 4.2.3.2 the expression $I(\mathbf{r}_m) = P_l X_l^2(\mathbf{r}_m)$ is the local irradiance [Wm^{-2}] of the laser beam at the position of the scatterer. Using these definitions, we can re-arrange Eq. (7.25) into a form that defines the quantity $\partial\sigma$, known as the *differential scattering cross-section*:

$$\partial\sigma_s(\hat{\mathbf{s}}_\mathbf{o}|\hat{\mathbf{s}}_i,\hat{\mathbf{e}}_\mathbf{l}) = \frac{1}{Y_o(\mathbf{r}_m)}\frac{P_{os}}{I\Omega_o} = \frac{1}{k^2}\frac{\langle E_l^I|\mathbf{S}^\dagger\mathbf{S}|E_l^I\rangle}{\langle E^l|E^l\rangle} \tag{7.27}$$

Physically, $\partial\sigma$ specifies the angular distribution of the scattered light: the optical power (for unit incident irradiance) scattered into a unit angle in a given direction. The factor $1/Y_o\,(\mathbf{r}_m)$ corrects for the misalignment of the receiver. For future use, an alternate form of Eq. (7.27) is

$$\partial\sigma_s(\hat{\mathbf{s}}_\mathbf{o}|\hat{\mathbf{s}}_i,\hat{\mathbf{e}}_\mathbf{l}) = \frac{ik^4}{6\pi}\left|\frac{\alpha}{\epsilon_o}\right|^2\frac{3}{8\pi}\Pi \quad \text{where} \quad \Pi = \langle e_l^I|\mathbf{\Pi}^\dagger\mathbf{\Pi}|e_l^I\rangle \tag{7.28}$$

where $\mathbf{\Pi}$ is the polarization transfer matrix, Eq. (7.20). The right hand side defines the *polarization factor* Π. Note that in general $\partial\sigma_s(\hat{\mathbf{s}}_\mathbf{o}|\hat{\mathbf{s}}_i,\hat{\mathbf{e}}_\mathbf{l})$ depends on the orientation of the scattering particle with respect to the propagation direction and polarization of the incident field. This should be kept in mind when subsequently using $\partial\sigma_s\,(\theta,\,\phi)$ as an abbreviation for $\partial\sigma_s(\hat{\mathbf{s}}_\mathbf{o}|\hat{\mathbf{s}}_i,\hat{\mathbf{e}}_\mathbf{l}) = \partial\sigma_s(\theta,\,\phi|\hat{\mathbf{s}}_i,\hat{\mathbf{e}}_\mathbf{l})$.

By integrating $\partial\sigma(\theta,\phi))$ over the 4π solid angle one obtains the *total scattering cross-section* σ_t:

$$\sigma_t(\hat{\mathbf{s}}_\mathbf{l},\hat{\mathbf{e}}_\mathbf{l}) = \int_{\theta=0}^{\pi}\int_{\phi=0}^{2\pi} \partial\sigma_s(\theta,\phi)\sin(\theta)\,d\phi\,d\theta \tag{7.29}$$

In an elementary interpretation, σ_t can be viewed as the effective area of the molecules exposed to a photon stream. (Thus, a scattering system that consists of identical particles at a mean concentration $\bar{c}$ will exhibit a scattering coefficient $\mu_s = \sigma_t\bar{c}$ (see Section 7.2.2.1). To illustrate the concept of cross-sections we cast Eq. (7.27) in a form, which can be interpreted in probabilistic terms. The probability $\mathcal{P}_{os} = P_{os}/P_l$ that a photon is scattered from the laser beam into the observation beam is

$$\mathcal{P}_{os} = \Omega_o \frac{\sigma_t}{\pi a_l^2}\, p(\hat{\mathbf{s}}_\mathbf{o}|\hat{\mathbf{s}}_\mathbf{l},\hat{\mathbf{e}}_\mathbf{l})\, Y_o(\mathbf{r}_m)Y_l(\mathbf{r}_m) \tag{7.30}$$

The ratio $\sigma_t/\pi a_l^2$ is the geometrical probability that a photon traveling in a beam of an effective cross-section πa_l^2 is scattered by a target of cross-section σ_d. The factor $Y_o(\mathbf{r}_m)Y_t(\mathbf{r}_m)$ accounts for misalignment of illumination and observation. The product $\Omega_o p(\hat{\mathbf{s}}_\mathbf{o}\,|\hat{\mathbf{s}}_\mathbf{l},\hat{\mathbf{e}}_\mathbf{l})$ is the probability that a scattered photon flies into the solid angle Ω_o around the direction $\hat{\mathbf{s}}_\mathbf{o}$; the corresponding angular probability density is

$$p(\hat{\mathbf{s}}_\mathbf{o}|\hat{\mathbf{s}}_\mathbf{l},\hat{\mathbf{e}}_\mathbf{l}) = p(\theta,\phi|\hat{\mathbf{s}}_\mathbf{l},\hat{\mathbf{e}}_\mathbf{l}) = \frac{\partial\sigma(\hat{\mathbf{s}}_\mathbf{o}|\hat{\mathbf{s}}_i,\hat{\mathbf{e}}_\mathbf{l})}{\sigma_t(\hat{\mathbf{s}}_\mathbf{l},\hat{\mathbf{e}}_\mathbf{l})} \tag{7.31}$$

Note that the probability density $p\left(\theta,\,\phi|\,\hat{\mathbf{s}}_1,\,\hat{\mathbf{e}}_1\right)$ is *per definitio* normalized such that $\int p\left(\theta,\,\phi|\,\hat{\mathbf{s}}_1,\,\hat{\mathbf{e}}_1\right)\sin(\theta)\,d\theta d\phi = 1$. Note also that $p\left(\theta,\,\phi|\,\hat{\mathbf{s}}_1,\,\hat{\mathbf{e}}_1\right)$ is a generalized version of the "phase function" that was introduced in Chapter 3. Consequently we also define the *asymmetry parameter*

$$g = \langle\hat{\mathbf{s}}_\mathbf{i}\cdot\hat{\mathbf{s}}_\mathbf{o}\rangle = \langle\cos(\theta)\rangle = \int_{\theta=0}^{\pi}\int_{\phi=0}^{2\pi} p(\theta,\phi)\cos(\theta)\sin(\theta)\,d\phi\,d\theta \tag{7.32}$$

This g is more often called "anisotropy factor," but we prefer "asymmetry parameter" in order to avoid confusion with anisotropy of shape and dielectric response of the particle. Even in isotropic systems, the phase function is an asymmetric function of the scattering angle θ, because of the θ-dependence of the scattering vector $\mathbf{q}$.

In the context of radiative transfer, Eq. (7.30) is often expressed in radiometric terms, namely in terms of the *radiant intensity* I_Ω and the irradiance I (Note the symbol E used for irradiance in other chapters):

$$I_\Omega(\hat{\mathbf{s}}_\mathbf{o}) = \partial\sigma_s(\hat{\mathbf{s}}_\mathbf{o}|\hat{\mathbf{s}}_\mathbf{l},\hat{\mathbf{e}}_\mathbf{l})\, I(\mathbf{r}_m)Y_o(\mathbf{r}_m) = p(\hat{\mathbf{s}}_\mathbf{o}|\hat{\mathbf{s}}_\mathbf{l},\hat{\mathbf{e}}_\mathbf{l})\,\sigma_t(\hat{\mathbf{s}}_\mathbf{l},\hat{\mathbf{e}}_\mathbf{l})\, I(\mathbf{r}_m)Y_o(\mathbf{r}_m) \tag{7.33}$$

Here $I_\Omega(\hat{\mathbf{s}}_\mathbf{o}) = P_{os}/\Omega_o$ is the radiant intensity [W/sr] of the beam wave selected from the scattered field to propagate in the direction $\hat{\mathbf{s}}_\mathbf{o}$ and $I(\mathbf{r}_m) = P_l Y(\mathbf{r}_m)/\pi a_l^2$ is the irradiance [W/m^2] in the laser beam at the position $\mathbf{r}_m$ of the scatterer; recall Eq. (4.40). The alignment factor $Y_o(\mathbf{r}_m)$ is usually neglected.

For future use in a Monte Carlo algorithm we note that the "phase function" can be directly expressed in terms of the *normalized* S-matrix:

$$\hat{\mathbf{S}} = \frac{1}{\sqrt{k^2\sigma_t}}\mathbf{S} \tag{7.34}$$

so that

$$p(\hat{\mathbf{s}}_\mathbf{o}|\hat{\mathbf{s}}_\mathbf{l}, \hat{\mathbf{e}}_\mathbf{l}) = \frac{\langle E_l^I \left| \hat{\mathbf{S}}^\dagger \hat{\mathbf{S}} \right| E_l^I \rangle}{\langle E_l^I | E_l^I \rangle} = \frac{\langle E_l^l \left| \mathbf{R}_2^{-1} \hat{\mathbf{S}}^\dagger \mathbf{S} \hat{\mathbf{R}}_2 \right| E_l^l \rangle}{\langle E_l^l | E_l^l \rangle} \tag{7.35}$$

A considerable simplification arises for particles that are *isotropic* and *non-chiral*, both, in shape and in dielectric response. Then the off-diagonal elements S_3 and S_4 of the S-matrix are zero (Section 13.6 in [9]), and the phase function can be written as

$$\begin{aligned} p(\theta,\phi) &= |e_{xl}^I|^2 |\hat{S}_1|^2 + |e_{yl}^I|^2 |\hat{S}_2|^2 \\ &= \tfrac{|\hat{S}_1|^2+|\hat{S}_2|^2}{2} + \tfrac{|\hat{S}_1|^2-|\hat{S}_2|^2}{2} \{\cos(2\phi)\left[|e_{xl}^l|^2 - |e_{yl}^2|^2\right] + 2\sin(2\phi)\Re(e_{xl}^{l*} e_{yl}^l)\} \end{aligned} \tag{7.36}$$

where e_{xl}^I, e_{yl}^I and e_{xl}^l, e_{yl}^l are the components of the polarization vector $\hat{\mathbf{e}}_\mathbf{l}$ in the I-system and L-system, respectively.

7.1.6 Rayleigh Scatterer

In the present section we consider a truly elementary scatterer, namely a particle that is small enough with respect to the light wavelength so that it can be regarded as a point. Thus, the oscillating charge distribution excited by the illuminating field is an infinitesimal dipole oscillator $\mathbf{d}$ localized at $\delta(\mathbf{r} - \mathbf{r}_m)$:

$$\mathbf{d} = \mathbf{A}_d \cdot \mathbf{E}^l(\mathbf{r}_m) e^{-i\omega t} \tag{7.37}$$

Here $\mathbf{A}_d$ is the dipole polarizability tensor (recall Eq. (4.41)) and $\mathbf{E}^l(\mathbf{r}_m)$ is the illuminating field from Eq. (7.10). The magnitude of $\mathbf{d}$ follows the field oscillations, but the direction $\hat{\mathbf{d}}$ does not necessarily coincide with the impinging polarization. For example, in the case of uniaxial anisotropy with anisotropy axis $\hat{\mathbf{a}}$ one may write explicitly

$$\mathbf{d} \sim \mathbf{A}_d \cdot \hat{\mathbf{e}}_l = \alpha_i \hat{\mathbf{e}}_l + \Delta\alpha\ \hat{\mathbf{a}}\,\hat{\mathbf{a}} \cdot \hat{\mathbf{e}}_l = \alpha \mathbf{A} \cdot \hat{\mathbf{e}}_l \tag{7.38}$$

where $\alpha = \text{Trace}(\mathbf{A}_\text{d})/3 = \alpha_\text{i} + \Delta\alpha/3$. In the absence of anisotropy, the normalized matrix $\mathbf{A}$ reduces to the unity matrix $\mathbf{I}$ (1 on diagonal, 0 otherwise). The oscillating dipole radiates the scattered field $|E^s\rangle$ (recall Eq. (4.43)) whose source is the current density distribution

$$\mathbf{j}(\mathbf{r},t) = -i\,\omega\delta(\mathbf{r}-\mathbf{r}_m)\,\mathbf{d}(\mathbf{r}_m) = -i\,\omega\delta(\mathbf{r}-\mathbf{r}_m)\,\mathbf{A}_d\ \cdot \mathbf{E}^l(\mathbf{r},t) \tag{7.39}$$

This $\mathbf{j}(\mathbf{r},t)$ is just a simple special version of the general current density from Eq. (7.5), with $\epsilon_o\boldsymbol{\chi}(\mathbf{r}-\mathbf{r}')$ being replaced by $\delta(\mathbf{r}-\mathbf{r}_m)\delta(\mathbf{r}-\mathbf{r}')\mathbf{A}_d$. Thus, starting from Eq. (7.13), one verifies that it is only necessary to replace the multipole polarizability tensor $\mathbf{A}_m$ with the dipole tensor $\mathbf{A}_d$. With the simple dipole, it is quite easy to get an explicit expression for the total scattering cross-section σ_d. The power P_s radiated in the scattered field is obtained upon inserting the dipole moment $\mathbf{d}$ from Eq. (7.37) into Eq. (4.44) and a little work produces

$$\sigma_d(\hat{\mathbf{e}}_l) = \frac{k^4}{6\pi}\frac{|\alpha|^2}{\epsilon_o^2}\left(\hat{\mathbf{e}}_l^*\cdot\mathbf{A}^\dagger\mathbf{A}\cdot\hat{\mathbf{e}}_l\right) \tag{7.40}$$

Notice in σ_d the k^4-dependence. The scattering intensity increases with $1/\lambda^4$, which is characteristic for the so-called Rayleigh scatterers. Notice also that $\sigma_d\left(\hat{\mathbf{e}}_l\right)$ depends in general on the orientation of an anisotropic molecule with respect to the impinging polarization $\hat{\mathbf{e}}_l$, as expressed by the anisotropy factor $\left(\hat{\mathbf{e}}_l^*\cdot\mathbf{A}^\dagger\mathbf{A}\cdot\hat{\mathbf{e}}_l\right)$. For example, in the uniaxial case

$$\left(\hat{\mathbf{e}}_l^*\cdot\mathbf{A}^\dagger\mathbf{A}\cdot\hat{\mathbf{e}}_l\right) = \frac{\alpha_i}{\alpha}\left\{1+\gamma\,(1+\gamma)\,|\hat{\mathbf{a}}\cdot\hat{\mathbf{e}}_l|^2\right\},\quad \text{where}\quad \gamma = \Delta\alpha/\alpha_i \tag{7.41}$$

When calculating the cross-section of the elements of the polarization transfer matrix $\boldsymbol{\Pi}$ (recall Eq. (7.20), e.g. $\hat{\mathbf{x}}\cdot\mathbf{A}\cdot\hat{\mathbf{y}}_\mathbf{I}$), one should not forget that all involved quantities must be in the same coordinate system. Since the tensor $\mathbf{A}$ is a property of the scattering sample, one would transform the vectors into the lab-frame. Such transformations make the anisotropic scattering a tedious algebraic exercise, best done by a computer. A thorough discussion of Rayleigh scattering, including explicit expressions for polarizability tensors of various types of Rayleigh particles, can be found in Chapter 5 in [9].

Here we content ourselves with a compilation of the explicit results for an isotropic dipole scatterer. We replace in all formulas the dimensionless polarizability tensor $\mathbf{A}$ by the unity matrix $\mathbf{I}$. This implies that in Eq. (7.40) for the total cross-section σ_d we set $\left(\hat{\mathbf{e}}_l^*\cdot\mathbf{A}^\dagger\mathbf{A}\cdot\hat{\mathbf{e}}_l\right) = \hat{\mathbf{e}}_l^*\cdot\hat{\mathbf{e}}_l = 1$. The elements of the isotropic polarization transfer matrix $\boldsymbol{\Pi}$ can be more or less looked up in Fig. 7.3, so that we just write

$$\boldsymbol{\Pi}^d = \begin{pmatrix} 1 & 0 \\ 0 & \cos(\theta) \end{pmatrix} \quad \mathbf{S}^d = \frac{ik^3}{4\pi}\frac{\alpha}{\epsilon_o}\begin{pmatrix} 1 & 0 \\ 0 & \cos(\theta) \end{pmatrix} \tag{7.42}$$

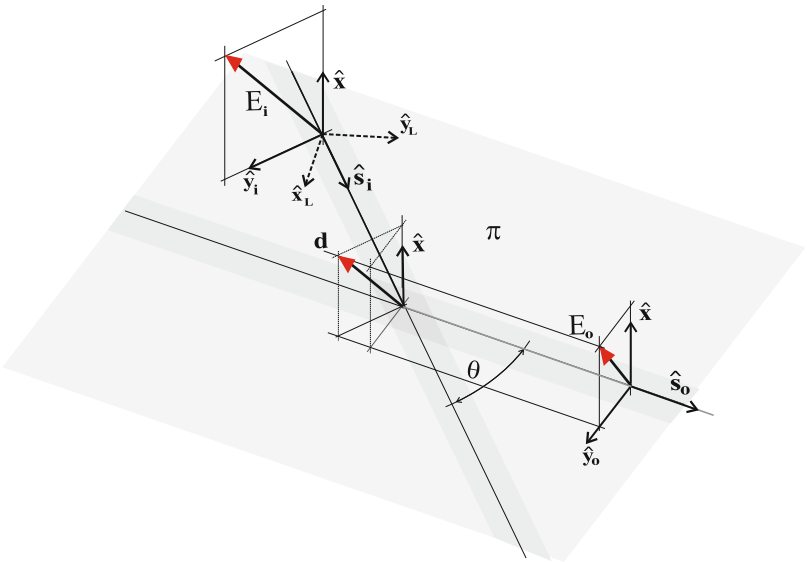


Fig. 7.3 Geometry of scattering from an isotropic dipole molecule

Such diagonal structure of the S-matrix, with $|S2/S1|^2 = \cos^2(\theta)$, is characteristic for single scattering from an isotropic system that fulfills the Born approximation. We also recall the normalized version of the S-matrix

$$\hat{\mathbf{S}}^d = \frac{1}{\sqrt{k^2\sigma_d}}\mathbf{S}^d = \sqrt{\frac{3}{8\pi}}\begin{pmatrix} 1 & 0 \\ 0 & \cos(\theta) \end{pmatrix} \tag{7.43}$$

from which we calculate the phase function of an isotropic dipole according to Eq. (7.36):

$$p_d(\theta,\phi|\hat{\mathbf{e}}_\mathbf{l}) = \frac{3}{8\pi}\Pi_d(\theta,\phi|\hat{\mathbf{e}}_\mathbf{l}) \tag{7.44}$$

where $\Pi_d(\theta,\phi|\hat{\mathbf{e}}_\mathbf{l})$ is the dipole polarization factor defined in Eq. (7.28):

$$\begin{aligned}\Pi_d(\theta,\phi|\hat{\mathbf{e}}_\mathbf{l}) &= |e^I_{xl}|^2 + |e^I_{yl}|^2\cos(\theta)^2 \\ &= \frac{1+\cos^2(\theta)}{2} + \frac{\sin^2(\theta)}{2}\left\{\cos(2\phi)\left[|e^l_{xl}|^2 - |e^l_{yl}|^2\right] + 2\sin(2\phi)\Re(e^{l*}_{xl}e^l_{yl})\right\}\end{aligned} \tag{7.45}$$

Here e^I_{xl}, e^I_{yl} and e^l_{xl}, e^l_{yl} are the components of the polarization vector $\hat{\mathbf{e}}_\mathbf{l}$ in the I-system and L-system, respectively. One can easily verify that $\int\int p(\theta,\phi^I)\sin(\theta)d\theta\, d\phi^I = 1$. We also note a couple of special cases. With a vertically polarized laser, so that $\hat{\mathbf{e}}_\mathbf{l} = \hat{\mathbf{x}}_L$, one obtains $\Pi_d(\theta,\phi|\hat{\mathbf{x}}_\mathbf{l}) = \cos^2(\phi) +$

$\sin^2(\phi)\cos^2(\theta)$. With a circularly polarized laser, only the first term in Eq. (7.45) survives. The same applies on average for so-called "natural" light, i.e. for randomly fluctuating polarization with uniform distribution of the polarization vectors $\hat{\mathbf{e}}_\mathbf{l}$.

7.2 Scattering from Interacting Rayleigh Particles

The simplest heterogeneous optical medium is a gas consisting of identical isotropic molecules. Now we perform a scattering experiment with a dilute gas as discussed in Section 7.1.1 and illustrated in Fig. 7.4. The only difference is that now the illuminated and observed region IOR formed by the profile overlap $X_o(\mathbf{r})\,X_l(\mathbf{r})$ contains many scatterers. The gas molecules are Rayleigh scatterers: when illuminated with a laser beam, they respond to $\mathbf{E}^l$ with dipole oscillations according to $\boldsymbol{\chi}_m = \alpha\mathbf{E}^l$. These oscillators are sources of molecular dipole fields $\mathbf{E}^m$ (Eqs. (4.43) and (4.42)), which in turn contribute to the excitation of neighbor molecules. However, since the polarizability α is small and a gas is dilute, we neglect these secondary fields assuming that $E^l \gg E^m$. In other words, we neglect all kinds of multiple scattering and assume that the molecular dipoles are driven by the unperturbed laser field $\mathbf{E}^l$, as it would exist in the absence of the scatterers. This is called the first order Born approximation.

In the old Maxwell's days, before the existence of molecules was established, one would describe this gas as a dielectric continuum with a susceptibility $\chi = \gamma c_{\text{mol}}$, where c_{mol} is the molar concentration (in mol/l) and γ is the molar specific susceptibility. Then one would use the constitutive relation Eq. (4.15): $\mathbf{P} = \epsilon_o \chi \mathbf{I}\cdot\mathbf{E}^l$, where $\mathbf{P}$ is the density of induced dipole moments. (The unity matrix $\mathbf{I}$ indicates explicitly that we restrict the discussion to isotropic response.) We are convinced of

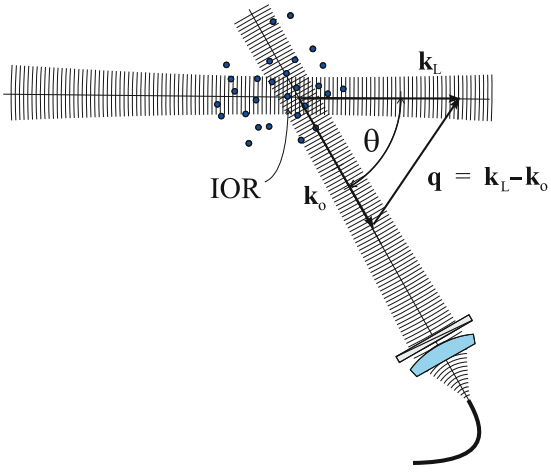


Fig. 7.4 Scattering experiment with single-mode receiver

the existence of molecules, but we can still use the continuum concept for a short hand notation:

$$\mathbf{P}(\mathbf{r},t) = \epsilon_o \chi(\mathbf{r},t)\mathbf{I} \cdot \mathbf{E}^l(\mathbf{r}) \quad \text{where} \quad \chi(\mathbf{r},t) = \frac{\alpha}{\epsilon_o} c_m(\mathbf{r},t) \tag{7.46}$$

Here $c_m(\mathbf{r},t)$ is the microscopic molecular concentration describing the distribution of the gas molecules. At the microscopic scale, the dilute gas is a highly inhomogeneous medium; $c_m(\mathbf{r},t)$ consists of δ spikes, a δ-function for each molecule m at its momentary position $\mathbf{r}_m(t)$:

$$c_m(\mathbf{r},t) \equiv \sum_m \delta[\mathbf{r} - \mathbf{r}_m(t)] \tag{7.47}$$

Because of thermal motion, the positions of the δ-spikes keep changing with time; $c_m(\mathbf{r},t)$ is a random and fluctuating distribution. The fluctuations of $c_m(\mathbf{r},t)$ are much slower than light oscillations $\exp(-i\omega t)$, which allows us to write the source of the scattered field, namely the current density distribution $\mathbf{j}(\mathbf{r},t) = \partial P/\partial t$ as

$$\mathbf{j}(\mathbf{r},t) = -i\,\omega c_m(\mathbf{r},t)\,\alpha\mathbf{I} \cdot \mathbf{E}^l(\mathbf{r}) = -i\,\omega\epsilon_o \chi(\mathbf{r},t)\mathbf{I} \cdot \mathbf{E}^l(\mathbf{r}) \tag{7.48}$$

This $\mathbf{j}(\mathbf{r},t)$ is a special version of the general current density from Eq. (7.5), with $\mathcal{X}(\mathbf{r}-\mathbf{r}')$ being replaced by $c_m(\mathbf{r})\alpha\mathbf{I}\delta(\mathbf{r}-\mathbf{r}')$. Thus, starting from Eq. (7.13), one finds that each of the elements T_{ol} of the $\mathbf{T}^{\mathbf{I}}_{\mathbf{lo}}$ matrix has the form

$$T_{ol}(t) = e^{ikL_{if}} \frac{ik}{2} \frac{\alpha}{\epsilon_o} \int X_o(\mathbf{r}) X_l(\mathbf{r})\, c_m(\mathbf{r},t)\, e^{i\,\mathrm{q}\cdot\mathbf{r}} d^3\mathbf{r}\, (\hat{\mathbf{e}}_\mathbf{i} \cdot \hat{\mathbf{e}}^*_\mathbf{o}) \tag{7.49}$$

$$= e^{ikL_{if}} \frac{ik}{2} \int X_o(\mathbf{r}) X_l(\mathbf{r})\, \chi(\mathbf{r},t)\, e^{i\,\mathrm{q}\,\cdot\,r} d^3 r\, (\hat{\mathbf{e}}_\mathbf{i} \cdot \hat{\mathbf{e}}^*_\mathbf{o}) \tag{7.50}$$

The second version is to keep in mind the continuum model, to which we shall return later. The Fourier integral in Eq. (7.49) expresses the interference of the random contributions of the individual scatterers to the scattering amplitude. Neglecting the broad IOR-profile $X_l(\mathbf{r})X_o(\mathbf{r})$, the amplitude is proportional to the $\mathbf{q}$-Fourier component of the particle distribution $c_m(\mathbf{r})$. This is fundamental to scattering experiments to elucidate the structure of matter. By choosing a certain $\mathbf{q}(\theta)$, one chooses the probing length scale, i.e. the wavelength $\lambda_q = 2\pi/q$ of a spatial harmonic in $c(\mathbf{r})$. Recall that $q^2 = k^2|\hat{\mathbf{s}}_\mathbf{o} - \hat{\mathbf{s}}_\mathbf{l}|^2 = 2\,k^2\,[1 - \cos(\theta)]$. Thus, with light scattering the shortest probing scale is $1/2\,k = \lambda/4\pi$, i.e. to roughly 50 nm in the visible range.

The wavelength λ_q corresponds to the spacing of fringes that are visible in the IOR in Fig. 7.2. As particles move randomly through the IOR, scattering amplitude fluctuates and this also results in fluctuations of the signal $P_{os}(t)$ (recall the definition in Eq. (7.23)). The measured quantity is the *time averaged* signal:

$$\langle P_{os}(t)\,\rangle = \langle |T_{xx}(t)|^2 \rangle |\mathcal{E}^I_{lx}|^2 + \langle |T_{yy}(t)|^2 \rangle |\mathcal{E}^I_{ly}|^2 \tag{7.51}$$

The mean square matrix elements are readily evaluated to

$$\langle |T_{ol}(t)|^2 \rangle = \frac{\sigma_d}{\pi a_l^2}\,\Omega_o\,\frac{3}{8\pi}|\hat{\mathbf{e}}_\mathbf{i}\cdot\hat{\mathbf{e}}_\mathbf{o}^*|^2\mathcal{S}(\mathbf{q}) \tag{7.52}$$

where

$$\mathcal{S}(\mathbf{q}) = \iint Y_{ol}(\mathbf{r})\langle c_m(\mathbf{r})c_m(\mathbf{r}')\rangle\, e^{i\,\mathrm{q}\cdot(\mathbf{r}-\mathbf{r}')}Y_{ol}(\mathbf{r}')\,d^3r\,d^3r' \qquad Y_{ol} = X_l X_o \pi a_l a_o \tag{7.53}$$

In the factor $(3/8\pi)|\hat{\mathbf{e}}_\mathbf{i}\cdot\hat{\mathbf{e}}_\mathbf{o}^*|^2$ we recognize the squared elements of the normalized S-matrix $\hat{\mathbf{S}}^d$, Eq. (7.43). Thus, recalling Eq. (7.36) we find that the normalized scattering signal $\mathcal{P}_{os} = P_{os}/P_l$ can be written as

$$\langle \mathcal{P}_{os}\rangle = \Omega_o\,\frac{1}{\pi a_l^2}\,\frac{k^4}{6\pi}\left|\frac{\alpha_d}{\epsilon_o}\right|^2 p_d(\hat{\mathbf{s}}_\mathbf{o}|\hat{\mathbf{s}}_\mathbf{l},\hat{\mathbf{e}}_\mathbf{l})\,\mathcal{S}(\mathrm{q}) \tag{7.54}$$

where $(k^4/6\pi)|\alpha_d/\epsilon_o|^2 = \sigma_d$ is the total differential cross-section of an elementary dipole and $p_d\,(\theta,\,\phi)$ is the dipole phase function from Eqs. (7.44) and (7.45). When comparing Eq. (7.54) with Eq. (7.30) for scattering from an arbitrary particle, one notices only one difference: the alignment factor $Y_o(\mathbf{r}_m)Y_l(\mathbf{r}_m)$ is replaced by the integral $\mathcal{S}\,(\mathbf{q})$.

7.2.1 Scattering and Fluctuations

Notice that in the integral in Eq. (7.53) a new quantity has appeared, namely the *spatial correlation function* $g(\mathbf{r},\mathbf{r}') = \langle c_m(\mathbf{r})c_m(\mathbf{r}')\rangle$ of the distribution $c_m(\mathbf{r})$. Recalling Eq. (7.47) we see that $g(\mathbf{r},\mathbf{r}')$ is the average of a double sum of δ-functions:

$$\langle c_m(\mathbf{r})c_m(\mathbf{r}')\rangle = \sum_m\sum_n \langle\delta[\mathbf{r}-\mathbf{r}_m(t)]\delta[\mathbf{r}-\mathbf{r}_n(t)]\rangle \tag{7.55}$$

This correlation function is central to the statistical physics of soft condensed matter and is likely to play an important role in the understanding of tissue optics. We can't go too deep into the enormous field of statistical physics, but a few hints are necessary to understand the role of fluctuations and interference in light scattering. First, we recall that any random quantity can be written as the sum of a mean and of fluctuations around mean. Thus, we write $c_m(\mathbf{r}) = \delta c(\mathbf{r}) + \bar{c}$, where $\bar{c} = \langle c_m(\mathbf{r})\rangle$ the macroscopic concentration that is defined as the ensemble of the microscopic distribution from Eq. (7.47). Thus, we decompose the density correlation into two terms, $g(\mathbf{r},\mathbf{r}') = \bar{c}^2 + g_\delta(\mathbf{r},\mathbf{r}')$, where $g_\delta(\mathbf{r},\mathbf{r}') = \langle\delta c_m(\mathbf{r})\delta c_m(\mathbf{r}')\rangle$ is the correlation function of density *fluctuations*. We assume for simplicity that the system is statistically homogeneous, which means that $g_\delta(\mathbf{r},\mathbf{r}') = g_\delta(\boldsymbol{\rho})$ depends only on the difference $\boldsymbol{\rho} = \mathbf{r}' - \mathbf{r}$ of the two positions $\mathbf{r}'$ and $\mathbf{r}$. (It is sufficient that we require statistical

homogeneity within IOR.) Reflecting on the meaning of the different terms in the double sum, Eq. (7.55), one finds that the fluctuation correlation function can be written as the sum of two terms, so that $g(\boldsymbol{\rho})$ consists of three terms

$$g(\boldsymbol{\rho}) = \bar{c}^2 + g_\delta(\boldsymbol{\rho}) = \bar{c}^2 + \bar{c}\,\delta(\boldsymbol{\rho}) + \bar{c}^2\,h(\boldsymbol{\rho}) \tag{7.56}$$

The first fluctuation term, the δ-correlation peak $\bar{c}\delta(\boldsymbol{\rho})$, comes from the $i = j$ terms of the double sum. This so-called self-correlation peak expresses the obvious fact that each particle is highly correlated with itself. The second fluctuation term $\bar{c}^2\,h(\boldsymbol{\rho})$, where $h(\boldsymbol{\rho})$ is the so-called pair correlation function, expresses the role of pairwise interactions of the particles. This is the contribution from the terms with $i \neq j$, concerning pairs of two distinct molecules. The positions of the particles are random, but the randomness has its limits, because the molecules do interact and their positions are therefore correlated. This correlation extends over a certain correlation length λ_c that corresponds to the characteristic width of the function $h(\boldsymbol{\rho})$.

Inserting $g(\boldsymbol{\rho})$ from Eq. (7.56) into Eq. (7.53), one obtains the Fourier integral $\mathcal{S}(\mathbf{q})$ as the sum of three contribution:

$$\mathcal{S}(\mathbf{q}) = \mathcal{S}_{\text{fwd}} + \mathcal{S}_{\text{inc}} + \mathcal{S}_{\text{coh}} \tag{7.57}$$

The first contribution S_{fwd} is due to the constant background $\bar{c}^2$:

$$\mathcal{S}_{\text{fwd}} = \bar{c}^2 \left| \int Y_{ol}(\mathbf{r}) e^{-i\,\mathbf{q}\cdot r} d^3 r \right|^2 \tag{7.58}$$

This integral is a 3D Fourier transform of the illuminated and observed region (IOR) profile, whose size is on the order of the beam radius a_w. It follows from properties of Fourier transformation that $\mathcal{S}_{\text{fwd}}$ will only contribute to forward scattering, i.e. at small q such that $q^2 < 1/a_w^2 = k^2\sigma_k^2$; recall Eq. (4.37). In this forward direction the receiver picks up predominantly the laser beam. Therefore forward scattering is omitted in standard scattering experiments. We can't do this when propagating light through the tissue. In fact, forward scattering is especially important, because it causes extinction. We shall continue this tricky topic in Section 7.2.2. For now we note an important finding: only the fluctuations $\delta c(\mathbf{r}, t)$ or $\delta\chi(\mathbf{r}, t)$ contribute to scattering at non-zero scattering angles.

The second contribution $\mathcal{S}_{\text{inc}}$ is due to the self-correlation peak $\bar{c}\delta(\boldsymbol{\rho})$:

$$\mathcal{S}_{\text{inc}} = \bar{c} \int Y_l(\mathbf{r}) Y_o(\mathbf{r}) d^3 r = \bar{c} V_{\text{IOR}} \tag{7.59}$$

Here V_{IOR} is the effective volume of the IOR and $\bar{c}V_{\text{IOR}}$ is the number of scatterers in the IOR. Each scatterer contributes independently according to its individual differential cross-section σ_d, viz. Eq. (7.52). There are no light interference effects and therefore this contribution is called "incoherent." Interference effects are

contained in the third contribution due to $\bar{c}h(\boldsymbol{\rho})$, which is therefore called the "coherent" term:

$$\mathcal{S}_{coh} = \bar{c}^2 \iint Y_{ol}(\mathbf{r}) Y_{ol}(\mathbf{r}+\boldsymbol{\rho}) h(\boldsymbol{\rho}) e^{-i\,\mathbf{q}\cdot\boldsymbol{\rho}} d^3r\, d^3\boldsymbol{\rho} \tag{7.60}$$

Here we introduce an additional assumption: the typical correlation length λ_c is much shorter than the size of the IOR, so that $h(\boldsymbol{\rho})$ can be regarded as a δ-function when compared with the broad profile $Y_{ol}(\mathbf{r})$. Then we can factor the double integration in $\int Y_{ol}^2(\mathbf{r}) d^3r \cdot \int h(\boldsymbol{\rho}) \exp(-i\,\mathbf{q}\cdot\ \boldsymbol{\rho}) d^3\boldsymbol{\rho}$, and the coherent contribution becomes

$$\mathcal{S}_{\text{coh}} = \bar{c} V_{\text{IOR}} \mathcal{H}(\mathbf{q})\, \bar{c} V_C = \mathcal{S}_{\text{inc}}\, \mathcal{H}(q)\, \bar{c} V_C \tag{7.61}$$

where

$$\mathcal{H}(\mathbf{q}) = \frac{1}{V_C} \int h(\boldsymbol{\rho}) \exp(-i\,\mathbf{q}\cdot\boldsymbol{\rho}) d^3\rho \qquad\qquad V_c = \int h(\boldsymbol{\rho}) d^3\rho \tag{7.62}$$

The function $\mathcal{H}(\mathbf{q})$ is normalized so that $\mathcal{H}(0) = 1$. The normalization constant V_c can be regarded as a *correlation volume* that expresses the tendency of molecules to cluster. Usually, the incoherent and coherent contributions are found combined into $\mathcal{S}_\delta = \mathcal{S}_{\text{inc}} + \mathcal{S}_{\text{coh}} = \bar{c} V_{\text{IOR}} S(\mathbf{q})$, where $S(\mathbf{q})$ is the so-called *structure factor*:

$$S(\mathbf{q}) = \frac{1}{\bar{c}} \int e^{i\,\mathbf{q}\cdot\boldsymbol{\rho}}\, g_\delta(\boldsymbol{\rho}) d^3\rho = 1 + \bar{c} V_c \mathcal{H}(\mathbf{q}) \tag{7.63}$$

The structure factor essentially contains all information about the structure obtained from elastic scattering experiments. In systems with isotropic interactions $h(\boldsymbol{\rho})$ depends only on $\rho = |\boldsymbol{\rho}|$ and the structure factor depends only on $q = |\mathbf{q}| = k\sqrt{2 - 2\cos(\theta)}$. In summary, neglecting forward scattering, the normalized time averaged signal $\langle \mathcal{P}_{os} \rangle_{\theta>0}$ can be written as

$$\langle \mathcal{P}_{os} \rangle_{\theta>0} = \Omega_o \frac{1}{\pi a_l^2} \frac{k^4}{6\pi} \left| \frac{\alpha_d}{\epsilon_o} \right|^2 p_d(\theta, \phi^I)\, S(\mathbf{q})\, \bar{c} V_{\text{IOR}} = \Omega_o\, \sigma_d p_d(\theta, \phi^I) S(\mathbf{q}) \bar{c} V_{\text{IOR}} \tag{7.64}$$

Because of interference, the measured intensity is not the sum of individual intensities. Neglecting the coherent contribution can't be justified by the vague argument of "randomness." Actually, the mutual arrangement is never quite random because all realistic particles do interact. What matters are the relations between three length scales: the mean inter-particle distance $d \approx \bar{c}^{1/3}$, their interaction range characterized by the correlation length $\lambda_C \approx V_C^{1/3}$ and the length scale $1/q$ probed in the scattering experiment. Coherent scattering can be neglected if the number of

molecules per correlation volume is small, i.e. if $\bar{c}V_c \ll 1$ and/or if $s(q) \ll 1$. The former condition is fulfilled in interstellar dust, but not in biological tissue, where the coherent term dominates in the optically accessible q-range.

7.2.2 Forward Scattering, Extinction and Scattering Coefficient

Since forward scattering is not forbidden in the propagation through a tissue, we must investigate the case $\mathbf{q} = 0$. The experiment is sketched in Fig. 7.5. A laser beam traverses a dilute cloud of interacting Rayleigh particles confined in a slab oriented perpendicularly to the propagation direction. After passing through the slab, the laser beam is attenuated, because of scattering. We also expect a slight phase shift with respect to the free propagation, because the cloud of dielectric particles may be regarded as a dielectric medium with mean susceptibility $\bar{\chi} = \bar{c}\alpha / \epsilon_o$.

To analyze the situation, we shall once again employ the projection formalism outlined in Chapter 4, Section 4.5 and in Section 7.1.2 of this chapter. We seek a Jones matrix $\mathbf{T}^t$ that characterizes the transmission from the plane at $z = 0$ to the plane $z = d$:

$$\begin{pmatrix} \mathcal{E}_x^t \\ \mathcal{E}_y^t \end{pmatrix} = \mathbf{T}^t \begin{pmatrix} \mathcal{E}_x^l \\ \mathcal{E}_y^l \end{pmatrix} \tag{7.65}$$

For simplicity we assume that the observation mode perfectly matches the laser beam:

$$\hat{\mathbf{E}}^{\mathbf{l}}(\mathbf{r}) = ...\sqrt{\frac{2}{c\epsilon}} X(\mathbf{r})\hat{\mathbf{e}}\, e^{ikz} \quad \hat{\mathbf{E}}^{\mathbf{o}}(\mathbf{r}) = ...\sqrt{\frac{2}{c\epsilon}} X(\mathbf{r})\hat{\mathbf{e}}\, e^{ik(z-d)} \tag{7.66}$$

Note that we choose a different phase reference plane for $\hat{\mathbf{E}}^{\mathbf{l}}$ and $\hat{\mathbf{E}}^{\mathbf{o}}$. (In this way we include the propagation phasor of the undisturbed laser beam in $\mathbf{T}^t$.) The laser field $|E^l\rangle$ is perfectly coupled into the receiver beam, but the receiver also picks

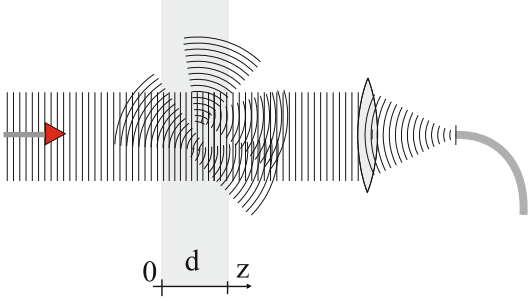


Fig. 7.5 Forward scattering and extinction

up the forwardly scattered field $|E^s\rangle$. In the transfer matrix, the superposition is expressed as

$$\mathbf{T}^t = \mathbf{I}_2 e^{i\,kd} + \mathbf{T}^S_{lo} \tag{7.67}$$

Here $\mathbf{I}_2$ is a two-dimensional identity matrix that represents the perfect coupling of the undisturbed laser beam. The phasor $\exp(ikd)$ accounts for the propagation over the distance d. The matrix $\mathbf{T}^s_{lo}$ is the usual scattering transition matrix, whose matrix elements are given by Eq. (7.49). However, now we are interested in the special case $\theta = 0$, i.e. $\mathbf{q} = \mathbf{0}$. This is a simplification, first of all because at $\theta = 0$ the polarization matrix $\mathbf{\Pi}_d$ simplifies to $\mathbf{I}_2$, which means that the two diagonal elements of $\mathbf{T}^s_{lo}$ are equal: $T^S_{xx} = T^S_{yy} = T^S$. The second simplification is $\mathbf{q} = \mathbf{0}$ in the integral in Eq. (7.49). Exploiting these simplifications, one finds that each row of Eq. (7.65) can be written as

$$\mathcal{E}^t(t) = [1 + i\mathcal{A}(t)]\; e^{ikd}\mathcal{E}^l \tag{7.68}$$

where $\mathcal{A}\,(t)$ is the *forward scattering amplitude*:

$$\mathcal{A}(t) = \frac{k}{2}\frac{\alpha}{\epsilon_o}\int_{z=0}^{d}\int_x\int_y |X(x,y)|^2\, c_m(\mathbf{r},t)dxdydz \tag{7.69}$$

The molecular distribution $c_m\,(\mathbf{r},\,t) \;=\; \bar{c} \;+\; \delta c\,(\mathbf{r},\,t)$ is a fluctuating quantity, and therefore $\mathcal{A}\,(t)$ fluctuates. However, when looking in the forward direction at $\mathbf{q} = \mathbf{0}$ the fluctuations are simply integrated over the IOR, and thereby more or less averaged out. Thus, in a good approximation we may replace $c_m(\mathbf{r}, t)$ with the time average $\bar{c}$. Recalling that $X(x, y)$ is quadratically normalized (Eq. (4.39)), the *pre-averaged* forward scattering amplitude is readily evaluated and the pre-averaged version of Eq. (7.68) reads

$$\mathcal{E}^t = \left(1 + i\frac{k\,d}{2}\frac{\alpha}{\epsilon_o}\bar{c}\right) e^{ik\Delta z}\mathcal{E}^l \tag{7.70}$$

Notice there is a serious problem: Eq. (7.70) is apt to violate energy conservation!! When increasing the slab thickness d, the forward scattering term will eventually dominate and grow towards infinity. This is because we neglected *multiple scattering*, i.e. the contribution of the secondary dipole fields to the field that drives the molecular oscillators. In the forward direction secondary waves are in phase, and thus amplified by constructive interference. We may neglect the secondary waves propagating backwards and sidewards, but we are not allowed to neglect them in the forward direction. In order to mend the problem, we sub-divide the slab into N thin layers of thickness $\Delta z = d/N$. Then, starting from the first layer, we include the forward scattered field from one layer in the driving field of the next layer. Mathematically, this is expressed as

$$\mathcal{E}^t(t,d) = \Pi_{j=1}^{N}\left[1 + i\mathcal{A}_j(t,\Delta z)\right]\; e^{ik\,d}\mathcal{E}^l \tag{7.71}$$

If we neglect fluctuations and assume that the medium is statistically homogeneous, then all the terms in the product are equal, and we can write

$$\mathcal{E}^t(d) = [1 + i\mathcal{A}(\Delta z)]^N \; e^{ik\,d}\mathcal{E}^l \tag{7.72}$$

Here we insert $\mathcal{A}(\Delta z) = k\Delta z\bar{\chi}/2$, where $\bar{\chi} = \bar{c}\alpha/\epsilon_o$ is the mean susceptibility of the medium. Since $\Delta z = d/N$, we remember that $\lim_{N\to\infty}(1 + 1/N)^N = e$, which yields:

$$\mathcal{E}^t(d) = e^{ikd\bar{\chi}/2}\, e^{ik\,d}\mathcal{E}^l = e^{ikd[1+\bar{\chi}/2]}\mathcal{E}^l \tag{7.73}$$

Since we expect that $\bar{\chi} \ll 1$, we can write $1 + \bar{\chi}/2 = \sqrt{1+\bar{\chi}} = \sqrt{\epsilon} = n$. Thus, finally we obtain

$$\mathcal{E}^t(d) = e^{inkd}\mathcal{E}^l \tag{7.74}$$

The argumentation was somewhat "handwaving," but it worked: the dilute scattering cloud behaves as an *effective medium* with refractive index $n > 1$. But where is the attenuation? It is not readily visible in Eq. (7.74), because so far we neglected the fact that the molecular polarizability α is a *complex* quantity: the driven molecular oscillators keep loosing energy by radiation and therefore there is a phase shift between the driving force $E^l \exp(-i\omega t)$ and the response, i.e. the oscillating dipole $d(t) = aE^l(t)$. In other words, we neglected the imaginary part of the polarizability α. The correct expression for the polarizability of a radiation damped oscillator is:

$$\frac{\alpha}{\epsilon_o} = i\left[1 - e^{i\delta}\right]\frac{3\pi}{k^3} \tag{7.75}$$

where δ is the so-called scattering phase (nothing to do with the "phase function" of radiative transfer), i.e. the phase shift of the radiation damped dipole oscillator with respect to the driving field. A simple algebra exercise yields the real and imaginary parts of the polarizability as

$$\Re\frac{\alpha}{\epsilon_o} = \frac{3\pi}{k^3}\sin(\delta) \qquad \Im\frac{\alpha}{\epsilon_o} = \frac{3\pi}{k^3}\left[1 - \cos(\delta)\right] \tag{7.76}$$

Upon inserting the complex susceptibility $\bar{\chi} = \bar{c}\,\alpha/\epsilon_o = \bar{\chi}_{re} + i\bar{\chi}_{im}$ into Eq (7.74), we obtain

$$\mathcal{E}^t(d) = e^{i\,n_{re}kd - kd\bar{\chi}_{im}/2}\mathcal{E}^l \tag{7.77}$$

where $n_{re} = \sqrt{1 + \bar{\chi}_{re}}$. Thus, by squaring the amplitude, we recover the Beer's law:

$$P_t(d) = e^{-k\bar{\chi}_{im}d}P_l \tag{7.78}$$

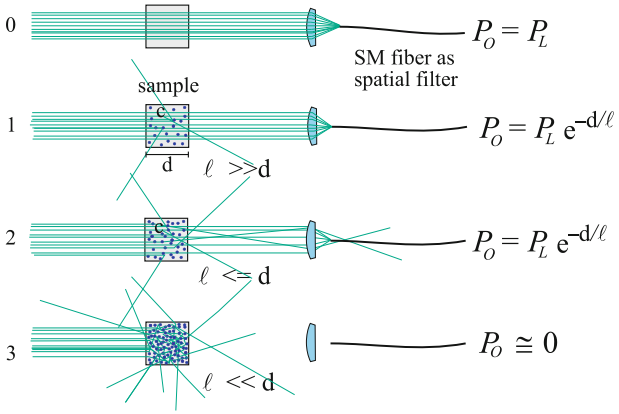


Fig. 7.6 Experiment demonstrating scattering regimes in a particle suspension with decreasing scattering mean free path $\ell = 1/1\sigma c$. (0) Reference; (1) Single scattering Beer's law; (2) Few scattering, still Beer's law; (3) Multiple scattering

The imaginary part of the susceptibility, that enters through the imaginary part of the forward scattering amplitude $\mathcal{A}$, is responsible for extinction. This is the essence of the so called *optical theorem* [11]. Usually one does not care about deep physical reasons of extinction and the Beer's law is formulated in terms of a measurable quantity, the so-called *scattering coefficient* μ_s:

$$P_t(d) = e^{-\mu_s d} P_l = e^{-d/\ell} P_l \quad \text{where} \quad \mu_s = k\bar{\chi}_{im} \tag{7.79}$$

The inverse $\ell = 1/\mu_s$ is the *scattering mean free path* (see Fig. 7.6).

Radiation is not the only way to loose energy of the molecular oscillator. Most often scattering is accompanied by absorption. Absorption can be included by an additional damping term in the oscillator equation, but a rigorous treatment requires quantum physics. We content ourselves with the corresponding modification of the Beer's law, i.e. we introduce the *extinction coefficient* μ_e (sometimes called "total attenuation coefficient," μ_t), that is the sum of μ_s and of the absorption coefficient μ_a:

$$\mu_e = \mu_s + \mu_a \tag{7.80}$$

We shall also use the extinction length $\ell_e = 1/\mu_e$. Since absorption is quite easily included in Monte Carlo simulations, we will forget about μ_a until the proper place.

7.2.2.1 Scattering Coefficient and Scattering Cross-Section

A particularly enlightening interpretation of the scattering coefficient is obtained if one realizes that

$$\Im\frac{\alpha}{\epsilon_o} = \frac{k^3}{6\pi}\frac{|\alpha|^2}{\epsilon_o^2} = \frac{1}{k}\sigma_d \tag{7.81}$$

where σ_d is the total scattering cross-section of an isotropic dipole oscillator. Then

$$\mu_s = \sigma_d \bar{c} \tag{7.82}$$

The scattering coefficient turns out to be the *density of scattering cross-section*. This interpretation, as well as Eq. (7.79), holds for an arbitrary scattering system. Note that the scattering coefficient from Eq. (7.82) is a purely incoherent superposition of the contributions of individual particles. There is no trace of interparticle interference. This is because interference effects other than constructive *forward* interference concern fluctuations and those we neglected when deriving Eq. (7.82). The rigorous inclusion of fluctuation is a demanding exercise and we content ourselves with a self-consistent approximation based on energy conservation.

According to Beer's law, the power removed from the beam upon passing a volume element δV is $\delta P = \mu_s I \delta V$ where I [W/m^2] is irradiation. Energy conservation requires that power removed equals the power radiated from this element by scattering. Energy conservation can be expressed as

$$\mu_s \, I \, \delta V = \int i_\Omega d\Omega \, \delta V \tag{7.83}$$

where i_Ω is the *density* of the radiant intensity [W sr^{-1} m^{-3}], i.e. $i_\Omega \delta V$ is the radiant intensity emanating from a volume element δV irradiated by I. The *differential* radiometric version of Eq. (7.64) is:

$$i_\Omega(\theta, \phi;\ \mathrm{r}) = \{\sigma_d \bar{c}\, p_d(\theta, \phi) S(\theta, \phi)\}\ I(\mathbf{r}) \tag{7.84}$$

As usual, $I(\mathbf{r}) = P_l Y(\mathbf{r})/\pi a_l^2$ is the local irradiance [Wm^{-2}] in the laser beam at the position $\mathbf{r}$ in the sample. The quantity in curly brackets can be identified as the *density of differential scattering cross-section*, abbreviated with $\{\partial\sigma\}$. Upon introducing i_Ω from Eq. (7.84) into Eq. (7.83), we find

$$\mu_s = \sigma_d \bar{c} \int p_d(\theta, \phi) S(\theta, \phi) \sin(\theta) d\theta d\phi \qquad p_R(\theta, \phi) = \frac{p_d(\theta, \phi) S(\theta)}{\int p_d(\theta, \phi) S(\theta, \phi) \sin(\theta) d\theta d\phi} \tag{7.85}$$

Here $S(\theta)$ is the structure factor, $p_d\,(\theta,\, \phi)$ is the phase function of an isolated dipole (Eq. (7.44)) and p_R is the phase function of interacting Rayleigh scatterers. These results are prototype for scattering problems in the Born approximation, i.e. in the absence of multiple scattering other than in the forward direction.

7.2.3 From Dilute Gas to Continuum and Back to Born Approximation

At this point we can make an attempt to be more specific about the meaning of "absence of multiple scattering" and about the validity of the Born approximation. It depends on the size d_{IOR} of the region of interest: one is in the single scattering regime if Beer's law, Eq. (7.78), applies for light beams traversing a path length d_{IOR} in an arbitrary direction through the sample (for illustration see Fig. 7.6). The criterion for the applicability of Born approximation is more strict.

In order to use the unperturbed laser beam in Eq. (7.46), we must require that $d_{\mathrm{IOR}} \ll \ell$. The Born approximation, in the strict sense as the absence of *elementary* multiple scattering, is not fulfilled in condensed matter. Even in such dilute gas as air at normal pressure the molecular concentration is as high as 10^8 molecules per $\mu\mathrm{m}^3$. The molecules are so close to each other that the electromagnetic interaction is dominated by the near field contribution $E^{\mathrm{near}} \sim d/(\epsilon_o r^3)$ (recall Eq. (4.42), where $r \approx \bar{c}^{1/3}$ is the mean intermolecular distance, and $d = \alpha E^l$ is the induced dipole moment. The elementary Born approximation requires that $E^{\mathrm{near}} \ll E^l$, which means that $\alpha\bar{c}/\epsilon_o = \bar{\chi} \ll 1$. This condition is well fulfilled in air where $\bar{\chi} = 0.0006$, but certainly not in water with $\bar{\chi} = 0.77$. In water one expects a significant contribution from multiple scattering, that is mediated by the near fields, instead of propagating photons. Fortunately, and somewhat paradoxically, this near field multiple scattering makes it possible to return to the Born approximation on a higher level: upon increasing the density, we end up with an electromagnetic quasi-continuum, whose susceptibility χ no longer reflects the polarizabilities of the individual molecules, but is a collective property of the molecules together with their neighbors. A rigorous theoretical description of the transition to electromagnetic continuum is a demanding task (see, e.g., [12]), but the practical consequence is favorable: the results as developed so far remain applicable as long as the fluctuations of the susceptibility are sufficiently small so that multiple scattering from the fluctuations can be neglected. The only modification we must make is to replace the vacuum permittivity ϵ_o with the effective medium permittivity $\epsilon_m = \epsilon_o(1 + \bar{\chi})$ and to multiply the vacuum wave number k with the corresponding refractive index n of the medium. The "unperturbed laser beam" is then to be understood as propagating in this homogeneous effective medium. A practical criterion for the applicability of the higher level of the first order approximation is Beer's law: the Born approximation becomes again applicable, if Beer's law applies over a length scale that is much larger than the region of interest. What is the "region of interest" depends on the context. The region of interest can be a sample cell, the IOR of a scattering experiment, or the structural correlation volume of size λ_C. In any case one must require $1/\ell = \mu_s \lll k$, which, according to Eq. (7.79), is equivalent to $\chi_{im} \ll 1$. This condition is well fulfilled with clean water whose scattering coefficient μ_s is as small as $0.0004\ \mathrm{m}^{-1}$ at 800 nm [13]. In fact, susceptibility fluctuations in water are sufficiently small that we can neglect them in present context, and regard water as a homogeneous medium. In this medium we suspend sophisticated species of highly structured molecules: proteins, polysaccharides, hydrophobic fats,

amphiphilic phospholipids. All these constituents will interact in manifold ways and self-organize into intriguing structures. We are about to create a biological tissue...

7.3 Scattering from Biological Tissues

7.3.1 Scattering from Particulate Systems

> Microstructures in biological tissue range from organelles 0.2–0.5 μm or smaller, to mitochondria 1–4 μm in length and 0.3–0.7 μm in diameter, to nuclei 3–10 μm in diameter, to mammalian cells 10–30 μm in diameter. The refractive index variation is about 0.04–0.10 for biological tissue with a background refractive index of $n_0 = 1.35$. [7]

In the present section we take the "tissue creation" seriously. We begin with a primordial soup containing the tissue constituents at low concentrations. The constituents of biological matter exhibit a strong tendency to aggregation that inevitably occurs upon increasing their concentration in water above a certain level. Phospholipids, as any surfactants, form micelles or vesicles, fats aggregate into droplets coated by a surfactant layer, amino-acids polymerize into peptides and protein macromolecules. These aggregates may be classified into four categories: R as Rayleigh, D as Debye, G as Gans and M for Mie.

Rayleigh particles: Aggregates of type R are small objects whose sizes a satisfy the inequality $ka << 1$. This small size has two consequences: (i) the internal structure of the aggregates can't be resolved by light scattering, whose shortest probing length $1/q$ is $1/2k$; (ii) they respond to electric field oscillations as simple dipoles with polarizability α. This so-called Rayleigh scattering is the subject of Sections 7.1.6 and 7.2. The only slight difference is that now we regard α as the *excess polarizability*, resulting from the mismatch of refractive index of the particle and the surrounding medium. For example, the excess polarizability of spherical fat droplets in water is well approximated by the Clausius-Massoti-Lorentz-Lorenz-formula

$$\frac{\alpha}{\epsilon_w} = 4\pi \frac{m^2 - 1}{m^2 + 2} a^3 \tag{7.86}$$

where $\epsilon_w = 1.33^2 \times \epsilon_o$ is the dielectric constant of water and $m = n/n_w$ the relative refractive index of the particle.

Debye particles: Aggregates of type D are large but fluffy objects, such as coils formed by linear macromolecular chains in good solvents. Since the coils are fluffy, we can neglect internal multiple scattering. Only the polarizabilities of the polymer units are slightly modified with respect to isolated monomers because of the near field coupling with their neighbors in the chain. Thus the scattering properties can be calculated using the Born approximation. However, the characteristic size a of the coils is large, so that $ka \geq 1$. Therefore one must take into account interferences of the dipole waves radiated by the individual segments of a single coil. These interferences are there, even if the polymer suspension is highly diluted.

Gans particles: Aggregates of type G are large compact objects whose refractive index differs only slightly from the surrounding medium. In other words: $ka \geq 1$ but $m \ll 1$. The scattering amplitude from such aggregates can be modeled as the coherent superposition of the contributions from internal dipoles driven by the external field E^l. This is quite similar to the Debye case, the only difference is of technical nature: in the Debye case one would use summation of contributions from discrete segments with polarizabilities α_s whereas with Gans particles one regards the interior of a particle as a continuum with an excess susceptibility χ_p.

RDG regime: The common feature of the RDG particles is that we may disregard intra-particle multiple scattering: the field acting on their molecular constituents is the external field E^l, e.g. the field of the illuminating laser, as it propagates through the homogeneous medium. In optics this is known as the Rayleigh-Debye-Gans-Born-approximation (RDGB).

Mie particles: Outside the RDG regime, scattering becomes quite complicated. Relatively simple expressions for particle scattering amplitudes of simple homogeneous objects such as spheres and ellipsoids are provided by Mie, Gans and followers, but even those need intense numerical calculations. More complicated shapes require ab initio numerics, for example the so-called T-matrix method. Fortunately, a biological tissue is not a dilute suspension of Mie particles; this does not even work for blood [14]. Therefore we can leave the subject "Mie and beyond" to standard texts [9, 10, 15, 16]. We shall venture into the Mie regime only briefly, in order to obtain a quantitative estimate of the applicability of the RDGB-approximation.

7.3.1.1 RDGB-Scattering

Consider a single RDD particle, such as a protein macromolecule. For the sake of generality we shall regard the macromolecule as a small fluctuating lump of a dielectric quasi-continuum, which scatters light because of its *excess susceptibility* χ_m with respect to the water background χ_w. The molecule is represented by the excess susceptibility distribution $\chi_m(\mathbf{r} - \mathbf{r}_\mathrm{m}, t)$, where $\mathbf{r}_m$ denotes the position of the molecule, perhaps the centroid of $\chi_m(\mathbf{r})$. For the sake of simplicity, we keep the position constant and assume that the size of the macromolecule, i.e. the size of the region where $\chi_m > 0$, is small with respect to the width a_l and a_O of the illumination and observation profiles. However, we allow the orientation, the form and the structure of the particle to fluctuate, in order to cover the case of flexible chain-like macromolecules.

In Section 7.1.1 we worked out the formal theory for a general particle illuminated and observed by a quasi-plane wave and now we can earn the fruits. The only difference is that now we employ the Born approximation, which means that the non-local susceptibility tensor $\epsilon_o \boldsymbol{\mathcal{X}}(\mathbf{r}, \mathbf{r}')$ is replaced by the local excess susceptibility $\epsilon_w \chi(\mathbf{r}) \delta(\mathbf{r} - \mathbf{r}')$. The subscript w is a reminder of the fact that now we work in water with susceptibility $\bar{\chi}_w$, which also enters the wavelength λ_w and all derived quantities. To further simplify the matters, we assume isotropy of the response so

that $\epsilon_w \chi(\mathbf{r})\delta(\mathbf{r}-\mathbf{r}') = \epsilon_w \chi_m(\mathbf{r})\delta(\mathbf{r}-\mathbf{r}')\mathbf{I}$, where $\mathbf{I}$ is the unit matrix. We insert this expression into Eq. (7.15) and calculate the polarizability tensor of the molecule:

$$\mathbf{A_m}(\mathrm{q}, t) = \mathbf{I} \int_V e^{i\,\mathbf{q}_w \cdot \mathbf{u}}\, \epsilon_w \chi_m(\mathbf{u}, t)\, d^3u \tag{7.87}$$

where $\mathbf{u} = \mathbf{r} - \mathbf{r}_m$. The next step consists in two definitions:

$$\alpha_m = \int \epsilon_w \chi_m(\mathbf{u}, t) d^3u \qquad \mathcal{F}(\mathbf{q}, t) = \frac{1}{\alpha_m} \int e^{i\,q_w \cdot \mathbf{u}}\, \epsilon_w \chi_m(\mathbf{u}, t) d^3u \tag{7.88}$$

Here α_m is the excess polarizability of the macromolecule. Note that α_m is the sum of the polarizabilities of the molecular constituents of the macromolecule and therefore is approximately constant in time. The quantity $\mathcal{F}(\mathbf{q}, t)$ is the so-called *form amplitude*. The polarization transfer matrix of the RGD-particle is that of an elementary dipole. Therefore the S-matrix is essentially a copy of Eq. (7.42), the only new feature is the form amplitude $\mathcal{F}(\mathbf{q}, t)$:

$$\mathbf{S}^{RDG}(t) = \frac{ik_w^3}{4\pi} \frac{\alpha_m}{\epsilon_w} F(\mathbf{q}_w, t) \begin{pmatrix} 1 & 0 \\ 0 & \cos(\theta) \end{pmatrix} \tag{7.89}$$

The form amplitude of a general RGD-particle is a fluctuating quantity. The measured signal depends on its average $F(\mathbf{q}) = \langle |\mathcal{F}(\mathbf{q}, t)|^2 \rangle$, which is called *form factor*, akin to the structure factor $S(\mathbf{q})$ from Section 7.2.1:

$$F(\mathbf{q}) = \langle |\mathcal{F}(\mathbf{q}, t)|^2 \rangle = \left| \frac{\epsilon_w}{\alpha_m} \right|^2 \int \int e^{i\,\mathrm{q}_w \cdot (\mathbf{u}-\mathbf{u}')}\, \gamma_m(\mathbf{u}, \mathbf{u}')\, d^3u d^3u' \tag{7.90}$$

where $\gamma_m(\mathbf{u}, \mathbf{u}') = \langle \chi_m(\mathbf{u}, t) \chi_m^*(\mathbf{u}', t) \rangle$is the susceptibility correlation function that is akin to the correlation function $g_\delta(\rho)$ of density fluctuations from Eq. (7.56). In terms of the form factor, the normalized time-averaged signal from a single particle is:

$$\langle \mathcal{P}^1_{rgd} \rangle_{\theta>0} = \Omega_w \frac{1}{\pi a_l^2} \left[\frac{k_w^4}{6\pi} \left| \frac{\alpha_m}{\epsilon_w} \right|^2 F(\mathbf{q}_w)\, p_d(\theta, \phi^I) \right] \tag{7.91}$$

Here $p_d(\theta, \phi)$ is the phase function of an isolated dipole, Eqs. (7.44) and (7.45). The bracketed quantity is the differential scattering cross-section $\partial\sigma_{rgd}$ of the RGD-particle. Correspondingly, the total scattering cross-section is the integral

$$\sigma_{\mathrm{rgd}} = \frac{k_w^4}{6\pi} \left| \frac{\alpha_m}{\epsilon_w} \right|^2 \int p_d(\theta, \phi) F(\theta, \phi) \sin(\theta) d\theta d\phi \tag{7.92}$$

Finally, the phase function p_R of RGD scatterers:

$$p_{\mathrm{rgd}}(\theta,\phi) = \frac{p_d(\theta,\phi)F(\theta,\phi)}{\int p_d(\theta,\phi)F(\theta,\phi)\sin(\theta)d\theta d\phi} \tag{7.93}$$

Thus, Eq. (7.91) can be abbreviated as

$$\langle \mathcal{P}^1_{\mathrm{rgd}} \rangle_{\theta>0} = \Omega_w \frac{\sigma_{\mathrm{rgd}}}{\pi a_l^2} p_{\mathrm{rgd}}(\theta,\phi) \tag{7.94}$$

Obviously, it is quite difficult to keep in the illuminated and observed region (IOR) of a single particle. Usually one works with a suspension of a certain number density $\bar{c}$. If the suspension is sufficiently dilute so that interactions between the particles can be neglected, then

$$\langle \mathcal{P}_{\mathrm{rgd}} \rangle_{\theta>0} = \Omega_w \frac{1}{\pi a_l^2} \left\{ \frac{k_w^4}{6\pi} \left| \frac{\alpha_m}{\epsilon_w} \right|^2 F(\mathbf{q}_w) p_d(\theta,\phi^I)\bar{c} \right\} V_{\mathrm{IOR}} = \Omega_w \frac{1}{\pi a_l^2} \left\{ \sigma_{\mathrm{rgd}} p_{\mathrm{rgd}}(\theta,\phi)\bar{c} \right\} V_{\mathrm{IOR}} \tag{7.95}$$

Notice in the curly brackets the density of differential scattering cross-section $\left\{\partial\sigma_{\mathrm{rgd}}\right\}$ of a dilute suspension of RGD particles. Interactions become significant as $\bar{c}$ increases and the work of Section 7.2 must be redone with RGD particles in place of the elementary dipoles. However, since a dilute suspension is a bad model for biological tissue, we content ourselves with Eq. (7.95).

7.3.1.2 Interpreting form Factors

For future use we will compile here the form factors of basic particle shapes as they can be found in many texts on small angle scattering. On the right hand sides of the following equations, we indicate the limiting behavior for large q, when probing length scales are much shorter than the size of the particle. To make this limiting behavior visible, the form factors in Fig. 7.7 are plotted in log-log-representation. All formulas assume that the scatterers are isotropic in the statistical sense, for example, because of free random rotation. Then $F(\mathbf{q})$ depends only on $q = |\mathbf{q}| = k\sqrt{2-2\cos(\theta)}$.

The upper curve in Fig. 7.7 represents homogeneous rods of length L:

$$F_{\mathrm{rod}}(q) = \frac{2}{qL}\left[\mathrm{Si}(qL) - \frac{1-\cos(qL)}{qL}\right] \quad \to \frac{1}{q} \tag{7.96}$$

The q^{-1} law at large q reflects the fact that a thin rod is a 1-dimensional object.

Thin homogeneous disc-like sheet of radius R:

$$F_{\mathrm{disc}}(q) = \frac{2}{q^2R^2}\left[1 - \frac{\mathrm{J}_1(2qR)}{qR}\right] \quad \to \frac{1}{q^2} \tag{7.97}$$

Again, q^{-2} law at large q reflects the dimensionality of the homogeneous structure.

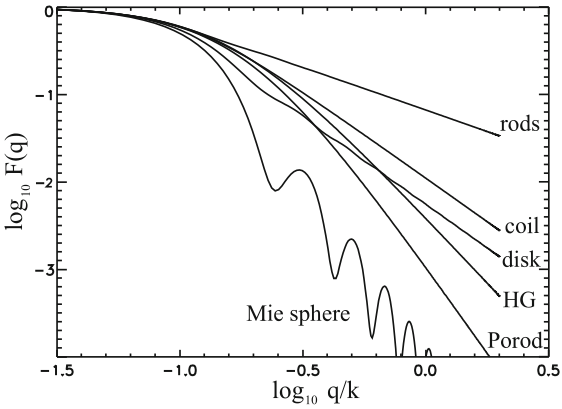


Fig. 7.7 Form factor of simple scatterers

The form-factor for thin spherical shell of radius R may be of a special interest because spherical vesicles are a good models of cell membranes:

$$F_{\text{disc}}(q) = \frac{\sin^2(qR)}{q^2R^2} \qquad \text{(smoothed)} \rightarrow \frac{1}{q^2} \tag{7.98}$$

Neglecting the oscillation, which are smoothed out in a polydisperse system, one again sees the q^{-2} behavior. This indicates the 2D sheet of which the shell membrane is made. Another frequently encountered object is a sphere, for example the nucleus of a cell. A homogeneous sphere of radius R:

$$F_{\text{sph}}(q) = \left\{ \frac{3[\sin(qR) - qR\cos(qR)]}{(qR)^3} \right\}^2 \qquad \rightarrow \frac{1}{q^4} \tag{7.99}$$

The simple dimensionality rule breaks down at dimension 3: there is no way to look into a homogeneous 3D structure (at large q) because of the lack of scattering. Only the surfaces of such a 3D object are visible. q^{-4} is characteristic for randomly oriented smooth interfaces. According to Porod and Debye, any two-phase system with smooth interfaces between the phases will scatter so that $F(q) \sim q^{-4}$ at large q.

We also consider two objects of fractal character. Fractals are structures that exhibit a certain degree of self-similarity [17]. When investigating the object on shorter and shorter length scales (i.e. at larger and larger q), and adjusting the magnification accordingly, one keeps finding the same picture, at least in the statistical sense. In practice, however, there are two cut-off lengths, L and l, above and below which the fractal picture breaks down: large L is the overall size of the object and small l is the size of its constituent monomers. A classical fractal is a random polymer coil with a characteristic size R_g:

$$F_{\text{coil}}(q) = \frac{2[\exp(-q^2R_g^2) + qR_g - 1]}{q^2R_g^2} \qquad \rightarrow \frac{1}{q^2} \tag{7.100}$$

When magnifying the polymer coil, one always sees the same randomly entangled loops. The coil is a "mass fractal." It is a 3D object, but it does not fill the 3D space. The fractal dimension of the coil is 2, and therefore q^{-2}. But the same large q behavior exhibits the 2D disk or shell. It is not easy to identify a true fractal [18]. Therefore we shall not dwell on the fractal point. For future use we will record only one more form factor, namely one that belongs to an UFO (Unidentified Fractal Object) that is also known under the name "HG":

$$F_{\mathrm{HG}}(q) = \frac{1}{(1+q^2L^2)^{3/2}} \quad \to \frac{1}{q^3} \tag{7.101}$$

An advanced reader interested in fractals will have noticed that this form factor is exactly on the border line between a mass fractal and a surface fractal.

Common to all six form factors is their limiting behavior at $q \to 0$, in the so-called Guinier regime at nearly forward scattering angles θ. Expanding Eq. (7.90) in a Taylor series in q, one finds that

$$F(q) = 1 - \frac{q^2R_g^2}{3} + \mathcal{O}(q^4R_g^4) \quad \text{where} \quad R_g^2 = \frac{\int\int \boldsymbol{\rho^2}\, \boldsymbol{\gamma}_m(\mathbf{u}, \mathbf{u}+\boldsymbol{\rho})\, d^3u d^3\boldsymbol{\rho}}{\int\int \gamma_m(\mathbf{u}, \mathbf{u}+\boldsymbol{\rho}) d^3u d^3\boldsymbol{\rho}} \tag{7.102}$$

Here $\gamma_m(\mathbf{u}, \mathbf{u}') = \langle \chi_m(\mathbf{u}, t)\chi_m^*(\mathbf{u}', t)\rangle$ is the susceptibility correlation function. R_g is the so-called radius of gyration, a universal measure of the size of the scattering object. All form factors in Fig. 7.7 correspond to the same R_g.

7.3.1.3 On the Applicability of the RDGB-Approximation

A single isolated Mie sphere may not be a good model for light scattering in tissue, but since Mie algorithms are readily available,[2] we can use it as a simple model to estimate the applicability of the Born approximation for modeling the optical properties of tissues. We consider a rather large sphere with a diameter $2a \approx 3.3\mu$m, which gives the size parameter $ka = 17.3$ when using a 800 nm laser. The refractive index of the sphere is 1.42 and the surrounding medium is water. The sphere may represent the nucleus of a mammalian cell.

In Fig. 7.8 we plot the Mie intensity $I(q) \sim \left(|S1|^2 + |S2|^2\right)$ together with the RDG-form factor from Eq. (7.99). Such a large sphere scatters predominantly forward, with an asymmetry parameter $g = 0.98$ (recall Eq. (7.32)). In the linear plot the difference between Mie and RGD is hardly discernable; and even at large angles, that are better visible in the log-log plot, the deviations appear to be marginal. Not as good, but still reasonable, is the agreement of the total differential cross-section. The exact Mie result is $\sigma_M = 2.12 \times \pi a^2$, whereas the RDG approximation, Eqs. (7.86) and (7.92), yields $\sigma_{\mathrm{RDG}} = 2.64 \times \pi a^2$. Since we are interested in polarization, we also plotted in Fig. 7.9 the ratio $\left|S2/S1\right|^2$. In the Born approximation, one expects

[2] We use an IDL version of Bohren-Huffman code [9] by B.T. Drain and P.J. Plateau.

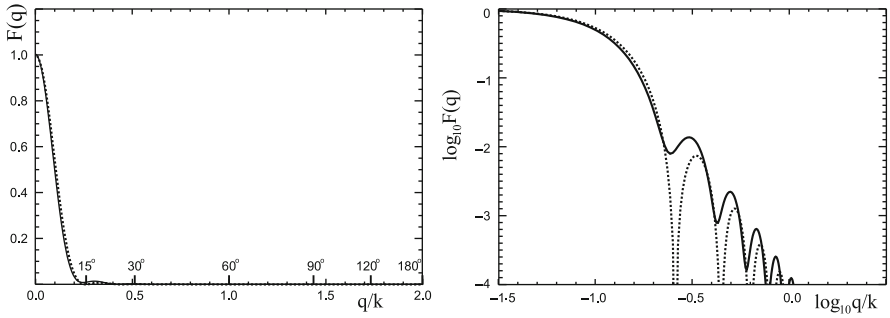


Fig. 7.8 Comparison of a Mie sphere (*solid line*) and RDG sphere (*dotted line*)

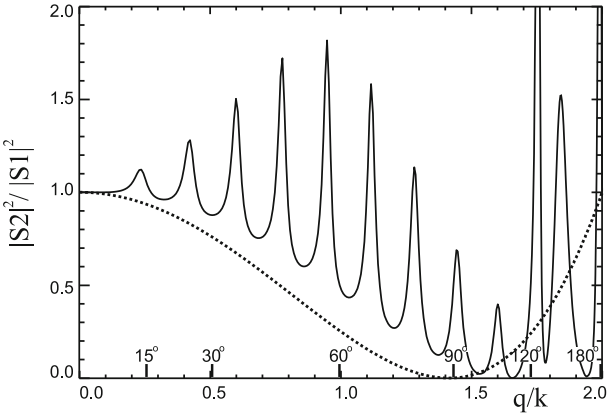


Fig. 7.9 Ratio $|S2/S1|^2$ of a Mie sphere (*solid line*) and a RDG sphere (*dotted line*)

$\left|S2/S1\right|^2 = \cos^2(\theta)$. The corresponding Mie curve is considerably wilder, indicating a resonant excitation of higher order multipoles. Nevertheless, even in the Mie range, the dipolar character of the polarization remains to some extent preserved. Except for the resonance peaks, that are easily smoothed out by size averaging, there is only a slight shift of the minimum towards large angles. Because of the strong forward anisotropy of the scattering, the large angles play a minor role. Thus, we conclude: when modeling the optical properties of tissues, it is hardly worth to bother with Mie spheres or more complicated Mie objects. RDGB approximations should be sufficient in most cases.

7.3.2 Scattering from a Continuum

While propagating through a tissue, a light wave can encounter structures with dimensions ranging from the width of a protein macromolecule (50–100 nm), to the diameter of the body of a cell (2–50 μm), to the length of a blood vessel (0.1 mm). Moreover, the constituents of tissues are tightly compacted, with their surfaces pressed together to form

contiguous units. Theories based on incoherent scattering among discrete particles cannot be applied at the microscopic level in tissue without violating their underlying assumptions. [19]

Equipped with the knowledge about scattering from dilute particulate systems, we return to the tissue creation laboratory. We increase the density of bio-matter in water to realistic volume fractions of about 10–20%. At such a high density, the bio-particles, such as membrane sheets, collagen rods and coils, protein spheres and coils, find themselves well within the range of their interaction forces. The picture of independently scattering particles is no longer meaningful. Thus, we revert to the continuum picture of a medium with fluctuating susceptibility $\chi(\mathbf{r}, t)$. We have already done most of the work and so we only need to recapitulate the results: (i) scattering at finite angles is due to *fluctuations* $\delta\chi(\mathbf{r}, t)$ (ii) forward scattering is included in the *scattering coefficient* μ_s and in the *mean refractive index* n_m of the medium where n_m corresponds to the real part of the mean susceptibility $\bar{\chi}$. Equipped with this knowledge, we return to Eq. (7.50) but express the matrix element $\mathbf{T}^{\mathbf{I}}_{\mathbf{lo}}$ in terms of the susceptibility fluctuations instead of molecular concentration:

$$T_{ol}(t) = e^{ik_m L_{if}}\, e^{-\mu_s L_{if}/2}\, \frac{i\,k_m}{2} \int X_o(\mathbf{r}) X_l(\mathbf{r})\, \delta\chi(\mathbf{r}, t)\, e^{i\,q_m \cdot r} d^3 r\, (\hat{\mathbf{e}}_\mathbf{i} \cdot \hat{\mathbf{e}}^*_\mathbf{o}) \tag{7.103}$$

The subscript m at k_m and q_m is a reminder of the fact that we now work in an effective medium with susceptibility χ_m, that also enters the wave number and wavelength λ_m. When calculating the time averaged signal that is proportional to $\langle |T_{ol}(t)|^2 \rangle$, one naturally encounters the correlation function of susceptibility fluctuations:

$$\gamma_\delta(\boldsymbol{\rho}) = \langle \delta\chi(\mathbf{r}, t) \delta\chi^*(\mathbf{r} + \boldsymbol{\rho}, t) \rangle \tag{7.104}$$

We assume statistical homogeneity within the IOR, i.e. $\gamma_\delta(\mathbf{r}, \mathbf{r} + \boldsymbol{\rho}) = \gamma_\delta(\boldsymbol{\rho})$ depends only on the displacement $\boldsymbol{\rho} = \mathbf{r}' - \mathbf{r}$. In the continuum model, $\gamma_\delta(\boldsymbol{\rho})$ replaces the correlation function $g_\delta(\boldsymbol{\rho})$ of density fluctuations of the interacting particle model in Section 7.2; recall Eq. (7.56). Recall also that $g_\delta(\boldsymbol{\rho})$ diverges for $\boldsymbol{\rho} = 0$, there is the self-correlation term $\bar{c}\delta(\boldsymbol{\rho})$. A similar behavior can be expected for $\gamma_\delta(\boldsymbol{\rho})$. Thus the fluctuation variance $\gamma_\delta(0) = \langle |\delta\chi|^2 \rangle$ is not a good measure of the strength of scattering, despite erroneous statements in some texts. As it turns out, a better measure of the strength of scattering is the integral of the correlation function

$$V_\chi = \lambda_\chi^3 = \int \gamma_\delta(\boldsymbol{\rho}) d^3\boldsymbol{\rho} \tag{7.105}$$

Note that susceptibility is a dimensionless quantity and therefore V_χ can be interpreted as the susceptibility correlation volume. Likewise, λ_χ is the susceptibility correlation length, i.e. the range of ρ where $\gamma_\delta(\boldsymbol{\rho})$ is non-zero. Having reflected on the properties of the susceptibility correlation function, we calculate the normalized

time averaged signal $\langle \mathcal{P}_{os} \rangle$. Recall that $\langle \mathcal{P}_{os} \rangle$ is the probability that a photon traveling in a laser beam from a plane at $\mathbf{r}_i$ is scattered from the IOR centered at $\mathbf{r}_f$ into the observation mode. Expressed in terms of $\gamma_\delta(\boldsymbol{\rho})$ the normalized signal reads

$$\langle \mathcal{P}_{os} \rangle_{\theta>0} = e^{-\mu_s L_{if}} \Omega_m \frac{1}{\pi a_l^2} \frac{k_m^4}{6\pi} \int \int Y_{ol}(\mathbf{r})\, \gamma_\delta(\mathbf{r}', \mathbf{r})\, e^{i\, \mathrm{q}_m \cdot (\mathbf{r}-\mathbf{r}')} Y_{ol}(\mathbf{r}')\, d^3 r\, d^3 r'\, p_d(\theta, \phi) \tag{7.106}$$

where $p_d(\theta, \phi)$ is the dipole phase function, which appears here, because the Born approximation preserves the dipole character of polarization transfer. If λ_χ is much shorter than the size of the IOR, then we factor the integration and write

$$\langle \mathcal{P}_{os} \rangle_{\theta>0} = e^{-\mu_s L_{if}} \Omega_m \frac{1}{\pi a_l^2} \frac{k_m^4}{6\pi} S_\chi(\mathbf{q})\, p_d(\theta, \phi)\, V_{\mathrm{IOR}} \tag{7.107}$$

where $V_{\mathrm{IOR}} = \int Y_o(\mathbf{r}) Y_l(\mathbf{r}) d^3 r$ is the effective volume of the IOR and $S_\chi(\mathbf{q})$ is the spatial power spectrum of the susceptibility fluctuations, i.e. Fourier transform of the correlation function $\gamma_\delta(\boldsymbol{\rho})$:

$$S_\chi(\mathrm{q}) = \int \gamma_\delta(\boldsymbol{\rho}) e^{i\, q \cdot \boldsymbol{\rho}} d^3 \boldsymbol{\rho} \tag{7.108}$$

Note that this $S_\chi(\mathbf{q})$ is analogous with the structure factor $S(\mathbf{q})$ from Eq. (7.63), but here the pre-factor $1/\bar{c}$ is missing. A convenient normalization that is appropriate for the present case is to use $F(\mathbf{q}) = S_\chi(\mathbf{q})/V_\chi$; recall that $F(\mathbf{0}) = 1$. So we write

$$\langle \mathcal{P}_{os} \rangle_{\theta>0} = e^{\mu_s L_{if}} \Omega_m \frac{1}{\pi a_l^2} \left\{ \frac{k_m^4}{6\pi} V_\chi F(\mathbf{q})\, p_d(\theta, \phi) \right\} V_{IOR} \tag{7.109}$$

Obviously, the quantity in curly brackets is the density of *differential* scattering cross-section $\{\partial \sigma_\chi\}$. Thus, the density of *total* scattering cross-section, i.e. the scattering coefficient μ_s is

$$\mu_s = \int \{\partial \sigma_\chi\} \sin(\theta) d\theta d\phi = \frac{k_m^4}{6\pi} V_\chi \int F(\mathbf{q})\, p_d(\theta, \phi) \sin(\theta) d\theta d\phi \tag{7.110}$$

and the phase function p_χ for scattering from susceptibility fluctuations reads, as usual

$$p_\chi(\theta, \phi) = \frac{F(\mathbf{q})\, p_d(\theta, \phi)}{\int F(\mathbf{q})\, p_d(\theta, \phi) \sin(\theta) d\theta d\phi} \tag{7.111}$$

where $p_d(\theta, \phi)$ is the dipole phase function. Recall that ϕ is the azimuth scattering angle in L-system. The abbreviated version of Eq. (7.109) reads

$$\langle \mathcal{P}_{os} \rangle_{\theta>0} = e^{\mu_s L_{if}} \Omega_m \frac{\mu_s}{\pi a_l^2} p_\chi(\theta, \phi) V_{IOR} \tag{7.112}$$

7.3.2.1 Quasi-Particles

When comparing Eq. (7.109) with the corresponding Eq. (7.95) for independently scattering RGD-particles, one finds only two differences. The first difference is the attenuation factor $\exp\left(-\mu_s L_{if}\right)$, which we introduced to account for the effective mediums complex susceptibility $\bar{\chi}$; recall Section 7.2.2. The second difference is one replacement:

$$V_\chi \leftrightarrow \left| \frac{\alpha_m}{\epsilon_w} \right|^2 \bar{c} \tag{7.113}$$

In the continuum model, the product of mean concentration and the square of particle polarizability is replaced with the susceptibility correlation volume V_χ. So it turns out that the RDGB theory of independent particle scattering is only a special case of light electrodynamics in the Born approximation, slightly complicated by the "simplification" of having a dilute suspension of discernable isolated particles. Thus there is no need to assume independent scattering. However, if one wishes, then one can imagine the IOR as being filled with independently scattering quasi-particles, whose scattering cross-section takes into account the correlations of the elementary scatterers with their neighbors. The size L_q of the quasi-particles is the correlation length λ_χ of the susceptibility fluctuations. Such quasi-particles are stochastic ghosts, reflecting the power spectrum (from Latin *spectrum*, apparition, ghost) of susceptibility fluctuations. They exist only on average, much like laser speckles formed by scattering of coherent light on rough surfaces. Because a quasi-particle is only an average ghost, we place it in an average position in the center of IOR. If we accept the existence of quasi-particles, then one would perhaps like to define their effective concentration $\bar{C}$. We refrain from such an attempt, since quasi-particles are only ghosts, and since their concentration cancels in the final result. The measurable quantities are the scattering coefficient μ_s and the phase function $p(\theta, \phi)$.

7.3.3 Scattering Matrices of Fluctuating Continuum

The scattering transition matrix $\mathbf{T}^{\mathbf{I}}_{\mathbf{lo}}(t)$ of a particle suspension or of a sample of biological tissue is a randomly fluctuating quantity, not well suited for practical use in, say, a polarized Monte Carlo simulation. However, thanks to the Born

approximation $\mathbf{T}^{\mathbf{I}}_{\mathbf{lo}}(t)$ is the product of a scalar complex amplitude $\mathcal{A}(t)$ and a polarization matrix $\mathbf{\Pi}^d$, which only concerns the polarization transfer:

$$\mathbf{T}^{\mathbf{I}}_{\mathbf{lo}}(t) = e^{ik_m L_{if}}\, e^{-\mu_s L_{if}/2}\, \mathcal{A}(t)\mathbf{\Pi}^d \tag{7.114}$$

where

$$\mathcal{A}(t) = \frac{i\,k_m}{2} \int X_o(\mathbf{r}) X_l(\mathbf{r})\, \delta\chi(\mathbf{r},t)\, e^{i\,\mathrm{q}_m\cdot r} d^3\mathrm{r} \qquad \mathbf{\Pi}^d = \begin{pmatrix} 1 & 0 \\ 0 & \cos(\theta) \end{pmatrix} \tag{7.115}$$

Fluctuations of $\mathcal{A}(t)$ concern only the angular distribution of the scattered light. The polarization state, that is operated on by the constant polarization matrix, does not fluctuate. In the Born approximation, the *single scattering process does not depolarize*; a pure polarization state is scattered into another pure polarization state. Therefore, in the practice of polarized Monte Carlo simulations we are allowed to employ the Jones formalism. Since we shall neglect interferences between different photon paths, it is sufficient to use the time or ensemble pre-averaged version of the transition matrix:

$$\overline{\mathcal{T}}^I = i\sqrt{\frac{\mu_s V_{IOR}}{\pi a_l^2}\Omega_o}\, i e^{ik_m L_{if}}\, e^{-\mu_s L_{if}/2} \left\{ \sqrt{\frac{F(\mathbf{q})}{\Phi}} \sqrt{\frac{3}{8\pi}} \begin{pmatrix} 1 & 0 \\ 0 & \cos(\theta) \end{pmatrix} \right\} \tag{7.116}$$

where $F(\mathbf{q})$ is an experimentally determined or modeled structure factor and Φ is the well known normalization constant

$$\Phi = \int F(\mathbf{q})\, p_d(\theta,\phi) \sin(\theta) d\theta d\phi \tag{7.117}$$

Recall that if the structure is statistically isotropic, then $F(\mathbf{q})$ depends only on $q = |\mathbf{q}| = k\sqrt{2 - 2\cos(\theta)}$. This is a welcome simplification, but structural isotropy can't be a priori expected in biological tissues. Most tissues exhibit both structural anisotropy and anisotropy of dielectric response. We leave this difficult case for future work.

The quantity in large curly brackets can be recognized as the normalized S-matrix $\hat{\mathbf{S}}$ that was defined in Eq. (7.34) as $\hat{\mathbf{S}} = \mathbf{S}/\sqrt{k^2\sigma_t}$. Instead of the total scattering cross-section σ_t we have now the density of the scattering cross-section μ_s, but still, the phase function can be calculated as

$$p(\theta,\phi) = \frac{\langle E_l^I \left| \hat{\mathbf{S}}^\dagger \hat{\mathbf{S}} \right| E_l^I \rangle}{\langle E_l^I | E_l^I \rangle} = \frac{\langle E_l^L \left| \mathbf{R}_2^{-1} \hat{\mathbf{S}}^\dagger \hat{\mathbf{S}} \mathbf{R}_2 \right| E_l^L \rangle}{\langle E_l^L | E_l^L \rangle} \tag{7.118}$$

If one wishes, one can insert into the curly bracket the normalized S-matrix of a Mie-sphere. Then one would interpret the scattering coefficient as $\mu_s = \sigma_t \bar{c}$, where

σ_t is the scattering cross-section and $\bar{c}$ the number density of the spheres. Everything is allowed in the world of "quasi-particles."

We leave it to the reader to verify that $\mathbf{T}^{\mathbf{I}}_{\mathbf{lo}}(t)$ from Eq. (7.114) and $\bar{\mathcal{T}}^{I}$ from Eq. (7.116) result in the same Mueller matrix $\mathbf{M}$. In this sense $\overline{\mathcal{T}}^{I} \equiv \sqrt{\mathbf{M}(\theta,\phi)}$. Such behavior can be expected for all systems with isotropic dielectric response, in particular for isotropic Mie spheres. Keep in mind, however, that the concept of the pre-averaged Jones matrix is not universally applicable. It would fail if the scattering system consists in an ensemble of anisotropic scatterers with randomly fluctuating orientation of the anisotropy axis $\hat{\mathbf{a}}(t)$. In this case, depolarization would occur and we would have to resort to the Mueller formalism. This is not because "Jones formalism can't treat unpolarized light," but because "a Mueller matrix is equivalent to an ensemble of Jones matrices" [20]. Unfortunately, single scattering depolarization makes the treatment of anisotropy in multiple scattering rather difficult. To include anisotropy, we need to know the polarization vector $\mathbf{e}$, but Mueller formalism yields only averages of products of its complex components. This difficult case of anisotropic media with fluctuating anisotropy axis is discussed in Ref. [21].

7.3.4 Modeling the Tissue Structure Factor

One way to approach tissue optics would be to construct the susceptibility correlation function $\gamma_\delta(\mathbf{r},\mathbf{r}')$ and the corresponding structure factor $F(\mathbf{q})$ from scratch, that is by ab initio modeling of the tissue structure. Today, in the computer age, the task is not impossible. Attempts for such ab initio modeling are reported in [22–24]. However, one must realize that the modeling task is exactly the same as physically constructing tissue. We let nature do the job and rely on empirical evidence. For a long time the work horse of tissue optics was the scalar radiative transfer (see Chapters 3, 5 and 6). With scalar light one does not distinguish between the structure factor and the phase function. The most commonly employed model is still the Henyey-Greenstein phase function [25]:

$$p(\hat{\mathbf{s}}_\mathbf{i},\hat{\mathbf{s}}_\mathbf{o}) = p(\theta) = \frac{1-g^2}{4\pi}\frac{1}{\left[1+g^2-2\,g\cos(\theta)\right]^{3/2}} \quad \text{where} \quad g = \langle\cos(\theta)\,\rangle \tag{7.119}$$

g is the asymmetry parameter. Since its introduction into tissue optics [26], the Henyey-Greenstein phase function describes the angular dependence of single scattering from biological tissue reasonably well, which is somewhat surprising since this origin of phase function is associated with "Diffuse radiation in the galaxy" [25]. In galactical as well as in tissue optics, the HG phase function is usually regarded as an empirical approximation for Mie scattering from independent particles with a distribution of sizes. This interpretation makes sense in the former branch of optics, but is questionable in the latter. Nevertheless, even in tissues HG remains a good and well accepted approximation.

A new era of tissue optics was initiated by Schmitt and Kummar in 1996 [19]. Using a phase contrast microscope, they recorded the spatial variations of the optical thickness of tissue samples (geometrical thickness 5 μm) and analyzed those variations in terms of two-dimensional spatial power spectra. They found the data to be well modeled with an expression used for modeling the light propagation through turbulent atmosphere:

$$\tilde{\gamma}_2(q) \sim \frac{1}{\left[1 + q^2 L^2\right]^m} \tag{7.120}$$

The exponents m found in four different types of tissues lay in a relatively narrow range 1.30–1.45 and cut-off length L_o is roughly 1 μm. (Since in [19] spatial frequency is defined as $\kappa = q/2\pi$, the L_O values in Table 1 in [19] must be divided by 2π.) At large q, so that $qL_O > 1$, the data are well described by a nice power law $q^{-2m}(2m \approx 2.5 - 2.9)$ that applies over more than three decades of the decay of $\bar{\gamma}_2\,(q)$. Similar results obtained with a similar technique are published in [27]; one finds somewhat lower exponents and less nice power laws, but the general picture is the same. It is not quite obvious if such two-dimensional data, obtained with sample thickness of the same order of magnitude as L_O, are indeed representative of the three dimensional distribution of susceptibility fluctuations. However, an independent confirmation of the power law behavior is provided by Popp et al. [28].

These authors employed a light scattering microscope to measure directly the angular distribution of scattering intensity from thin tissue layers. Their data shown in Fig. 7.10 indicate power law exponents $2m$ in the range 2.3–2.7. (An exception is skin: the exponent $2m = 3.6$ approaches the exponent 4 of Porod's law.)

Power law decays of power spectra, especially fractional power laws, are suggestive of fractals. The fractal idea triggered an intense burst of activities in optical modeling [7, 23, 24, 29–32], but such modeling is beyond the scope of the present work. Fractals are a matter of taste. One can as well imagine an average structure

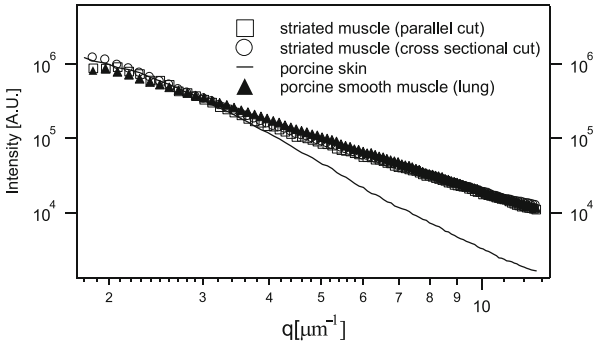


Fig. 7.10 Form factors of different tissue types exhibit power law decays. Adapted from [28] with authors' permission

(quasi particle) resulting from a polydisperse mixture of elementary shapes, rods, spheres, discs ... Consider for example the HG phase function. When translated into the present scattering language, the HG power spectrum reads

$$\tilde{\gamma}(q) \sim \frac{1}{\left[1+q^2L^2\right]^{3/2}} \quad \text{where} \quad L^2 = \frac{1}{k^2}\frac{2\,g}{(1-g)^2} \tag{7.121}$$

where L is the size of the quasi-particle, or the correlation length. Note that the asymptotic power law exponent of 3 is not too far from the experimental values 2.5–2.9. This may explain the long lasting success of the HG phase function. As already mentioned, the q^{-3} power spectrum can be interpreted as a border line fractal, either an extreme mass fractal or an extreme surface fractal. However, the same power spectrum can be and has been interpreted as resulting from polydisperse Mie spheres with broad size distribution. (One can assume a power law [30].) Therefore, instead of promoting the fractal model, we content ourselves with a practical proposal. In order to acknowledge the new developments in the practice of radiation transfer in tissues, one should include the "fractal" exponent as a parameter in the scattering law (i.e. in the structure factor). A simple but rather general and mathematically tractable model for the *isotropic* case is easily found, if one realizes that $q\bar{\gamma}\,(q)$ and $\boldsymbol{\rho}\gamma\,(\boldsymbol{\rho})$ are a pair of Fourier sinus-transforms. When skimming through the corresponding table in Gradshteyn Ryzhik [33], one immediately stumbles over the following formula:

$$\tilde{\gamma}(q) = \frac{L^3}{\left[1+(Lq)^2\right]^{\nu+3/2}} \quad \text{where} \quad \nu > -1 \tag{7.122}$$

The appearance of the exponent 3/2 is of course reminiscent of HG, which is recovered when $\nu = 0$. The extreme $\nu = -1$ would correspond to a 1D structure, recall the form factor of rods, Eq. (7.96). The corresponding correlation function reads:

$$\gamma(\boldsymbol{\rho}) = \frac{\sqrt{\pi}}{2^{\nu+1}}\frac{(\boldsymbol{\rho}/L)^{\nu}}{\Gamma(\nu+3/2)}K_{\nu}(\boldsymbol{\rho}/L) \tag{7.123}$$

Note that for $\nu \leq 0$ the correlation function diverges at $\boldsymbol{\rho} = \mathbf{0}$. In other words, the variance of susceptibility fluctuation $\langle \delta\chi^2 \rangle$ appears to be infinite. This is because we did not include the short scale cut-off l_m, namely the finite size of the molecular constituents. This could be mended by convoluting Eq. (7.123) with a suitable molecular profile, for example with a Gaussian $\exp\left(-\boldsymbol{\rho}^2/l_m^2\right)$, which would then multiply $\bar{\gamma}\,(q)$ by the Gaussian $\exp\left(-q^2/l_m^2\right)$. Not much would change because $l_m \ll L$; the molecular length scale is not observable in light scattering experiments. Recall that the scattering intensity is not proportional to $\gamma(0) \approx \langle \delta\chi^2 \rangle$ but to $\int \gamma(\boldsymbol{\rho})d^3\rho = \int \langle \delta_\chi(\mathbf{0})\delta_\chi(\mathbf{r})\rangle d^3\rho$, which is finite. A more detailed discussion of the properties of the correlation function (7.123) can be found in a recent paper by

Sheppard [32], with whom we share the practical proposal, without stressing the fractal issue.

Perhaps even more important than fractalization of the tissue micro-optics is a good model of anisotropy, which is present in most kinds of tissue. Two kinds of anisotropy can be expected to play a role. We have already encountered the anisotropy of dielectric response, which is associated with birefringence. Dielectric anisotropy affects the dipole S-matrix, as already discussed in Section 7.1.6. The second kind of anisotropy is the anisotropy of structure, leading to anisotropic correlation function and thus also to anisotropic light propagation (not to be confused with the θ-asymmetry of the phase function). A striking example is found in [28]. Usually, the structural and dielectric anisotropy are coupled. Both kinds of anisotropy can be incorporated in MC simulations, but this is a rather tiring exercise in geometry.

Remember, finally, that in order to apply the present version of the continuum scattering theory, one must be able to find a region of interest that is much larger than the correlation length $\lambda_C \approx L$, but much smaller that the mean free scattering length $\ell = 1/\mu_s$. Thus, when determining the structure factor experimentally, one should always verify the existence of the single scattering limit. This is done by measuring with tissue slices with decreasing thickness d. As long as $d > \ell = 1/\mu_s$, the measurement is polluted by multiple scattering so that the parameter values, say L and ν, are not reliable. Thus, one takes a thinner slice and repeats the experiment, recording the new values. If the parameters do not change, then we are in the single scattering limit. If there is change, we repeat the procedure once again, and so on. Does the procedure converge? If not, then we are forced to revise the theory.

7.4 Multiple Scattering

In tissue optics, the "region of interest" is not a microscopic volume of the size of few micrometers, but a macroscopic sample that is much larger than the scattering mean free path $\ell = 1/\mu_s$. We must take multiple scattering into account. The rigorous way would be to proceed to a second order approximation, which means adding to the laser field $\mathbf{E}^I$ the first time scattered field $\mathbf{E}^{(1)}$, generated by interaction of $\mathbf{E}^I$ with the susceptibility fluctuations $\delta\chi$. This is the second step in the following hierarchy:

$$\delta\mathbf{P}^{(1)}(\mathbf{r},t) = \epsilon\delta\chi(\mathbf{r},t)\mathbf{E}^I(\mathbf{r}) \quad \delta\mathbf{P}^{(1)}(\mathbf{r},t) \to \mathbf{E}^{(1)}(\mathbf{r},t) \tag{7.124}$$

$$\delta\mathbf{P}^{(2)}(\mathbf{r},t) = \epsilon\delta\chi(\mathbf{r},t)\left[\mathbf{E}^I(\mathbf{r}) + \mathbf{E}^{(1)}(\mathbf{r},t)\right] \quad \delta\mathbf{P}^{(2)}(\mathbf{r},t) \to \mathbf{E}^{(2)}(\mathbf{r},t) \tag{7.125}$$

In the first order approximation, we were concerned with the fluctuations of the dielectric polarization density $\delta\mathbf{P}^{(1)}(\mathbf{r},t)$ as the result from interaction of the laser field $\mathbf{E}^I(\mathbf{r})$ (as it propagates in homogeneous media of susceptibility $\bar{\chi}$) with the susceptibility fluctuations $\delta\chi(\mathbf{r},t)$. This density of oscillating dipole moments corresponded to the current density distribution $j(\mathbf{r},t) = -i\omega\delta\mathbf{P}^{(1)}$, which was inserted

into the Eq. (4.122) to calculate the amplitude $\mathcal{E}^o$ of the beam wave selected by the receiver from the scattered field.

Now we postpone this step and instead use $\delta\mathbf{P}^{(1)}(\mathbf{r},t)$ to calculate the first order scattered field $\mathbf{E}^{(1)}(\mathbf{r},t)$ itself. This $\mathbf{E}^{(1)}(\mathbf{r},t)$ is a superposition of spherical waves radiating from the dipoles that constitute the distribution $\delta\mathbf{P}^{(1)}(\mathbf{r},t)$. Those dipole fields consist of near- and radiation-contributions, as given in Eqs. (4.42) and (4.43). Since we keep assuming that $k\ell = k/\mu_s \gg 1$, we can neglect the near-field, but the expression for $\mathbf{E}^{(1)}(\mathbf{r},t)$ still looks difficult, even in the simplest case of isotropic dielectric response:

$$\mathbf{E}^{(1)}(\mathbf{r},t) = \frac{1}{4\pi}k^3 \int_V \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{k|\mathbf{r}-\mathbf{r}'|}\left[\mathbf{E}^{(l)}(\mathbf{r}') - \frac{(\mathbf{r}-\mathbf{r}')(\mathbf{r}-\mathbf{r}')\cdot\mathbf{E}^{(l)}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|^2}\right]\delta\chi(\mathbf{r}')\,d^3\mathbf{r}' \tag{7.126}$$

Insert $\mathbf{E}^{(1)}$ into Eq. (7.125), and the resulting $\delta\mathbf{P}^{(2)}(\mathbf{r},t)$ in turn into Eq. (4.122), and then the task would be to re-do the work we have done so far. Thereby we would be facing rather ugly monsters, namely higher order fluctuation correlation functions, up to $\langle\delta\chi(\mathbf{r}_1)\delta\chi(\mathbf{r}_2)\delta\chi(\mathbf{r}_3)\delta\chi(\mathbf{r}_4)\rangle$. In view of such an outlook, we immediately interrupt the rigorous procedure. Since in our case $\ell = 1/\mu_s$ is much smaller than the size of the sample, it would be necessary to iterate the procedure a great many times to achieve a realistic result. The iterative scheme may be straightforward, but the convergence is very slow and the procedure only works in few special cases [34]. The main problem is the infinite range of the radiation interaction: all involved integrals must be evaluated over the macroscopically large volume of the sample. An elegant, efficient and successful, yet rather haphazard, solution to this problem is offered by two seemingly disparate but de facto identical "theories": radiative transfer theory that describes the photon exchange in terms of integro-differential equations, and Monte Carlo modeling that simulates the energy transport as a random flight of photons.

7.4.1 Radiative Transfer

Radiative transfer relies heavily on the particle picture of light. One forgets about electromagnetic waves and regards photons as localized point-like particles that propagate along straight lines that are called *rays* in the context of geometrical optics. Occasionally, a photon bounces with a material particle and changes the direction of the flight. The only role of the wave optics is to provide rules for such photon-particle encounters, namely the scattering cross-section σ_t and the phase function $p(\theta, \phi)$. The remaining part of the radiative transfer theory consists in the accurate accounting of rays and photon fluxes, under the strict observation of the law of energy conservation. The presumptions of Radiative Transfer can be formulated as follows: (i) Light fields can be replaced by a manifold of geometrical rays. (ii) Nevertheless, scattering and absorption processes can be characterized by Beer's law and by the *pre-averaged* phase function $p\,(\theta, \phi)$ (recall Section 7.3.3). (iii) Light propagating along different rays does not interfere.

We shall now investigate if and how these presumptions are compatible with the electromagnetic theory of light. The key idea needed to justify the ray-presumption is to realize that a *ray* is the limiting case of a *beam-wave* (recall Section 4.2.3.1). A ray is a beam wave whose effective radius a_b and effective solid angle $\Omega_b = 4\pi \big/ k^2 a_b^2$ have shrunk to zero, while the Rayleigh range $z_r = k a_b^2$ extends to infinity. The only way to reach this limit without violating the uncertainty principle is to shrink the wavelength λ to zero. In this sense, geometrical optics is the short wavelength limit of wave optics [8]. Geometrical ray optics precludes all kind of interference. Fortunately, however, the transition into the geometrical limit is unnecessary. A beam appears as a ray for all practical purposes if the following conditions are fulfilled: (i) The Rayleigh range z_r (recall Fig. 4.2) is much larger than the size L of the region of interest. (ii) The effective radius a_b, that eventually determines the resolution, is much smaller than L. These conditions can be quite easily fulfilled with visible light in stellar and planetary atmospheres: a beam of a radius $a_b = 1$ and Rayleigh range $z_r = 6 \times 10^9$ km (at $\lambda = 1\ \mu$m) will certainly appear like a ray when looked at from Proxima Centauri.

Having accepted beams in place of geometrical rays, we are able to analyze the scattering processes in a rigorous wave-optical way, as we have done in previous sections. Thereby, one realizes that the conditions on the "beam-rays" that are needed to justify presumption II, are actually more strict: we must require that $a_b \gg \lambda_C$, where λ_C is the structural correlation length of the scattering system. On the other hand, in order to discern *single scattering events* we must require that $a_b \ll \ell_e$, where ℓ_e is the extinction length of Beer's law. Fortunately, we may somewhat release the condition on the Rayleigh range: it is sufficient to require that $z_r \gg \ell_e$. All this is certainly well fulfilled in the optics of atmospheres, but less clear is the situation in biological tissues (recall the concluding remarks in Section 7.3.4).

The third presumption of Radiative Transfer concerns the absence of interference of light waves propagating in different "beam-rays." To justify this presumption, we must first reflect on the meaning of the phrase "absence of interference." Recall from Section 4.2.3.1 that any propagating light field can be expressed as the superposition of plane waves. Thereby, the total power carried by the light field is the sum of the partial powers carried by the individual partial waves.[3] Expressed in the language of mathematics, plane waves represent a complete set of *orthogonal* base vectors in the space of the solutions of Maxwell Equations. The construction of a beam wave, Eq. (4.29), is just one example. Since a beam-wave in Rayleigh range is a quasi-plane wave, one can expect that a propagating light field can be as well constructed as a superposition of orthogonal unit beam waves, at least locally and in a good approximation. Recalling the notion of the unit beam wave from Section 4.5.2, we write the resulting field as

$$|E\rangle = \sum_b |b\rangle\, \mathcal{E}_b \tag{7.127}$$

[3] Since plane waves are delocalized, their individual power contributions are infinitesimally small.

Since the orthogonal modes do not interfere, the power $P_E = \langle E \mid E \rangle$ is the sum of powers carried in the individual modes:

$$P_E = \sum_b |\mathcal{E}_b|^2 \tag{7.128}$$

Note that "absence of interference" does not mean that the light fields of the partial modes would not interfere *locally*, forming fringes or speckles wherever they overlap. The term "absence of interference" merely asserts the validity of Eq. (7.128): there are no *cross-terms*, i.e. no *interference* between the orthogonal modes. Note that the term "orthogonal" does not only imply the orthogonality of the polarization vectors $\hat{\mathbf{e}}_b$ or the propagation vectors $\hat{\mathbf{s}}_b$ of the base modes. Two beam waves $|b_1\rangle$ and $|b_2\rangle$ are orthogonal if $\langle b_1 \mid b_2 \rangle = 0$. The scalar product $\langle b_1 \mid b_2 \rangle$ is defined in Eq. (4.120). Using the paraxial approximation, Eq. (4.121), one can verify that two Gaussian beam waves $|b_1\rangle$ and $|b_2\rangle$ are nearly orthogonal as soon as the angle between their propagation vectors $\hat{\mathbf{s}}_1$ and $\hat{\mathbf{s}}_2$ exceeds a couple of the beams divergence angles $\beta = 1/ka_b$. Two beams are also nearly orthogonal as soon as the shortest distance between their central rays exceeds a couple of effective beam radii a_b. (Gaussian beams may be not a best set of orthogonal base vectors, but sufficient for the present plausibility argument.) The sketch in Fig. 7.11a illustrates a selected few of orthogonal "beam-rays" crossing a certain surface Π.

Note that each of the beams exhibits a small yet finite effective solid angle Ω_b, which is not easy to visualize because Ω_b is small and therefore the Rayleigh range z_r rather long.

Having understood the concept of the superposition of orthogonal modes, we are in the position of making the transition to Radiative Transfer plausible. Consider a planar surface of an area A that is much larger than the mode area πa_b^2. This plane is irradiated by the field $|E\rangle$, that is a superposition of a large number of the orthogonal beam waves. For simplicity we assume that all modes have the same effective cross-section πa_b^2 and the same effective solid angle Ω_b. The modes are distinguished by two parameters: the intersection $\boldsymbol{\rho}_b$ of their principal ray with the plane Π and the propagation vector $\hat{\mathbf{s}}_b$. If one could record the irradiance in the Π plane,

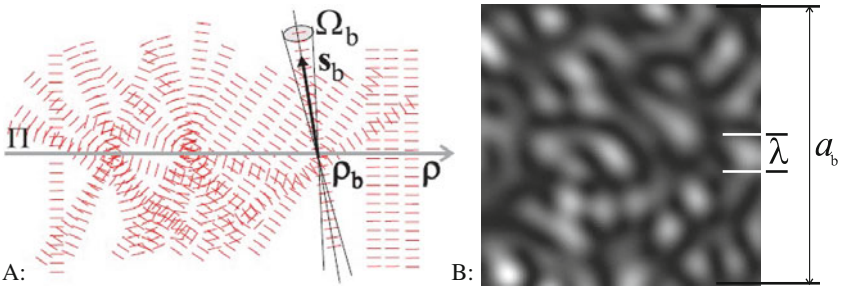


Fig. 7.11 (**a**) Sketch of beam-waves passing through a plane Π. (**b**) Speckle pattern that results from a superposition of many such partial waves

one would observe a speckled pattern, illustrated in Fig. 7.11b. This pattern reflects *local*[4] interference effects resulting from the superposition of the beam waves with a broad distribution of propagation vectors $\hat{\mathbf{s}}_b$. Since this distribution extends over the entire 4π solid angle, the average size of a speckle corresponds roughly to one wavelength λ. In Radiative Transfer we are not interested in such details, but only in length scales that are larger than a_b. Thus, we pre-average the pattern over an area of a couple of πa_b^2. Because of the orthogonality of the involved beam-waves, the pre-averaged irradiance is the sum of the irradiances of the individual beams as they are crossing Π. Each beam with a propagation vector $\hat{\mathbf{s}}_b$ carries the power P_b that is distributed over a small effective area $A_b = \pi a_b^2/(\hat{\mathbf{s}}_b \cdot \hat{\mathbf{n}})$, where $\hat{\mathbf{n}}$ is the surface normal. The mean irradiance within the area A_b is $\bar{I}_b = P_b/A_b = (P_b/\pi a_b^2)\,\hat{\mathbf{s}}_b \cdot \hat{\mathbf{n}}$. Each mode radiates into an effective solid angle $\Omega_b = \lambda^2/\pi a_b^2$. To take this into account, we express the *irradiance* $\bar{I}_b$ [W/m^2] in terms of the *radiance* L_b [W/m^2/sr] of the beam:

$$\bar{I}_b = L_b \Omega_b \quad \text{where} \quad L_b = \frac{\bar{I}_b}{\Omega_b} = \frac{P_b}{\lambda^2} \tag{7.129}$$

In the world of beam waves, this is just a harmless rearranging trick, which allows us to write the power transmitted across the surface Π as

$$P = \sum_{\boldsymbol{\rho}_b} \sum_{\hat{\mathbf{s}}_b} L_b(\boldsymbol{\rho}_b, \hat{\mathbf{s}}_b)\, \hat{\mathbf{s}}_b \cdot \hat{\mathbf{n}}\ \Omega_b\, \pi a_b^2 \tag{7.130}$$

In the world of rays, Ω_b and πa_b^2 are regarded as infinitesimal quantities. Thus, we replace the summation by integration:

$$P = \int_{\boldsymbol{\rho}} \int_{\hat{\mathbf{s}}} L(\boldsymbol{\rho}, \hat{\mathbf{s}})\, \hat{\mathbf{s}} \cdot \hat{\mathbf{n}}\ \delta\Omega\, \delta\rho^2 \tag{7.131}$$

Recall that in electrodynamics the transmission of power across a surface Π is expressed as $P = \int \mathbf{S} \cdot \hat{\mathbf{n}}\, d^2\rho = \int S\hat{\mathbf{s}} \cdot \hat{\mathbf{n}}\, d^2\rho = v \int u\, \hat{\mathbf{s}} \cdot \hat{\mathbf{n}}\, \delta\rho^2$, where $\mathbf{S}$ is the local Poynting vector, u is the local energy density and v is the speed of energy transport. By the mode pre-averaging we arrived at the ray version of this relation, which can be expressed as

$$P = \int \mathbf{F} \cdot \hat{\mathbf{n}}\, d^2p = \int \phi \bar{\mathbf{s}} \cdot \hat{\mathbf{n}}\, d^2p = v \int u\, \bar{\mathbf{s}} \cdot \hat{\mathbf{n}}\, \delta\rho^2 \tag{7.132}$$

where

$$\mathbf{F}(\mathbf{r}, t) = \int_{4\pi} L(\mathbf{r}, t, \hat{\mathbf{s}})\hat{\mathbf{s}} d\Omega, \quad \phi(\mathbf{r}, t) = \int_{4\pi} L(\mathbf{r}, \mathbf{t}, \hat{\mathbf{s}}) d\Omega, \quad u = \frac{\phi(\mathbf{r}, t)}{v} \tag{7.133}$$

[4] In Section 7.5 we shall also reflect on *global* interferences of photon paths. Those are not apparent here, since Radiative Transfer is a *local* theory.

Here $\mathbf{F}(\mathbf{r},t)$ is the pre-averaged *energy flux density* and $\phi(\mathbf{r}, t)$ is the so-called *fluence rate*. In Eq. (7.132) $\bar{\mathbf{s}} = \mathbf{F}(\mathbf{r}, t)/\phi(\mathbf{r}, t)$ is the mean propagation vector. What remains to establish the Radiative Transfer is to invoke energy conservation, that can be expressed through the Equation of Continuity:

$$\begin{aligned} &\frac{\partial \bar{u}}{\partial t} = -\nabla \cdot \mathbf{F} - \text{sinks} + \text{sources} \\ &\frac{1}{v}\int_{4\pi} \frac{\partial L(\mathbf{r},\hat{\mathbf{s}})}{\partial t} d\Omega = -\int_{4\pi} \hat{\mathbf{s}} \cdot \nabla L(\mathbf{r}, t, \hat{\mathbf{s}}) d\Omega - \text{sinks} + \text{sources} \end{aligned} \tag{7.134}$$

The last step is to recall that the symbol $L(\mathbf{r}, \hat{\mathbf{s}})$ actually represents a beam wave passing through $\mathbf{r}$ in the direction $\hat{\mathbf{s}}$. During the propagation, energy is removed from each beam wave by absorption and scattering, as expressed by Beer's law with the extinction coefficient $\mu_e = \mu_a + \mu_s$. However, there is also energy input by scattering from the crossing beams from a direction $\hat{\mathbf{s}}'$ into the direction $\hat{\mathbf{s}}$. This applies independently for all beams, and therefore we can remove the Ω-integration in Eq. (7.134). Thus we write:

$$\frac{1}{v}\frac{\partial L(\mathbf{r},\hat{\mathbf{s}})}{\partial t} + \frac{\partial L(\mathbf{r},\hat{\mathbf{s}})}{\partial s} = -\mu_e L(\mathbf{r},\hat{\mathbf{s}}) + \mu_s \int_{4\pi} p(\hat{\mathbf{s}}|\hat{\mathbf{s}}')\, L(\mathbf{r},\hat{\mathbf{s}}') d\Omega' + S_L \tag{7.135}$$

Here we employed the identity $\hat{\mathbf{s}} \cdot \nabla L(\mathbf{r},\hat{\mathbf{s}}) = \partial L(\mathbf{r},\hat{\mathbf{s}})/\partial s$. The source term S_L is the reminder of the laser injecting energy somewhere into the system. Equation (7.135) is the *Equation of Radiative Transfer* in the time dependent version. Note that it is quite similar to Boltzmann's Equation for the transport of the molecules in gas, simplified by the fact that all photons propagate with the same speed. The physical meaning of Eq. (7.135) can be easiest understood in the stationary case, when $\partial L/\partial t = 0$.

$$\frac{\partial L(\mathbf{r},\hat{\mathbf{s}})}{\partial s} = -\mu_e L(\mathbf{r},\hat{\mathbf{s}}) + \mu_s \int_{4\pi} p(\hat{\mathbf{s}}|\hat{\mathbf{s}}')\, L(r,\hat{\mathbf{s}}') d\Omega' + S_L \tag{7.136}$$

On the right hand side, the first term represents Beer's law, i.e. energy removal from the beam wave over the infinitesimal length ∂s. The second term accounts for photons that have been scattered within ∂s from all other modes $\hat{\mathbf{s}}'$ into the state $\hat{\mathbf{s}}$. Recall also the laser energy input S_L.

The only difference between the scalar transport theory and its polarized version is that not only the propagation vector $\hat{\mathbf{s}}$, but also the polarization vector $\hat{\mathbf{e}}$, specify the photon state. Correspondingly, the "rules for photon-particle encounters" are a little bit more complex. In the scalar theory, scattering rules are contained in the phase function $p\left(\hat{\mathbf{s}}, \hat{\mathbf{s}}'\right)$ and the coefficients μ_S, μ_a. In reality, however, photon exchange depends on the propagation vectors $\hat{\mathbf{s}}$ and $\hat{\mathbf{s}}'$of the involved photons, and their polarization states $\hat{\mathbf{e}}$ and $\hat{\mathbf{e}}'$. Since the polarization modes span a 2D complex space, it is advantageous to express the polarized rules in the matrix formalism of Mueller [1] or Jones [35]. The next task is to specify the initial and boundary conditions, including the laser light source, and solve the integro-differential

Eq. (7.135). This is a formidable amount of mathematical work, and therefore we give the radiative transfer up, just like we have given up the use of rigorous multiple scattering theory. Fortunately, there is an alternative approach, much more intuitive and conceptually simple than the abstract mathematics of radiative transfer.

7.4.2 Photon Random Flight and Monte Carlo Simulations

Since we have given up using electromagnetic waves, we will represent the energy transport in the multiply scattering medium as a random flight of photons. In this picture, photons are regarded as a kind of billiard balls that fly through the scattering medium with the transport speed v, bouncing thereby randomly into some random obstacles. However naive, this picture is an exact representation of scalar radiative transfer. It is somewhat surprising that Chandrasekhar does not refer to this simple model, having contributed significantly to the theory of random flights [36]. The most likely reason for his negligence is the lack of computing power: he did not consider the possibility, because in the 1940s the only way of treating physical problems was setting up and solving equations. Today we have the means to solve the transport equation directly, that is by Monte Carlo simulations of random flight. The great feature of the MC simulation is its dual nature: it is a flexible technique for solving transport equation with arbitrary shapes of the scattering medium and arbitrary phase functions; but it is at the same time a simple, intuitive and lucid model of the physical reality. With scalar photons, the MC simulation is exceedingly simple: one launches a photon into tissue at a certain initial position $\mathbf{r}_0$ and lets it fly a distance d_0 in the direction $\hat{\mathbf{s}}_0$ until it bounces into a scatterer at $\mathbf{r}_1$. This encounter changes the flight direction to $\hat{\mathbf{s}}_1$ and the procedure is repeated. The photon pursues a highly irregular path through the medium, until it leaves the sample to vanish in space. Thereby Beer's law is implemented in a probabilistic fashion, by throwing a dice: the flight distance between two interaction events is distributed according to the exponential distribution

$$p_d(d)\delta d = \frac{1}{\ell_e} e^{-d/\ell_e}\delta d \quad d \in \{0, \infty\} \tag{7.137}$$

where $\ell_e = \langle d \rangle = 1/\mu_e$ is the interaction mean free path, $\mu_e = \mu_s + \mu_a$ is the extinction coefficient (alias total attenuation coefficient). Thus, to generate an interaction event, one generates a random number as described in Chapter 5. After the interaction, one more random number is needed to simulate the absorption: with the probability μ_s/μ_e the photon proceeds its irregular flight; with the complementary probability μ_a/μ_e it is absorbed. When absorption has occurred, one marks this event in a 3D matrix that will finally contain the energy deposition in the tissue. Otherwise the change of flight direction in the scattering process is simulated according to the chosen phase function $p\left(\hat{\mathbf{s}}, \hat{\mathbf{s}}'\right)$. Thus, one generates two random numbers from the joint distribution $p_{\theta,\phi}(\theta, \phi)\sin(\theta)\delta\theta\delta\phi$, such that $\theta \in \{0, \pi\}, \phi \in \{0, 2\pi\}$. Note that the source term $\mu_s \int_{4\pi} p(\hat{\mathbf{s}}|\hat{\mathbf{s}}')\, L(\mathbf{r}, \hat{\mathbf{s}}')d\Omega$ does not

explicitly enter MC simulations; it is automatically realized after having generated a large number of random photon paths. Much more about such MC simulations can be read in Chapter 5. Finally, we note that from random flight there is only a short step to the diffusion approximation. As shown, e.g. by Chandrasekhar [36], the diffusion process is the limit of a random flight with an infinitesimally short scattering mean free path ℓ, but with a correspondingly large number of steps n, so that L_n =const while $1/\ell = \mu_s \to \infty$. (Thereby keep μ_a finite; absorbed photons do not propagate.) Within this limit, the transfer equation, Eq. (7.135), can be replaced by the diffusion equation for the energy density $u(\mathbf{r}, t)$, as discussed in detail in Chapter 6.

7.5 Modeling Light Propagation as Polarized Random Flight

Monte Carlo simulations of disordered light propagation rely on the concept of the photon path. This path approach involves serious approximations with respect to the exact electromagnetic theory, and one may ask why Monte Carlo is so successful. A rigorous transition from electrodynamics to radiative transfer and Monte Carlo simulation is a formidable task [37], well beyond the scope of the present contribution. However, we can at least try to make the approximations involved in this transition explicit. At the same time we take the opportunity for a brief summary of the concepts and tools which we have prepared in previous sections. Unlike most polarized MC simulations previously published in the context of tissue optics [38–43], we will not use Mueller formalism, but work on the amplitude level. We thus follow the path that is also taken by researchers interested more in fundamental than in applied aspects of light propagation [44–46]. The reasons have been mentioned earlier: (i) single scattering usually does not depolarize and therefore there is no reason to employ Mueller; (ii) Jones formalism works on the level of amplitudes and is therefore nearer to the underlying physics than Mueller's averaged products of complex amplitude components; and (iii) we would like to keep the door open for future developments that would involve interferences of the photon paths.

7.5.1 Photon Path

As we decided to work at the amplitude level, we find it convenient to employ a pedestrian version of Feynman's concept of photon path. We shall follow Feynman's path rules, as given in [47], or, in the pedestrian version, in the enlightening booklet "*QED – a strange theory of light and matter*" [48]. (Feynman rules adapted for the present purpose can be found in [6], Chapter 20.) A photon path is a chain of localized scattering events connected by propagation. A photon is injected into a state $|l\rangle$ (this initial state is the impinging laser beam) and propagates to undergo its first scattering at $\mathbf{r}_1$, which transfers the photon into a new state $|\pi_1\rangle$. Then it propagates until a second scattering event occurs at $\mathbf{r}_2$, and so on, from $\mathbf{r}_{i-1}$ to $\mathbf{r}_i$,

until the photon exits the scattering medium in a state $\left|\pi_p\right\rangle$ or is absorbed. There is a certain chance that an exiting photon ends up in a detector. In Feynman's spirit, we associate each path with a *path amplitude* $\mathcal{F}_p$ for the transition from an initial state $|l\rangle$ to a final state $\left|\pi_p\right\rangle$:

$$\left|\pi_p\right\rangle = \mathcal{F}_p \, |l\rangle \tag{7.138}$$

In our case, the states $\left|\pi_p\right\rangle$ are polarized beam waves characterized by a polarization vector $\hat{\mathbf{e}}_\mathbf{p}$, propagation vector $\hat{\mathbf{s}}_\mathbf{p}$ and a scalar complex amplitude ε_p. Working with a succession of transverse coordinate systems, the path amplitude $\mathcal{F}_p$ can be represented by a 2×2 Jones matrix and the states $\left|\pi_p\right\rangle$ or $|l\rangle$ as Jones vectors. However, Feynman's approach is quite general, and therefore we keep using the general bra ket notation.

The *path probability* is the probability that a photon injected into the state $|l\rangle$ will make it into the state $\left|\pi_p\right\rangle$. The path probability is defined as:

$$\mathcal{P}_p = \frac{\langle \pi_p | \pi_p \rangle}{\langle l | l \rangle} = \frac{\langle l \left| \mathcal{F}_p^\dagger \mathcal{F}_p \right| l \rangle}{\langle l | l \rangle} = \langle l \left| \mathcal{F}_p^\dagger \mathcal{F}_p \right| l \rangle \tag{7.139}$$

On the right hand side we recall that the initial state is normalized, i.e. $\langle l \mid l \rangle = 1$. The path amplitude is the product of segment amplitudes, i.e. the photon state after the last scattering event is represented by

$$\left|\pi_p\right\rangle = \mathcal{F}_p \, |l\rangle = \Pi_{i=1}^{np} \mathcal{F}_i \, |l\rangle \tag{7.140}$$

where the segment amplitude $\mathcal{F}_i$ is the amplitude for the transition of a photon from a state $|\pi_{i-1}\rangle$ to a state $|\pi_i\rangle$, such that $|\pi_i\rangle = \mathcal{F}_i \, |\pi_{i-1}\rangle$. Obviously, the path probability $\mathcal{P}_p$ can be expanded into a product of *segment probabilities*, $\mathcal{P}_i$. We only indicate the first step of the expansion:

$$\mathcal{P}_p = \frac{\langle \pi_p | \pi_p \rangle}{\langle l | l \rangle} = \frac{\langle \pi_{p-1} | \mathcal{F}_p | \pi_{p-1} \rangle}{\langle \pi_{p-1} | \pi_{p-1} \rangle} \frac{\langle \pi_{p-1} | \pi_{p-1} \rangle}{\langle l | l \rangle} = ... = \Pi_{i=1}^{np} \mathcal{P}_i \tag{7.141}$$

Here $\mathcal{P}_i$ is the probability that a photon that entered the photon state $|\pi_{i-1}\rangle$ at $\mathbf{r}_{i-1}$ will make it to $\mathbf{r}_i$ and there becomes scattered into a state $|\pi_i\rangle$:

$$\mathcal{P}_i = \frac{\langle \pi_i | \, \pi_i \rangle}{\langle \pi_{i-1} | \, \pi_{i-1} \rangle} = \frac{\langle \pi_{i-1} | \, \mathcal{F}_i^\dagger \mathcal{F} \, | \pi_{i-1} \rangle}{\langle \pi_{i-1} | \, \pi_{i-1} \rangle} \tag{7.142}$$

Notice that in the present context "probability" is to be understood in the quantum sense, as the complex square of an amplitude. We are now dealing with pure states and the fluctuations in the scattering medium are not yet considered.

A real experiment involves a receiver and a detector. We are interested in the normalized expected signal, i.e. in the detection probability $\mathcal{P}_d = N_d / N$, where N_d is the number of photons that are likely to be detected, when N photons were injected

into the system. For simplicity, we assume an ideal detector which registers with a probability 1 each photon that enters into the receiver. The receiver is a polarized single-mode beam filter (Section 4.5), which projects each of the incoming states $|\pi_p\rangle$ on the receiver mode $|o\rangle$ (recall that $\langle o \mid o\rangle = 1$). The received state $|\pi_d\rangle = |o\rangle\, \mathcal{E}_d$ is "filled" with contributions from all paths:

$$|\pi_d\rangle = |o\rangle\, \mathcal{E}_d = |o\rangle \sum_p \langle o|\mathcal{F}_p\, |l\rangle \tag{7.143}$$

The receiver mode $|o\rangle$ and the initial state $|l\rangle$ are normalized, so that we write the detection probability as

$$\mathcal{P}_d = \frac{\langle \pi_d|\pi_d\rangle}{\langle l|l\rangle} = \sum_{p=1}^{N_p}\sum_{q=1}^{N_p} \langle l|\, \mathcal{F}_p^\dagger\, |o\rangle\, \langle o|\, \mathcal{F}_q\, |l\rangle \tag{7.144}$$

Note that there are two types of terms in the double sum: The N_p terms with $p = q$ give the incoherent sum of the detection probabilities of the individual path $P_{dp} = |\, \langle o|\, \mathcal{F}_p\, |l\rangle\, |^2$. The cross-terms with $p \neq q$ express the *interference of paths*. In Section 7.4.1 we did not consider this type of interference, because in Radiative Transfer the paths are not explicitly visible; all one has is the *local* transfer equation, Eq. (7.135). However, we discussed the *local* interference that leads to the speckle pattern shown in Fig. 7.11b. The same pattern would appear on an area detector positioned in place of the single mode receiver. Integrating the signal over the detector area A, the speckles will average out: because a multimode detector receives a large number of states, one can invoke in Eq. (7.144) the closure relation $\sum_o |o\rangle\langle o| \approx \mathbf{I}$ and assume that the detected signal is the sum of probabilities of paths terminating in the detector area:

$$\mathcal{P}_A \sim \sum_{A,p=1}^{N_p} \langle l|\mathcal{F}_p^\dagger \mathcal{F}_p\, |l\rangle = \sum_{A,p=1}^{N_p} \mathcal{P}_p \tag{7.145}$$

This kind of detection is employed in the classical scalar MC simulations. However, the role of interference and fluctuations is precisely what we would like to address in future work. Therefore we use a single-mode detector and the general expression for the detected signal, Eq. (7.144). Thus, we must explicitly carry out the last step of the path analysis, namely the time averaging. All scattering media of practical interest exhibit thermal fluctuations, which keep rearranging their microscopic structure. Thus, the amplitude of a path through a fixed set of scattering locations $\{\mathbf{r}_i\}$ fluctuates in time, $\mathcal{F}_p = \mathcal{F}_p(t)$. The measured detection probability is the time average:

$$\overline{\mathcal{P}_d} = \frac{1}{T}\int_0^{\mathcal{T}} \mathcal{P}_d(t)\, dt = \sum_p \sum_q \overline{\langle l|\, \mathcal{F}_p^\dagger(t)\, |o\rangle\, \langle o|\, \mathcal{F}_q(t)\, |l\rangle} \tag{7.146}$$

Here the cross terms represent *paths correlations*. Strictly speaking, only if all paths are uncorrelated, then we are allowed to write the detected signal as the sum of the detection probabilities of the individual paths:

$$\overline{\mathcal{P}_d} \approx \sum_p \overline{|\langle o|\,\mathcal{F}_p(t)\,|l\rangle\,|^2} \tag{7.147}$$

To justify the decorrelation we must start filling the bra-kets with their contents.

7.5.2 Initial Photon State, Launching a Photon

A realistic initial photon state is a collimated or moderately focused laser beam. We model the beam as a paraxial superposition of plane waves which span a certain range of the wave vectors around a principle wave vector $\mathbf{k}_l$ and which are synchronized at the focal point $\mathbf{r}_l$:

$$|l\rangle = |\widehat{\mathbf{e}_\mathbf{l}}, \mathbf{k_l}, \mathbf{r}_l\rangle \equiv \sqrt{\frac{2}{c\epsilon_m}}\, X(\mathbf{r}; \mathbf{r}_l, \mathbf{k}_l)\widehat{\mathbf{e}_l}\, e^{i\mathbf{k_l}\cdot(\mathbf{r}-\mathbf{r}_l)} \tag{7.148}$$

The beam profile $X(\mathbf{r}, \mathbf{r}_l, \mathbf{k}_l)$ is for example a Gaussian $X(\mathbf{r}) = 1/a(z)\exp(-\rho^2/2\mathrm{a}(z)^2)$, where $a(z) = a_l\sqrt{1 + z^2/z_r^2}$ is the effective beam radius, that varies with the distance z from the beam focus, but within the Rayleigh range $z < z_r = ka_l^2$ it is approximately constant. We assume that the length z_r is much larger than the extinction length $\ell_e = 1/\mu_e$, so that we can regard the beam as a quasi plane wave in the relevant range of propagation. Taking a typical $\ell_e \approx 30\,\mu\mathrm{m}$, the condition $z_r \gg \ell_e$ is fulfilled with a beam as narrow as $a_l = 10\,\mu m$.

Into the normalized beam $|l\rangle$ we start launching photons. The squared profile $|X(\rho)|^2$ can be interpreted as the probability density for finding a photon within a certain distance ρ from the beam axis. Thus, we choose a unique initial propagation vector $\hat{\mathbf{s}}_\mathbf{0}$ and launch the photons with an offset ρ with respect to the beam axis $\hat{\mathbf{z}}$ so that ρ_x and ρ_y are random numbers from the distribution $X_l^2(\rho_x, \rho_y)$. A rigorous treatment of launching through a strongly focused laser beam or using a fiber-optic optode would require more thinking. As a first order approximation, one can adopt the approach of radiative transfer, replacing the light fields with polarized photon streams.

7.5.3 Scattered Photon States, Beams and Paths

The scattered fields are more complicated. They result from the coupling of dipole fields of a large number of elementary microscopic scattering processes. In other words, scattering involves a large number of elementary photon paths. Such a coupled dipole approach has been successfully used to calculate the scattering from

small particles [34], but it is entirely inadequate for light propagation in macroscopic disordered media. The basic idea of the MC simulation is essentially to "preprocess" the maze of elementary paths into manageable sets of path bundles. In most MC simulations one implicitly assumes that the scattering medium consists of isolated spherical particles. The scattering events are thought to be localized in their centers and a photon propagates along ray-segments connecting the individual particles. We shall keep this picture, but replace the tangible Mie-spheres with the ghost quasi-particles from Section 7.3.2.1, i.e. with fuzzily localized regions in space surrounding each scattering event. Moreover, we will replace the rays by orthogonal beam waves, in the same spirit as we have already done in Section 7.4.1, while "deriving" the Equation of Radiative Transfer. In other words, we will decompose the spherical wave radiating from a quasi-particle into a broad range of wave vectors, into a more or less complete set of orthogonal partial beam waves, as illustrated in Fig. 7.12. Each of the "beam-rays" radiates into a small solid angle $\delta\Omega$. The procedure is analogous to the rigorous expansion of a radiating dipole wave into plane waves [49], but we prefer quasi-plane beam waves, in order to preserve the localization of the path. Admittedly, the proposed kind of decomposition is rather fuzzy. Nevertheless, if the procedure succeeds, then we provide a sound electromagnetic ground to Monte Carlo simulations in multiply scattering media.

The focal point of each beam is the location $\mathbf{r}_i$ of the scattering event i, and the principle wave vector $\mathbf{k}_i$ is directed along the path segment vector $\mathbf{r}_{i+1} - \mathbf{r}_i$. These scattered beams may be thought to have a certain finite waist radius δd (equivalent to the Gaussian beam radius $\delta\Omega = 4\pi(k\delta d)^2$. In each beam we collect a bundle of elementary photon paths that are synchronized in $\mathbf{r}_i$ and that propagate in the solid angle $\delta\Omega$. In a scattering event the photon "chooses" one of such path bundles, as illustrated in Fig. 7.12. One may imagine the scattering system as being subdivided into "sample cells," formed at the intersections of the partially scattered beam waves. In between the cells the beam wave propagates in a macroscopically homogeneous effective medium, the only reminder of the microscopic inhomogeneity being the extinction of the beam wave due to scattering or absorption. The cell C_i is centered in the ith-scattering event, that is in the focal point of the beam wave $|\pi_i\rangle$. In

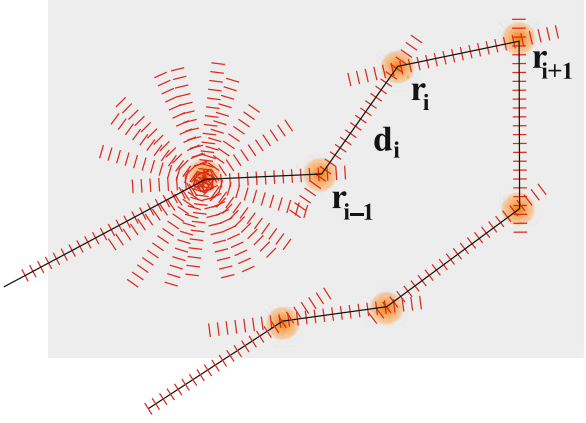


Fig. 7.12 Multiple scattering model: photon propagation through a system of paths of beam-rays connecting virtual sample-cells

each realization of path a new subset of such cells is created. The volume V_i of each cell is given by the waist radius of the scattered beam waves, i.e. $V_i \approx \pi(\delta d)^3$. As we shall see later, the exact size does not matter. However, there are certain restrictions: the waist radius δd should be larger or at least comparable with the structural correlation length λ_s, but much smaller than the mean distance ℓ between the scattering events, so that only single scattering is likely to occur in each cell. In addition, the Rayleigh range $z_r = k(\delta d)^2$ of the partial beam wave should be large compared with ℓ, i.e. $k\,(\delta d)^2 \;>\; \ell$. At this point we may summarize the requirements concerning the length scales:

$$1 \ll k\ell \quad (k\lambda_s)^2 << k\ell < (k\delta d)^2 \quad k\lambda_s << k\delta d << k\ell \tag{7.149}$$

If these relations can't be fulfilled, then our MC simulation likely fails to describe experimental reality.

7.5.4 Averages and Correlations

The last but not least fundamental approximation involved in modeling the photon random flight is path decorrelation which results from the decorrelation of individual segments. A path segment consists of the propagation of the photon from $\mathbf{r}_{i-1}$ to $\mathbf{r}_i$ and of the scattering event at $\mathbf{r}_i$. Having accepted the photon path picture in Fig. 7.12, we apply the results developed in Section 7.3.3. The segment amplitude matrix $\mathcal{F}_i$ is concisely written as

$$\mathcal{F}_i(t) = \mathcal{A}_i(t)\mathbf{\Pi}^d(\theta_i)\mathbf{R}_2(\phi_i) \cdot \mathbf{J}e^{ik_m d}e^{-\mu_e d/2} \tag{7.150}$$

In the present section we disregard the propagation that is included in the phasors and in the Jones matrix $\mathbf{J}$. We also disregard the matrix $\mathbf{\Pi}^d(\theta_i)\,\mathbf{R_2}(\phi_i)$, which concerns only the polarization. We assume, as we discussed in Section 7.3.3, that the dielectric response does not fluctuate, so that an incoming pure polarization state remains a pure state upon scattering. The only fluctuating quantity is the scalar amplitude

$$\mathcal{A}_i(t) = \frac{i\,k_m}{2} \int X_{i-1}(\mathbf{r})X_i(\mathbf{r})\,\delta\chi(\mathbf{r},t)\,e^{i\,q_i\cdot\mathbf{r}}d^3r \tag{7.151}$$

Here $\delta\chi(\mathbf{r},t)$ represents *susceptibility fluctuations*, $\mathbf{q}$ is the scattering vector and $X_{i-1}(\mathbf{r}), X_i(\mathbf{r})$ are the profiles of the incoming and outgoing beams. Note that $\overline{\mathcal{A}_i\,(t)} \;=\; 0$, since the time average of $\delta\chi(\mathbf{r},t)$ is zero by definition. To investigate the decorrelation we return to Eq. (7.146) which we rewrite explicitly in terms of the amplitudes $\mathcal{F}_{pi}$ of the i-th segment in the p-th path:

$$\overline{\mathcal{P}_d} = \sum_p \sum_q \overline{\langle l|\,\Pi_{i=1}^{np}\mathcal{F}_{pi}^{\dagger}\,|o\rangle\langle o|\,\Pi_{j=1}^{q}\mathcal{F}_{qj}\,|l\rangle} \tag{7.152}$$

The time averaging concerns the products $\overline{\mathcal{A}^{\dagger}_{pi}(\mathbf{q}_i, \mathbf{r}_i, t)\mathcal{A}_{qj}(\mathbf{q}_j, \mathbf{r}_j, t)}$. Recall the meaning of "uncorrelated": the average of the product of two independent, or *uncorrelated*, random variables can be factored into the product of their averages. The decorrelation approximation is grounded in the following fact: two scalar segment amplitudes $\mathcal{A}_{pi}\left(\mathbf{q}_i, \mathbf{r}_i, t\right)$ and $\mathcal{A}_{qj}\left(\mathbf{q}_j, \mathbf{r}_j, t\right)$ are uncorrelated unless they belong to the same scattering volume (the same quasi-particle) at $\mathbf{r}_i$ and the same scattering vector $\mathbf{q}_i$. More precisely, as soon as $|\mathbf{r}_i - \mathbf{r}_j| > \delta d$ and/or $|\mathbf{q}_i - \mathbf{q}_j| > 1/\delta d$, then

$$\overline{\mathcal{A}^{\dagger}_{pi}(\mathbf{q}_i, \mathbf{r}_i, t)\mathcal{A}_{qj}(\mathbf{q}_j, \mathbf{r}_j, t)} = \overline{\mathcal{A}^{\dagger}_{\pi}(\mathbf{q}_i, \mathbf{r}_i, t)}\;\overline{\mathcal{A}_{qj}(\mathbf{q}_j, \mathbf{r}_j, t)} = 0 \tag{7.153}$$

The lack of correlation between two non-overlapping volumes is obvious. The $\mathbf{q}_i\mathbf{q}_j$-decorrelation requires a bit of work [50], but finally one finds that two amplitudes decorrelate as soon as the difference of the scattering angles is larger than the divergence of $1/ka$ of the involved beam wave. We conclude: two paths are uncorrelated if there is at least one distinct segment satisfying Eq. (7.153). Such a situation is illustrated in Fig. 7.13a. Uncorrelated segments can be found in a large majority of paths. Therefore we can neglect the cross-terms with $p \neq q$ in Eq. (7.153) and use the decorrelation approximation, Eq. (7.147).

It is obvious that the lack of inter-segment correlation justifies using the pre-averaged version $\bar{\mathcal{T}}$ of the transition matrix (recall Eq. (7.116)). Neglecting the rare path loops indicated in Fig. 7.13b, we can write the time averaged path probability (recall Eq. (7.139)) as the product of time averaged segment probabilities:

$$\overline{\mathcal{P}}_p = \overline{\Pi_{i=1}^{np}\mathcal{P}_i} = \Pi_{i=1}^{np}\overline{\mathcal{P}}_i \tag{7.154}$$

Recall, however, the restriction at the end of Section (7.3.3): if the quasi-particles consist in an ensemble of anisotropic scatterers with randomly fluctuating

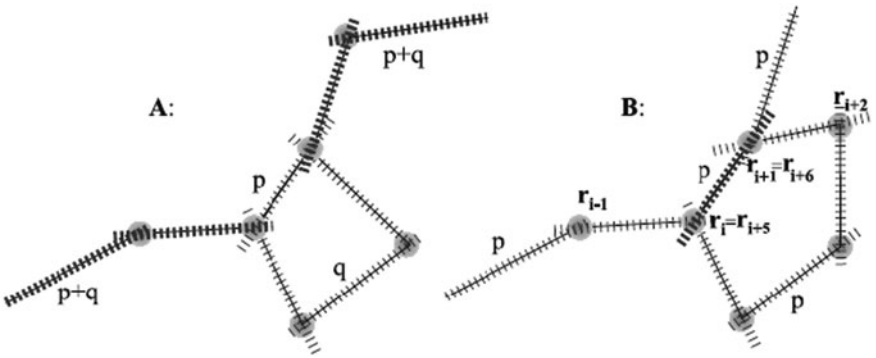


Fig. 7.13 Path decorrelation. (**a**) Paths p and q are decorrelated because of the distinct branching. (**b**) In the path p there is a loop with one correlated segment. We neglect such loops

orientation of the anisotropy axis $\hat{\mathbf{a}}$, we would probably have to resort to Mueller formalism.

The de-correlation approximation is not entirely obvious in the case of a medium with frozen microscopic heterogeneity, as found in most biological tissues. Lacking a better theory, we content ourselves with the assumption that even in such a non-ergodic case we are allowed to use an ensemble pre-averaged segment amplitude. In any case, we are likely to miss certain effects of interference of neighboring paths. For example, the coherent back scattering that results from interference of counter-propagating but otherwise identical paths [45] is excluded by the decorrelation approximation.

7.5.5 Polarized Segment Amplitude and Probability

Having simplified the problem by decorrelation of the paths and segments, we may focus on the effect of polarization on a single segment. The geometry of the segment is illustrated in Fig. 7.14. Notice that the angles ϕ_i and θ_i are Euler angles (as defined in [51]) of rotations relative to the coordinate system i–1. We take it for granted that the transformations from a local coordinate system to the laboratory system have been carried on together with the evolution of the photon path. In other words, we assume that we know a 3D rotation matrix $\mathbf{R_i}$ constructed from the succession of ϕ and θ rotations:

$$\mathbf{R_i} = \Pi_{j=0}^{i} \mathbf{R}_\theta(\theta_j)\mathbf{R}_\phi(\phi_j) \tag{7.155}$$

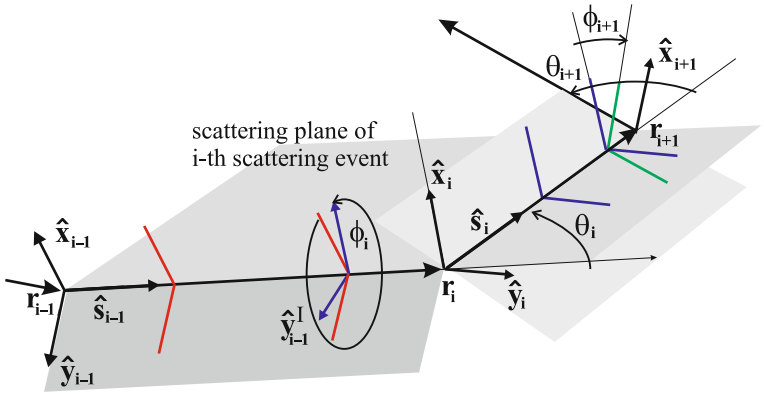


Fig. 7.14 Geometry of a path segment

The element $j = 0$ includes the transformation concerning the initial state $|l\rangle$, i.e. the aiming and polarization alignment of the laser beam. Recall that the segment amplitude matrix $\mathcal{F}_i$ can be concisely written as

$$\mathcal{F}_i(t) = \mathcal{A}(t)\Pi^d(\theta_i)\,\mathbf{R_2}(\phi_i)\ \cdot\ \mathbf{J}\,e^{ik_m d}\,e^{-\mu_e d/2} \tag{7.156}$$

The dot · separates the factors that concern propagation of the photon from those that concern scattering. We first discuss the propagation.

7.5.5.1 Propagation

The factor $\exp(ik_m d)$ is the phasor associated with the optical path $k_m d$ where $d = |\mathbf{r}_i - \mathbf{r}_{i-1}|$. The subscript m in k_m is the reminder of the assumption that the beam propagates in an effective medium of refractive index n_m. The factor $\exp(-\mu_e d/2)$ describes in a crudely simplified fashion the removal of photons from the partial beam $|\pi_{i-1}\rangle$ by scattering and absorption. The extinction coefficient $\mu_e = \mu_s + \mu_a$ includes both processes. The evolution of the polarization state during the propagation is effected by the Jones matrix $\mathbf{J}$, such that $\mathbf{J}^\dagger\mathbf{J} = \mathbf{I}$, where $\mathbf{I}$ is the 2D identity matrix ($\mathbf{J}$ is a unitary matrix). In a plain isotropic medium $\mathbf{J} = \mathbf{I}$. As discussed in Section 4.2.2.2, a chiral medium causes a phase shift between the two circular polarizations, and therefore a rotation of linear polarization. The chiral propagation matrix is given by Eq. (4.69):

$$\mathbf{J} = \begin{pmatrix} \cos(\delta k\, d) & +\sin(\delta k\, d) \\ -\sin(\delta k\, d) & \cos(\delta k\, d) \end{pmatrix} \tag{7.157}$$

where $\delta k = k_L - k_R$. Circular dichroism can also be included in $\mathbf{J}$; i.e. different attenuation for left and right circularly polarized beam waves.

Considerably more challenging is the case of anisotropy of the dielectric response, because in the course of their random flight the photons propagate in arbitrary angles with respect to the optical axis $\hat{\mathbf{a}}$. This is much unlike the propagation through a well aligned retarder plate in Section 4.3. In Section 4.2.2.1 we discussed the uniaxial anisotropy with arbitrary aligned axis, so that we now appreciate the difficulties. A major problem is the birefringence, i.e. the splitting of the beam wave into an ordinary and an extraordinary beam. The photon has two choices of $\hat{\mathbf{s}}$ to propagate. Thus, in the MC simulation we should throw a dice to decide which way to take. But in this way we would miss the interference between the two polarization states and therefore also the possibility of producing circular or elliptical polarization (recall Section 4.3.4.3). Fortunately, the anisotropy of the biological tissue is not too large, so that the expected angles between the propagation vectors of the ordinary and extraordinary beam (Eq. (4.25)) are perhaps smaller than the divergence angle $1/k\delta d$ of the beams. Thus, a good approximation may be to give up the birefringence and propagate the ordinary and extraordinary photon in the same beam wave. This approximation is used, for example, by Gangnus et al. [46] in their

polarized MC simulation. Even so, the inclusion of anisotropy is rather demanding, since the propagation matrix must be newly constructed in each segment of the random flight.

7.5.5.2 Scattering

In the scattering factor $\mathcal{A}(t)\,\mathbf{\Pi}^d\mathbf{R_2}$, the first operator to be applied is the 2D rotation matrix $\mathbf{R}_2(\phi_i)$ which actuates the rotation from the coordinate system i–1 into the intermediate coordinate system I in the scattering plane of the i-th scattering event (see Fig. 7.14 and recall Section 7.1.1 for a more detailed explanation). The matrix $\mathbf{\Pi}^d$ is the dipole polarization matrix, as introduced in Eq. (7.42). For convenience we recall that

$$\mathbf{R}_2(\phi) = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin(\phi) & \cos(\phi) \end{pmatrix} \quad \mathbf{\Pi}^d = \begin{pmatrix} 1 & 0 \\ 0 & \cos(\theta) \end{pmatrix} \tag{7.158}$$

Finally, $\mathcal{A}(t)$ represents the fluctuating scalar scattering amplitude. In the continuum model we write

$$\mathcal{A}(t) = \frac{i\,k_m}{2} \int X_{i-1}(\mathbf{r}) X_i(\mathbf{r})\,\delta\chi(\mathbf{r},t)\,e^{i\,\mathbf{q}_i\cdot\mathbf{r}} d^3r \tag{7.159}$$

Since no depolarization occurs in the single scattering process (recall Sections 7.3.3 and 7.5.4), we simplify matters by replacing $\mathcal{A}(t)$ with its pre-averaged version that is included in the time averaged transition matrix $\bar{\mathcal{T}}$ in Eq. (7.116). Furthermore, recall the cell model discussed in Section 7.5.5 and insert in Eq. (7.116) the following symbols:

$$V_{\mathrm{IOR}} \approx (\pi\,\delta d)^3, \;\; \pi a_l^2 = \pi\,\delta d^2 \text{ and } \Omega_o = \delta\,\Omega = 4\pi/(k\delta d)^2.$$

Thus, we replace Eq. (7.150) by

$$\overline{\mathcal{F}}_i = i\sqrt{\mu_s \delta d\,\delta\,\Omega}\,\hat{\mathbf{S}}\mathbf{R_2}(\phi_i)\cdot\mathbf{J}\,e^{ik_m d}\,e^{-\mu_e d/2} \tag{7.160}$$

Here $\hat{\mathbf{S}}$ is the normalized amplitude scattering matrix. In the Born approximation:

$$\hat{\mathbf{S}} = \sqrt{\frac{F(\mathbf{q}_i)}{\Phi}}\sqrt{\frac{3}{8\pi}}\,\mathbf{\Pi}^d(\theta_i) \tag{7.161}$$

where $F(\mathbf{q}_i) = F(\theta_i, \phi_i)$ is an experimentally determined or modeled structure factor and Φ is the already well known normalization constant, namely the integral of $F(\theta, \phi)$ weighted with the dipole phase function from Eq. (7.44):

$$\Phi = \int F(\theta,\phi)\,p_d(\theta,\phi)\sin(\theta)d\theta d\phi \tag{7.162}$$

Recall that in a structurally isotropic medium (in statistical sense), the structure factor $F(\mathbf{q})$ depends only on $|\mathbf{q}| = k_m\sqrt{2-2\cos(\theta)}|$.

At this point we would like to recall that, unlike all previously published polarized MC simulations, we work here with polarized RDGB scattering instead of scattering from Mie-spheres. Of course, a dilute suspension of monodisperse Mie particles is an excellent test system, good for a comparison of a simulation with a well controlled experiment. If desired, then one can employ the normalized $\hat{\mathbf{S}}$-matrix of a Mie particle instead of $\hat{\mathbf{S}}$ from Eq. (7.161) (recall the discussion in Section 7.3.3). However, we prefer to regard tissue as a fluctuating continuum that can be treated) in RDGB approximation. This allows us to factor the scattering amplitude into the structure amplitude $\sqrt{F(\theta)}$ that reflects the shape of the fluctuation correlation function, i.e. the mean structure of the medium, and in the matrix $\mathbf{\Pi}$ that accounts for the polarization. This factoring is a convenient starting point for future developments that will include anisotropy. Recall that there are two kinds of anisotropy to be encountered in the tissue: (i) Anisotropy of structure, which concerns the structure factor $F(\mathbf{q})$; and (ii) anisotropy of the dielectric response, that concerns the polarization matrix $\mathbf{\Pi}$. A generalization that incorporates both kinds of anisotropy into Monte Carlo simulations would be straightforward. However, anisotropic F and $\mathbf{\Pi}$ depend on the orientation of the local coordinate system with respect to the laboratory frame, which keeps changing at every step of the path. The implementation of anisotropies is an exercise in geometry that requires a considerable portion of patience and concentration.

7.5.5.3 Segment Probability

It remains to calculate the segment probability, $\mathcal{P}_i$, i.e. the probability that a photon that entered the photon state $|\pi_{i-1}\rangle$ at $\mathbf{r}_{i-1}$ will make it to $\mathbf{r}_i$ and there becomes scattered into a state $|\pi_i\rangle$ where it propagates in the direction $\hat{\mathbf{s}}_\mathbf{i}$. We introduce the segment amplitude $\mathcal{F}_i$ from Eq. (7.160) into Eq. (7.142) and rearrange the terms to obtain

$$\overline{\mathcal{P}_i}(\theta,\phi|\hat{\mathbf{e}}_{\mathbf{i-1}}) = p_i(\theta,\phi|\hat{\mathbf{e}}_{\mathbf{i-1}})\delta\Omega\cdot\frac{\mu_s}{\mu_e}\cdot\mu_e e^{-\mu_e d_i}\delta d \tag{7.163}$$

Equation (7.163) deserves to be framed, since it represents the foundation of polarized Monte Carlo simulations. Note that neither δd nor $\delta\Omega$ are used in the actual MC simulation. They only occur as concepts to *derive* Eq. (7.163) instead of simply assuming it.

It is important to realize that the segment probability $\mathcal{P}_i(\theta,\phi|\hat{\mathbf{e}}_{\mathbf{i-1}})$ is the product of three probabilities that are separated by $\cdot$ dot. Reading from right to left, the first factor is the probabilistic representation of Beer's law: $\mu_e\exp(-\mu_e d_i)\delta d$ is the probability that a photon survives propagation over a distance $d_i = |\mathbf{r}_i - \mathbf{r}_{i-1}|$ from $\mathbf{r}_{i-1}$ to $\mathbf{r}_i$, but there it interacts with the matter within a small volume of size δd. The second factor $\mu_s/\mu_e = \mu_s/(\mu_s+\mu_a)$ is the probability that the interaction event is a scattering event, not an absorption. With the probability μ_s/μ_e the photon continues

its path. With the probability $1-\mu_s/\mu_e$ the path is terminated and the energy of the photon is deposited in the matter. The third factor represents the phase function, i.e. the angular distribution of the scattered photons into different outcoming beam waves with propagation vectors $\hat{\mathbf{s}}_\mathbf{i} \equiv (\phi_i, \theta_i)$ and solid angle $\delta\Omega$:

$$p_i(\theta,\phi|\hat{\mathbf{e}}_{\mathbf{i-1}}) = \frac{\langle \pi'_{i-1}| \mathbf{R}_2^{-1}\hat{\mathbf{S}}^\dagger\hat{\mathbf{S}}\mathbf{R}_2 |\pi'_{i-1}\rangle}{\langle \pi'_{i-1}|\pi'_{i-1}\rangle} = \frac{F(\theta,\phi)}{\Phi}\frac{3}{8\pi}\frac{\langle \pi'_{i-1}| \mathbf{R}_2^{-1}\mathbf{\Pi}^\dagger\mathbf{\Pi}\mathbf{R}_2 |\pi'_{i-1}\rangle}{\langle \pi'_{i-1}|\pi'_{i-1}\rangle} \tag{7.164}$$

where $|\pi'_{i-1}\rangle = \exp(-\mu_e d_i/2)\mathbf{J}\,|\pi_{i-1}\rangle$. In Eq. (7.164) one recognizes the polarization factor of a dipole: $\langle \pi'_{i-1}| \mathbf{R}_2^{-1}\mathbf{\Pi}^{d\dagger}\mathbf{\Pi}^d\mathbf{R}_2 |\pi'_{i-1}\rangle / \langle \pi'_{i-1}|\pi'_{i-1}\rangle$. We refrain from recalling the explicit result, since in a simulation algorithm one operates with matrices and vectors instead with lengthy scalar formulas. For the structure factor $F(\theta, \phi)$ we shall insert the proposed general expression from Eq. (7.122). One must also calculate the constant Φ, Eq. (7.162). In the special case of HG structure factor, this work has been already done by Liu and Weng [52], who also provide a re-interpretation of the asymmetry parameter g that takes polarization into account. These authors are probably the first to combine the HG phase function with the dipole scattering matrix; naturally, in the context of light propagation in the atmosphere.

7.5.6 Detecting a Photon

Photon detection is an important stage of the simulation process since only now is the photon actually realized, re-emerging, so to speak, from the maze of the virtual path system into the physical reality. We would like to model the reality as closely as possible and convenient, and therefore we consider a realistic experimental setup for polarization imaging that is outlined in Fig. 7.15. Note that we are employing a telecentric imaging system so that the sample is everywhere imaged from the direction that is perpendicular to the sample surface. There is ample space between the two lenses to insert the optical elements necessary for polarization analysis. Thus, each of the imager channels is a combination of a *beam filter* (Section 4.5) and a *polarization filter* (Section 4.3.4). Only one observation beam is indicated by shading. Recall from Section 4.2.3.1 that the polarization state remains preserved upon

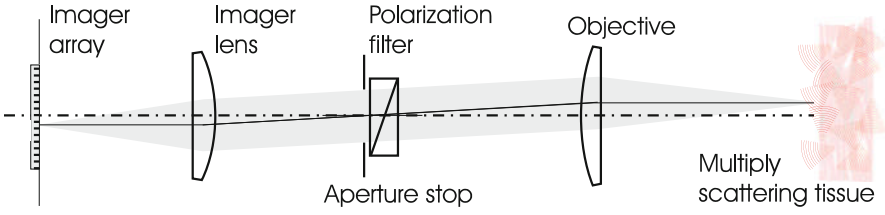


Fig. 7.15 Setup for polarization imaging

moderate focusing by a good lens, but take care to keep the angle of all observation channels low with respect to the optical axis. Not shown in Fig. 7.15 is the laser beam. A clever scheme of beam launching is crucial in order to avoid specular reflection. Examples of practical realizations of polarization imaging can be found in Ref. [53].

With this setup in mind, we must form an image from the simulated polarized photons emerging from the tissue. There are two ways to understand image formation in terms of "physical optics," i.e. optics of electromagnetic waves, as opposed to the optics of geometric rays. The easier approach is due to Rayleigh. One imagines in the object plane infinitesimal light sources, small particles that scatter the illuminating light. Each of the particles radiates a spherical wave, from which a beam wave is filtered out by the objective aperture and re-focused into the geometrical image point in the image plane. The image is formed by the superposition of the amplitudes of many such beams. Rayleigh's approach is simple and intuitive; the only problem is that in our case it is difficult to discern any small particles in the object plane. We are dealing with a random superposition of waves that are radiated from the depth of the tissue. Thus, we must consider an alternative approach that was proposed by Abbe. Abbe forms the image out of waves only, without needing light sources in the object plane. He regards the object as a grating which modulates the phase and amplitude of the transmitted illuminating light. The resulting "complex disturbance" in the object plane is then propagated through the imaging optics according to Huygens-Fresnel principle and Kirchhoff-Fresnel theory of diffraction. To treat polarization, one should use the rigorous vectorial theory of diffraction (see e.g. [8]).

Of course, the two approaches are fully equivalent. Interestingly, they correspond to our two representations of the scalar product of the observation mode with the observed field: whereas Rayleigh employs the source representation (Eq. (4.122)), Abbe prefers the surface representation (Eq. (4.120)) or rather the approximate version, Eq. (4.121). Note that using the virtual observation beam greatly simplifies Abbe's approach: instead of propagating through the imaging optics a complicated superposition of waves, we propagate a well behaved selected beam. The idea is sketched in Figs. 7.15 and 7.16.

We associate each pixel of the imager with a single-mode receiver beam. This is of course an approximation, since CCD pixels are neither single-mode fibers (not yet) nor points. But the approximation is good, if one takes care to match the pixel pitch with the resolution, which is given by the waist radius of the focused receiver beam in the object plane. Mathematically, this is expressed with our usual equation for the observation mode in the Rayleigh range:

$$|o\rangle \equiv \hat{\mathbf{E}}^{\mathbf{o}}(\mathbf{r}) = \sqrt{\frac{2}{c\epsilon_o}} X_o(\mathbf{r}-\mathbf{r}_o)\hat{\mathbf{e}}_{\mathbf{o}}\, e^{ik\,\hat{\mathbf{s}}_{\mathbf{o}}\cdot(\mathbf{r}-\mathbf{r}_o)} \tag{7.165}$$

where $X(\mathbf{r}-\mathbf{r}_o)$ is the amplitude profile of the beam focused in the object point $\mathbf{r}_o$. The propagation vector $\hat{\mathbf{s}}_{\mathbf{o}}$ defines the alignment of the receiver. We chose $\hat{\mathbf{s}}_{\mathbf{o}} = -\hat{\mathbf{z}}_{\mathbf{L}}$, where $\hat{\mathbf{z}}_{\mathbf{L}}$ defines the positive z-direction in the laboratory coordinate

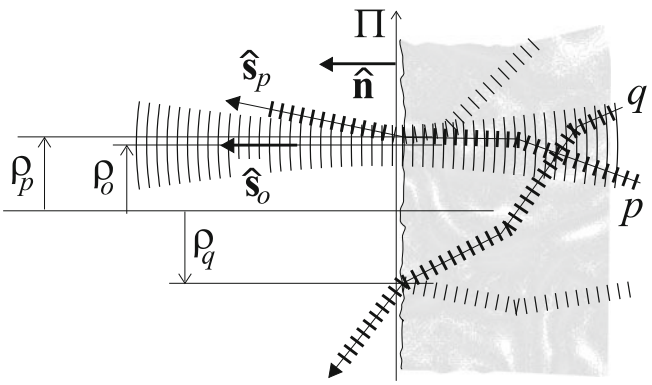


Fig. 7.16 Receiving the paths

system. Finally $\hat{\mathbf{e}}_\mathbf{o}$ represents one of the two unit vectors $\hat{\mathbf{x}}_\mathbf{o}$ and $\hat{\mathbf{y}}_\mathbf{o}$ that specify the linearly polarized base of the observation beam. A reasonable choice is $\hat{\mathbf{x}}_\mathbf{o} = \hat{\mathbf{x}}_\mathbf{L}$ and $\hat{\mathbf{y}}_\mathbf{o} = -\hat{\mathbf{y}}_\mathbf{L}$.

Knowing the light field $\mathbf{E}^s(\mathbf{r})$ as it emerges from the sample in the plane Π, one can calculate, for each pixel, the projection amplitude as the overlap integral (Eq. (4.121)) of $\mathbf{E}^s(\mathbf{r})$ with the pixels receiver mode from Eq. (7.165):

$$\langle o|E^s\rangle \approx \frac{c\epsilon_o}{4}\int_\Pi \hat{\mathbf{E}}^{o*}(\boldsymbol{\rho}-\boldsymbol{\rho}_o)\cdot\mathbf{E}^s(\boldsymbol{\rho})\left[\hat{\mathbf{s}}_s(\boldsymbol{\rho})+\hat{\mathbf{s}}_o\right]\cdot\hat{\mathbf{n}}\,d^2\rho \tag{7.166}$$

Here ρ represents a 2D vector in the Π-plane. Recall from previous sections that we regard the emerging field as being formed by the superposition of partial beam waves of the individual paths. Since the paths are assumed to be decorrelated, we can detect them individually, inserting the beam wave $\mathbf{E}^p(\boldsymbol{\rho})$ of a path p in place of $\mathbf{E}^s(\boldsymbol{\rho})$ in Eq. (7.166). Only two paths are illustrated in the sketch in Fig. 7.16. Note that the paths exit the sample from a medium of refractive index n_m into air. Thus, before considering the detection, we would have to return to Section 4.3.5.1 to implement Fresnel's refraction and reflection in the Jones matrix for the transmission across the (corrugated) interface. Here we only consider the Jones matrix for the transition from the transmitted beam-ray of the path p into the observation mode. Recall from Section 7.1.2 that

$$\begin{pmatrix}\mathcal{E}_x^o\\ \mathcal{E}_y^o\end{pmatrix} = \begin{pmatrix}\langle o_x|E^p\rangle\\ \langle o_y|E^p\rangle\end{pmatrix} = \mathbf{T}_{po}\begin{pmatrix}\mathcal{E}_x^p\\ \mathcal{E}_y^p\end{pmatrix} \tag{7.167}$$

Using the scalar product from Eq. (7.166), one obtains

$$\mathbf{T}_{po} = \mathcal{B}_o(\boldsymbol{\rho}_p,\hat{\mathbf{s}}_p)\begin{pmatrix}\hat{\mathbf{x}}_\mathbf{o}\cdot\hat{\mathbf{x}}_\mathbf{p} & \hat{\mathbf{x}}_\mathbf{o}\cdot\hat{\mathbf{y}}_\mathbf{p}\\ \hat{\mathbf{y}}_\mathbf{o}\cdot\hat{\mathbf{x}}_\mathbf{p} & \hat{\mathbf{y}}_\mathbf{o}\cdot\hat{\mathbf{y}}_\mathbf{p}\end{pmatrix} \tag{7.168}$$

Note that in order to calculate the polarization transfer matrix, all polarization unit vectors must be given in the same coordinate system, preferably in the lab system. Fortunately, we kept updating the 3D rotation matrix $\mathbf{R}_i$ during the evolution of the path (and in Fresnel refraction at the interface) so that now we can apply the operator $\mathbf{R}_p^{-1}$ to the photon polarization vector expressed in the last transverse coordinate system.

The scalar factor $\mathcal{B}_o\left(\boldsymbol{\rho}_p, \hat{\mathbf{s}}_p\right)$ is the overlap integral of the beam-ray p with the observation beam:

$$\mathcal{B}_o(\boldsymbol{\rho}_p, \hat{\mathbf{s}}_p \approx \frac{\left[\hat{\mathbf{s}}_o + \hat{\mathbf{s}}_p\right] \cdot \hat{\mathbf{n}}}{2} \int_{\Pi} X_o(\boldsymbol{\rho} - \boldsymbol{\rho}_o) X_p(\boldsymbol{\rho} - \boldsymbol{\rho}_p) e^{i\mathbf{q}\cdot\boldsymbol{\rho}} d^2\rho \tag{7.169}$$

where $\mathbf{q} = k(\hat{\mathbf{s}}_p - \hat{\mathbf{s}}_o)$ and $X_o(\boldsymbol{\rho} - \boldsymbol{\rho}_o,\ X_p(\boldsymbol{\rho} - \boldsymbol{\rho}_p)$ are the beam profiles within Π, centered at $\boldsymbol{\rho}_o$ and $\boldsymbol{\rho}_p$, as indicated in Fig. 7.16. In a MC simulation we use a simplified version of the beam selection. We include the polarization transfer in the rigorous way, but simplify the beam overlap. Each of the pixels of the CCD array in Fig. 7.15 corresponds to a rectangular pixel on the Π-plane, whose area $\Delta x_o \Delta y_o$ corresponds to the effective cross-section πa_o^2 of the observation beam. The emerging photon states are beams with some effective cross-section $\pi(\delta d)^2$, which we can choose more or less at will, since δd is not used in MC simulation. Therefore we skip the calculation of the overlap integral. A photon is accepted if two conditions are met: (i) in its ultimate jump the photon crosses the interface within the pixel area $\Delta x_o \Delta y_o$ around x_o, y_o; (ii) the angle θ between the photon propagation vector, $\hat{\mathbf{s}}_\mathbf{p}$, and the receiver axis, $\hat{\mathbf{s}}_o$, is lower than the angular aperture of the receiver. Now we insert polarization filters and measure the polarization, as discussed in Sections 4.4 and 4.6.

7.5.7 *From Path Probabilities to Monte Carlo*

We summarize the formal development of the path model. The time-averaged probability $\bar{\mathcal{P}}_d\left(|l\rangle\right)$ that a photon injected in the state $|l\rangle$ will be detected in a single-mode detection channel can be expressed as $\overline{\mathcal{P}}_d = \langle|\mathcal{E}_x^d|^2\rangle_t + \langle|\mathcal{E}_y^d|^2\rangle_t$. Here $\mathcal{E}_x^d$ and $\mathcal{E}_y^d$ are the components of the detected Jones vector, which may be regarded as a vectorial *detection amplitude*. On its way from the laser beam through the tissue into the receiver the photon may propagate on many different paths p, each of them consisting of different numbers np of scattering events at different positions $\{\mathbf{r}_i\}_p$ in the sample. The total detection amplitude is therefore the sum of the amplitudes of individual paths. In our case this so-called *path integral* [54] reads:

$$\begin{pmatrix} \mathcal{E}_x^d \\ \mathcal{E}_y^d \end{pmatrix} = \mathbf{F} \sum_p^{N_p} \mathbf{T}_{po}\, \mathbf{T}_{pt} \left\{ \Pi_{j=1}^{np} \overline{\mathcal{F}_{jp}} \right\} \begin{pmatrix} e_x^l \\ e_y^l \end{pmatrix} \tag{7.170}$$

Here e_x^l and e_y^l are the components of the input state $|l\rangle$ such that $\langle l \mid l\rangle = 1$. The product $\Pi_{j=1}^{np} \overline{\mathcal{F}_{jp}}$ represents a chain of np scattering events with segment amplitudes $\mathcal{F}_{jp}$. We are interested in time averages; and since we employ the decorrelation approximation, we use the pre-averaged segment amplitudes from Section 7.5.5. The matrix $\mathbf{T}_{pt}$ represents a special scattering event, namely the transmission of the photon across the tissue-air interface (recall Section 4.3.5.1). The matrix $\mathbf{T}_{po}$ is the beam filter matrix, Eq. (7.168). Finally, $\mathbf{F}$ is one of the filters from Eqs. (4.60), (4.61) and (4.62).

The key task in Feynman's path analysis would be to calculate the path integral. In our case this would first mean finding *all possible paths* that lead from the laser beam into the detector. This is only slightly simplified by the fact that we do not need to add the amplitudes but only the probabilities. (See [54] for an attempt to apply the path integral approach to radiative transfer.) The problem is that, when searching for *all* paths, there are a great number of paths which are highly improbable. Instead of searching all paths, we select a relatively small sample of paths that are probable enough to bring some of the photons into the detector. Thus, instead of searching all paths and then weighting them with their probabilities, we *sample* the paths by choosing the distances d_i, the absorption events, and the angles ϕ_i, θ_i randomly, according to their probability *distributions*. Keep also in mind that in each step of the path one must examine if the photon makes an attempt to cross the tissue-air interface. If yes, a special sampling procedure is required to decide if the photon is reflected from the interface or if it is transmitted.

Examining Eq. (7.164), one finds a feature that significantly distinguishes the polarized MC simulation from the scalar version and which makes sampling of the angles θ and ϕ in polarized MC technically more demanding: the polarized phase function $p(\theta, \phi; \pi_{i-1})$ depends on the polarization state of the incoming photon, which keeps changing in every segment of the path. Moreover, different initial polarization states $|l\rangle$ prefer different paths. This feature has two practical consequences. First, the only practicable way to sample the angles θ and ϕ is the acceptance-rejection method [55]. However, the problem is to find an efficient *instrumental distribution*, i.e. a simple phase function, $\mathrm{p}'(\theta)$, which forms a tight envelope of $p(\theta, \phi; \pi_{i-1})$ for any incoming state π_{i-1}. This is quite a challenge if one works with Mie spheres [42], but for our dipole scattering matrix the envelope function can be rather easily deduced. The second consequence is that one needs some patience when simulating the Müller matrix. A new simulation is needed for each of the four required initial polarization states.

Note that each *sampling act* corresponds to a *measurement* in the sense of quantum mechanics: each sampling act renders the probability into reality. The photon takes the new path segment, just as a billiard ball would do. Thus, after each step i of the path (including transition across tissue-air interface) has been sampled, one must renormalize the photon state $|\pi_i\rangle$, so that $\langle \pi_i \mid \pi_i\rangle = 1$. No sampling is needed in the final detection steps: if the photon makes it into the receiver, then we assign it the weight $\langle \pi_o | \mathbf{\Pi}^\dagger \mathbf{F}^\dagger \mathbf{F} \mathbf{\Pi} | \pi_o \rangle$, where $\mathbf{F}$ is the filter matrix and $\mathbf{\Pi}$ the polarization matrix from Eq. (7.168). Except for this final step, we refrain from weighting

or re-weighting a previously sampled path. In particular, we *sample* the absorption events, thereby terminating the path, instead of propagating the photon all the way through and then multiplying with $\exp(-\mu_a L)$. We also refrain from re-using a path or its parts. (In particular we generate a new path for each initial state.) Such path re-weighting and re-using is more likely to cause statistical artifacts, rather than a true improvement of the accuracy.

References

1. Chandrasekhar S. *Radiative transfer*. Dover, New York (1960).
2. Debye P, Anderson HR, and Brumberger H. Scattering by an inhomogeneous solid. The correlation function. *J. Appl. Phys.*, 28(6):679–683 (1957).
3. Ishimaru A *Wave propagation and scattering in random media*. Academic, New York (1978).
4. Glatter O and Kratky O. *Small angle X-ray scattering*. Academic, London (1982).
5. Roe R-J. *Methods of X-Ray and neutron scattering in polymer science*. Topics in Polymer Science. Oxford University, New York, Oxford (2000).
6. Lindner P and Zemb Th (eds). *Neutrons, X-rays and light: Scattering methods applied to soft condensed matter*. North-Holland Delta Series. Elsevier, Amsterdam (2002).
7. Xu M and Alfano RR. Fractal mechanisms of light scattering in biological tissue and cells. *Opt. Lett.*, 30(22):3051–3053 (2005).
8. Born M and Wolf E. *Principles of optics: Electromagnetic theory of propagation, interference and diffraction of light*. University Press, Cambridge, England, 7th edition (1999).
9. Bohren CF and Huffman D R. *Absorption and scattering of light by small particles*. Wiley, New York, 2nd print edition (1998).
10. van de Hulst HC. *Light scattering by small particles*. Dover, New York, reprint, Dover edition (1981).
11. Jackson JD. *Classical electrodynamics*. Wiley, New York, 2nd edition (1975).
12. Vinogradov AP and Aivazyan AV. Scaling theory for homogenization of the Maxwell equations. *Phys. Rev. E*, 60(1):987–993 (1999).
13. Mobley CD. *Light and water: Radiative transfer in natural waters*. Academic, London (1994).
14. Nilsson AMK, Alsholm P, Karlsson A, and Andersson-Engels S. T-matrix computations of light scattering by red blood cells. *Appl. Opt.*, 37(13):2735–2748 (1998).
15. Barber PW and Hill SC. *Light scattering by particles: Computational methods*, Vol. 2. Advanced Series in Applied Physics. World Scientific, Singapore, NJ, London, Hong Kong (1990).
16. Kerker M. *The scattering of light and other electromagnetic radiation*. Academic, New York (1969).
17. Mandelbrot BB. *The fractal geometry in nature*. Freeman, San Francisco (1982).
18. Beaucage G. Small-angle scattering from polymeric mass fractals of arbitrary mass-fractal dimension. *J. Appl. Crystallogr.*, 29:134–146 (1996).
19. Schmitt JM and Kumar GG. Turbulent nature of refractive-index variations in biological tissue. *Opt. Lett.*, 21(16):1310–1312 (1996).
20. Kim K, Mandel L, and Wolf E. Relationship between Jones and Mueller matrices for random-media. *J. Opt. Soc. Am. A*, 4(3):433–437 (1987).
21. van Tiggelen B and Stark H. Nematic liquid crystals as a new challenge for radiative transfer. *Rev. Mod. Phys.*, 72(4):1017–1039 (2000).
22. Saidi IS, Jacques SL, and Tittel FK. Mie and Rayleigh modeling of visible-light scattering in neonatal skin. *Appl. Opt.*, 34:7410–7418 (1995).
23. Dunn A and Richards-Kortum R. Three-dimensional computation of light scattering from cells. *IEEE J. Sel. Top. Quantum Electron*, 2:898–905 (1996).

24. Schneiderheinze DHP, Hillman TR, and Sampsonr DD. Modified discrete particle model of optical scattering in skin tissue accounting for multiparticle scattering. *Opt. Exp.*, 15:15001–15010 (2007).
25. Henyey LG and Greenstein JL. Diffuse radiation in the galaxy. *Astrophys. J.*, 93(1):70–83 (1941).
26. Jacques S, Alter C, and Prahl S. Angular dependence of HeNe laser light scattering by human dermis. *Lasers Life Sci.*, 1:309–333 (1987).
27. Lue N, Bewersdorf J, Lessard MD, Badizadegan K, Feld MS, and Popescu G. Tissue refractometry using hilbert phase microscopy. *Opt. Lett.*, 32(24):3522–3524 (2007).
28. Popp AK, Valentine MT, Kaplan PD, and Weitz DA. Microscopic origin of light scattering in tissue. *Appl. Opt.*, 42(16):2871–2880 (2003).
29. Gelebart B, Tinet E, Tualle JM, and Avrillier S. Phase function simulation in tissue phantoms: A fractal approach. *Pure Appl. Opt.*, 5:377–388 (1996).
30. Wang RK. Modelling optical properties of soft tissue by fractal distribution of scatterers. *J. Mod. Opt.,* 7(1):103–120 (2000).
31. Schmitt JM and Kumar G. Optical scattering properties of soft tissue: A discrete particle model. *Appl. Opt.*, 37:2788–2797 (1998).
32. Sheppard CJR. Fractal model of light scattering in biological tissue and cells. *Opt. Lett.*, 32(2):142–144 (2007).
33. Gradshteyn IM, Ryzhik IM, and Jeffrey A. *Table of integrals, series, and products*. Elsevier, Amsterdam, 7th edition (2007).
34. Yurkin MA and Hoekstra AG. The discrete dipole approximation: An overview and recent developments. *J. Quant. Spectrosc. Radiat. Transf.*, 106:558–589 (2007).
35. Fymat AL and Vasudevan R. New approach to radiative-transfer theory using Joneses vectors. *Astrophys. Space Sci.*, 38(1):95–124 (1975).
36. Chandrasekhar S. Stochastic problems in physics and astronomy. *Rev. Mod. Phys.*, 15(1):0001–0089 (1943).
37. Sheng P. *Introduction to wave scattering, localization and mesoscopic phenomena*. Springer, Berlin, 2nd edition (2006).
38. Bartel S and Hielscher AH. Monte Carlo simulations of the diffuse backscattering Mueller matrix. *Appl. Opt.*, 39(10):1580–1588 (2000).
39. Ramella-Roman JC, Prahl SA, and Jacques SL. Three Monte Carlo programs of polarized light transport into scattering media: Part i. *Opt. Express*, 13(12):4420–4438 (2005).
40. Wang XD and Wang LHV. Propagation of polarized light in birefringent turbid media: A Monte Carlo study. *J. Biomed. Opt.*, 7(3):279–290 (2002).
41. Wang XD, Yao G, and Wang LHV. Monte Carlo model and single-scattering approximation of the propagation of polarized light in turbid media containing glucose. *Appl. Opt.*, 41(4):792–801 (2002).
42. Jaillon F and Saint-Jalmes H. Description and time reduction of a Monte Carlo code to simulate propagation of polarized light through scattering media. *Appl. Opt.*, 42(16):3290–3296 (2003).
43. Wood MFG, Guo XX, and Vitkin IA. Polarized light propagation in multiply scattering media exhibiting both linear birefringence and optical activity: Monte Carlo model and experimental methodology. *J. Biomed. Opt.*, 12(1):014029 (2007).
44. Martinez AS and Maynard R. Faraday-effect and multiple-scattering of light. *Phys. Rev. B*, 50(6):3714–3732 (1994).
45. Lenke R and Maret G. Magnetic field effects on coherent backscattering of light. *Eur. Phys. J. B*, 17(1):171–185 (2000).
46. Gangnus SV, Matcher SJ, and Meglinski IV. Monte Carlo modeling of polarized light propagation in biological tissues. *Laser Phys.*, 14:886–891 (2004).
47. Feynman RP. Space-time approach to non-relativistic quantum mechanics. *Rev. Mod. Phys.*, 20(2):367–387 (1948).

48. Feynman RP. *QED the strange theory of light and matter*. University Press, Princeton, NJ (2006).
49. Keller O. Attached and radiated electromagnetic fields of an electric point dipole. *J. Opt. Soc. Am. B*, 16:835–847 (1999).
50. Schätzel K. Suppression of multiple scattering by photon cross-correlation techniques. *J. Mod. Opt.*, 38:1849–1865 (1991).
51. Goldstein H. *Classical mechanics*. Addison-Wesley, Reading, MA, USA, 2nd edition (1980).
52. Liu QH and Weng FZ. Combined Henyey-Greenstein and Rayleigh phase function. *Appl. Opt.*, 45(28):7475–7479 (2006).
53. Duncan DD, Jacques SL, and Johnson PC (eds). *Laser-tissue interaction XII: Photochemical, photothermal, and photomechanical*, Vol. 4257. Proc. SPIE. Polarized Light Interactions (2001).
54. Miller SD. Stochastic construction of a Feynman path integral representation for Green's functions in radiative transfer. *J. Math. Phys.*, 39:5307–5315 (1998).
55. Robert CP and Casella G. *Monte Carlo statistical methods*. Springer, New York (2004).

Chapter 8
Measurement of Ex Vivo and In Vivo Tissue Optical Properties: Methods and Theories

Anthony Kim and Brian C. Wilson

8.1 Introduction

In this chapter, the various experimental techniques that have been developed to measure the optical scattering and absorption properties of tissues are discussed, together with the theory underlying these methods. The fundamental optical properties of interest are the absorption coefficient, μ_a, scattering coefficient, μ_s, total attenuation coefficient, $\mu_t = \mu_a + \mu_s$, scattering phase function, $p(\cos\theta)$, scattering anisotropy, g, reduced scattering coefficient, $\mu_s' = \mu_s(1 - g)$, and the tissue refractive index, n. These optical properties are parameters in the radiation transport equation that describes the propagation of light in tissue. Another parameter often measured is the effective attenuation coefficient, μ_{eff}, which describes the exponential attenuation of light with depth in tissue.

There are several ways to classify techniques for measuring tissue optical properties. The majority of these are photometric techniques, using light to probe the tissue. Measuring the tissue optical properties using photometric techniques can best be classified in terms of ex vivo and in vivo methods. Ex vivo methods can be considered to employ direct or indirect measurements. Direct measurements refer to those in which a particular microscopic coefficient is measured in a way that does not require a model of light propagation. In contrast, indirect methods involve measuring parameters (e.g. reflection and transmission) from which the fundamental coefficients are obtained by solving an "inverse problem" based on a model of light propagation in tissue, such as the adding-doubling model [1, 2] or diffusion approximation [3]. That is, the optical properties are placed in a light propagation model and values for the measured parameters are fitted to the optical properties [4–6].

Significant research into in vivo photometric methods during the past decade has resulted in a plethora of techniques that can be further divided in several different ways. In vivo techniques can be classed in terms of measurement geometry

B.C. Wilson (✉)
Department of Medical Biophysics, University of Toronto and Ontario Cancer Institute, Toronto, ON M5G 2M9, Canada
e-mail: wilson@uhnres.utoronto.ca

A.J. Welch, M.J.C. van Gemert (eds.), *Optical-Thermal Response of Laser-Irradiated Tissue*, 2nd ed., DOI 10.1007/978-90-481-8831-4_8,

(reflectance [3], transmittance [7]), the quantity being measured (intensity, fluence rate [8] or radiance [9]), or how the measurement is resolved (measurements can be made in the spatial [3, 10], temporal [11] or frequency domains [12], among others). In vivo photometric measurements are fundamental in the deployment of optically-based diagnostic and therapeutic procedures; hence, this broad class of methods is of considerable interest in biophotonics research and applications.

Finally, one can consider separately photometric methods compared with photothermal [13–15] methods. In photometric techniques, light itself is measured, either as it propagates within the tissue or is diffusely reflected or transmitted. Photothermal methods rely on detecting secondary signals produced by the thermal changes in tissue following absorption of light. These various approaches to tissue optical measurement each have their own advantages and limitations, as will be discussed.

8.2 Photometric Techniques

8.2.1 Ex Vivo Methods

Ex vivo measurements of tissue optical properties can be accomplished in three different ways:

1. Direct measurement of light absorbed or scattered by optically thin (single scattering) tissue sections;
2. Indirect measurement of the (macroscopic) diffuse reflectance and transmittance in optically thick (multiple scattering) tissue sections and solution of the inverse problem to derive the fundamental (microscopic) optical properties;
3. Indirect measurements in homogenized bulk tissue samples, using the "added absorber" method in which a macroscopic parameter, such as the effective attenuation coefficient or diffuse reflectance, is measured as a function of the known concentration of an optical absorber, which is added uniformly to the tissue; as with method 2, this requires application of a light propagation model to relate the measured parameters to the required optical properties.

8.2.1.1 Direct Measurements in Optically Thin Tissue Sections

The optical absorption and scattering properties of tissue can, in principle, be measured directly in optically thin, i.e., single scattering, samples, as illustrated in Fig. 8.1. The total attenuation coefficient is given by

$$\mu_t = -\frac{1}{d} \ln T_c \tag{8.1}$$

where T_c is the fraction of normally incident collimated flux that is transmitted through the sample, of thickness d, with minimal scattering. This requires that the

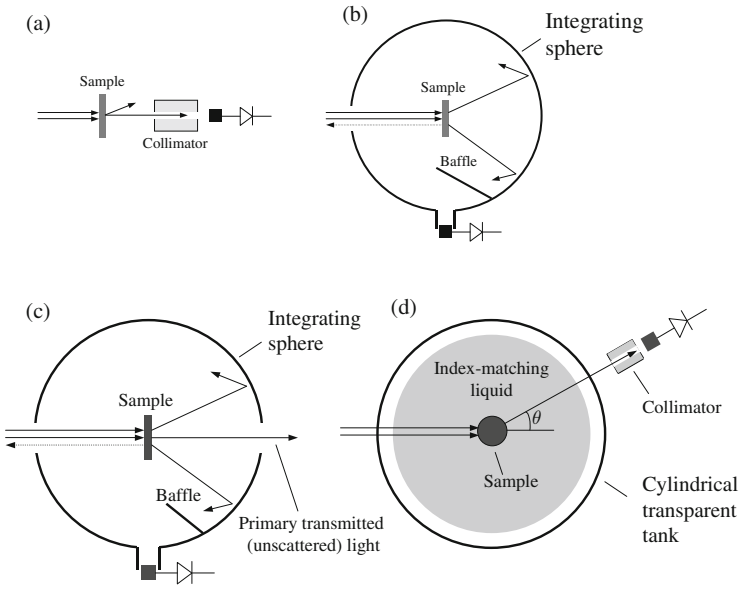


Fig. 8.1 Techniques for measuring optical properties directly in optically thin tissue sections: (**a**) μ_t, measuring the pencil beam transmission using a collimated detector. (**b**) μ_a, measuring the transmitted and scattered light using an integrating sphere. The specular reflectance is rejected through the input port, and the baffle prevents light scattered from the sample reaching the detector directly without re-scattering from the sphere. (**c**) μ_s, measuring the total scattered light. The unscattered primary photons exit via a small coaxial port. (**d**) $p(\cos\theta)$, measuring the angular distribution of singly-scattered light using a collimated detector mounted on a rotating arm

detector be strongly collimated so as to eliminate as many scattered photons as possible (Fig. 8.1a).

The absorption coefficient is determined by measuring all photons transmitted or scattered by the sample, so that the only loss is due to absorption. This is most conveniently done using an integrating sphere which has a highly diffuse reflective coating and which detects a fraction, η, of all photons within it. Then

$$\mu_a = -\frac{1}{d} \ln\left(\frac{N_a}{\eta N_0}\right) \tag{8.2}$$

where N_0 is the number of photons incident on the sample and N_a is the number detected (Fig. 8.1b). Similarly, for the scattering coefficient, if N_s is the number of scattered photons detected (Fig. 8.1c), then (assuming that $\mu_s \gg \mu_a$)

$$\mu_s = -\frac{1}{d} \ln\left(1 - \frac{N_s}{\eta N_0}\right) \tag{8.3}$$

Collimated photons must be allowed to escape and not be detected along with the scattered photons for this equation to be valid.

The scattering phase function, $p(\cos\theta)$, can be measured by rotating a collimated detector around the sample (Fig. 8.1d), which is held within a cylindrical tank containing a transparent liquid of approximately the same refractive index as that of the tissue so as to minimize distortions due to refraction of the scattered light. The value of $p(\cos\theta)$ is then proportional to the detected signal, divided by $\cos\theta$ in order to normalize for the acceptance solid angle of the detector.

Although simple in principle, it is extremely difficult to make these direct measurements accurately in practice, for numerous reasons. A fundamental common problem relates to the thickness of the sample. In order that multiple scattering in the sample be negligible, the condition $d \ll 1/\mu_s$ must hold. Since, typically, the optical scattering coefficient for soft tissues is in the range 100–1000 cm^{-1}, the sample must be less than about 10 μm thick. This has three main consequences. First, such thin samples must be prepared either by microtome sectioning of frozen tissue, or by homogenizing the tissue to a semi-liquid state for use in a microcuvette. In either case, the handling may alter the optical properties. Thin tissue slices dehydrate quickly, shrink and leak blood. As well, freezing and thawing can change the scattering properties of the tissue. Second, since such thin sections must be mounted either in a cuvette or between glass slides to provide mechanical support, it is difficult to ensure that the tissue surfaces are optically smooth. Thus, surface scattering artifacts may contribute spuriously to the signal. Third, since the interaction probability in the sample is necessarily low, the detected signal is very weak and may easily be masked by unavoidable fluctuations in the incident light flux, by ambient lighting, ambiguity between collimated and forward-directed light, or, in the case of μ_a and μ_s, by non-uniformity of response of the integrating sphere.

Several approaches have been taken to minimize these effects. As seen from Eq. (8.1) if the single scattering condition is valid, then plotting $-\ln T_c$ versus d should give a straight line of slope μ_t. An example of such a plot is shown in Fig. 8.2a, and illustrates two artifacts. As d increases, the data deviate from linearity. This is due to the detection of multiply scattered photons. The problem can be reduced by using stronger detector collimation, but cannot be completely eliminated since there are always some photons which are coaxial after multiple scattering. There is also a practical limit to the collimation, since the signal decreases as the collimation is increased. For reasonably achievable acceptance solid angles of, say 10^{-5} *sr*, multiple scattering starts to affect the measurements for tissue thicknesses greater than about 10% of the reduced scattering mean free path ($1/\mu_s'$). A second experimental artifact is seen if the straight line fit is extrapolated to $d = 0$, since the fraction transmitted is not 1.0 as it should be. This effect has been attributed to loss of light due to surface scattering by the sample. (Note that this occurs even if, as is required, corrections are made for the known specular reflection losses from the cuvette or glass slides holding the tissue). Similar effects are observed in the case of the absorbance, A (related to the absorption coefficient by $A = \log_{10}[\exp(\mu_a)]$, measured using an integrating sphere (Fig. 8.2b). Measurements over a range of sample thicknesses thus allow these effects to be observed, and the correct optical properties estimated.

Since the "signals" are small in these experiments, it is very important to minimize sources of uncertainty. Ideally, the light source should be stable over the

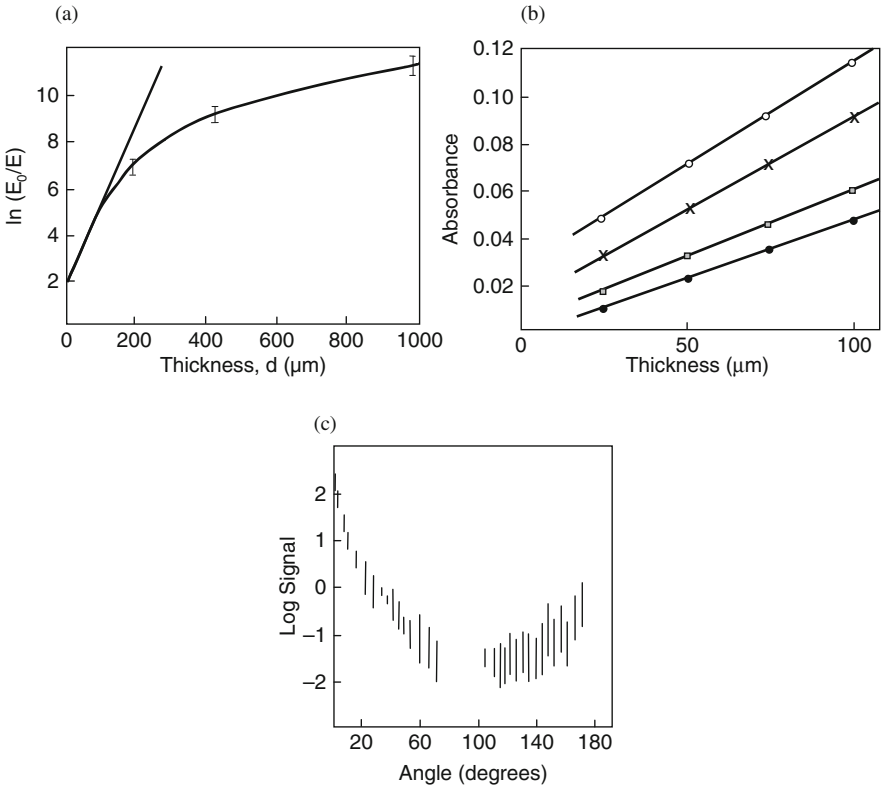


Fig. 8.2 Examples of ex vivo direct measurements in optically thin tissue samples: (**a**) total attenuation measurement for muscle tissue (chicken) at 633 nm (as in Fig. 8.1a) [19]. The *straight line* is the best linear fit to the measurements before significant multiple scattering takes place. The offset from the origin may be due to artifacts from the interface between the tissue and the glass cells holding the sample in place. (**b**) absorbance measurements for human lung tissue samples at 635 nm (as in Fig. 8.1b) with each line representing a different subject, using a non-absorbing diffuser as a reference signal [20]. (**c**) phase function measurement for muscle (chicken) at 633 nm, with the signal being proportional to $p(\theta)$ (as in Fig. 8.1) [19]. (Adapted with permission from Refs. [19, 20])

period of the measurement, both in output power and in beam position and profile. Often the output is monitored, for example, by splitting off a fraction of the incident beam using a glass slide at 45°, which is then measured by a separate calibrated detector. Care must be taken when using a laser source to ensure that the polarization of the output beam is constant, since this influences the fraction of light split off. This is not a problem with a lamp source; but if an arc lamp is used, the slight fluctuations in the arc position may be significant, especially in the measurements for total attenuation and phase function, where the precise position of the beam is critical because the acceptance angles have to be so small in order to eliminate multiple scattering. With either type of source, it is best to match the detection aperture and incident beam size in order to get the maximum signal with the minimum scatter

contribution. Likewise, the detector sensitivity and the light source power should be properly matched in order to maximize the signal-to-noise photodiode detector. Incident powers in the μW–mW range are usually adequate, and this avoids any heating of the sample. It is useful to chop the beam and measure the detector signal via a lock-in amplifier, since this minimizes effects due to ambient lighting.

A major problem arises in using integrating spheres for measurements such as those shown in Fig. 8.1(b, c) because of the fact that real integrating spheres do not have a perfectly uniform response. That is, the detected signal can depend on the spatial distribution of incident or scattered light within the sphere, even if a baffle is used (as shown) in order to prevent photons scattered by the sample reaching the detector directly. These variations (typically a few percent) can be large compared to the changes in signal due to, say, the absorption in the sample. Consider, for example, measuring μ_a in a typical soft tissue in the visible wavelength range, for which $\mu_a \sim \mu_s/100$. Since $d \ll 1/\mu_s$, in order to avoid multiple scattering, the loss of photons due to absorption, representing the signal to be detected, will be only a fraction of a percent. Measurements of μ_a have been reported in which the tissue sample is mounted side-by-side with a non-absorbing light diffuser [21]. The difference in signal between the two can then be attributed to absorption in the sample, assuming that the tissue and the diffuser both produce equivalent distributions of scattered light. In principle, this can be achieved by using well-defined scattering suspensions, if the scattering coefficient and phase function of the tissue are known or can be estimated beforehand. A further problem which has not been solved is that, since scattered photons strike the sphere and are reflected until a steady-state diffuse irradiance is established (of which a fraction is detected as the signal), some of the diffuse photons will be absorbed by the sample. This can be minimized by using an integrating sphere which is large compared to the tissue sample size; for example, Vogel et al. have used a 64 cm diameter sphere to measure the absorption coefficient of scleral tissue with samples of 15×15 mm [22]. The sample alignment is also critical to ensure that the specular reflection from the surface exits via the input port and is not detected. As a result of these many difficulties in the use of integrating spheres, some of which are discussed by Pickering et al. [23], direct measurements of μ_a and μ_s in tissue are very sparse.

There are also problems in measuring the scattering phase functions directly by the method shown in Fig. 8.1d, which is sometimes known as goniometry or nephelometry. The measured phase functions are usually incomplete. Ideally, the sample would be cylindrical in form, but due to the difficulty of preparing such samples, the studies reported to date have all used flat samples, for which the measurements around ±90° are missing. The smallest polar angle at which measurements can be made ($\theta \sim 0°$) is determined by the diameter of the incident beam, while the largest angle ($\theta \sim 180°$) is limited by the detector blocking the incident beam. It is also extremely difficult to account for all the factors that affect the absolute magnitude of the detected signal, such as the reflection and refraction losses, the acceptance solid angle, and the detector sensitivity. The measurements can be normalized by integrating the data over 0–180° (allowing for the solid angle) and equating this to the scattering coefficient, if known. This requires interpolation for the missing data

in the angular distribution, for example, by fitting the data to a theoretical phase function such as the Henyey-Greenstein form (see Chapter 3). As in the other cases, it is possible to measure the phase functions for different sample thicknesses in order to ensure that multiple scattering does not distort the distributions.

8.2.1.2 Indirect Measurements in Optically Thick Tissue Sections

Figure 8.3 shows the principle of using the diffuse reflectance, R_d, and diffuse transmittance, T_d, in tissue sections in which there is multiple photon scattering to determine tissue optical properties. If only R_d and T_d are measured, then it is not possible to determine all three values, μ_a, μ_s, and g. However, under conditions where the "Similarity Principle" (see Chapter 6) applies, it is possible to estimate the two independent quantities, μ_a and μ_s'. In this situation it is assumed that the reflection and transmission for a slab of tissue with optical properties μ_a, μ_s, and $g \neq 0$ are the same as those for the same slab with optical properties μ_a, μ_s' and $g = 0$. This is only approximately true for tissue sections of finite thickness, and breaks down increasingly as the sections become thinner and comparable to the effective penetration depth ($1/\mu_{\text{eff}}$).

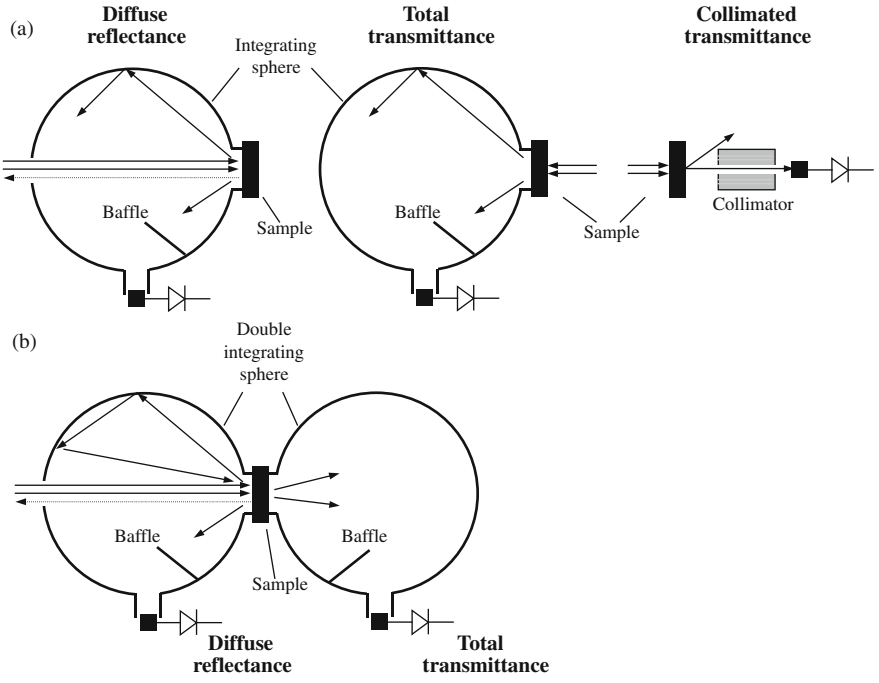


Fig. 8.3 Diffuse reflectance and total transmittance measurements: (**a**) experimental set-up using a single integrating sphere to measure the diffuse reflectance and total transmittance in order to determine μ_a and μ_s'. The collimated transmittance can also be measured to also determine g. (**b**) set-up using a double integrating sphere. Note that the light incident on the sample in this case is the primary beam plus a fraction of the diffusely reflected light

The relationships between R_d and T_d and the optical properties have been studied in two ways: (1) by approximate analytic models such as multiple-flux models [24, 25], diffusion theory [17, 18], or the adding-doubling method [1, 2]; and (2) by Monte Carlo modeling [4, 5, 16] (see Chapter 5). The analytic models are more elegant in principle, since they allow R_d and T_d to be expressed in closed form in terms of μ_a and μ_s'; however, the accuracy in determining the coefficients is limited by the ability of the model to describe the specific experimental conditions. On the other hand, Monte Carlo calculation of R_d and T_d allows incorporation of any source-tissue-detector geometry and the accuracy is limited only by the accuracy of the simulation geometry with respect to tissue and computation. The analytic approach would be especially advantageous if it were possible to derive analytic, or at least computable, expressions for the inverse solution, i.e., to write closed-form equations for μ_a and μ_s' in terms of R_d and T_d. However, this has not been achieved to date, so that, for both the analytic and Monte Carlo methods, it is necessary either (a) to "forward calculate" R_d and T_d over the expected range of optical property values and, as illustrated in Fig. 8.4a, interpolate from

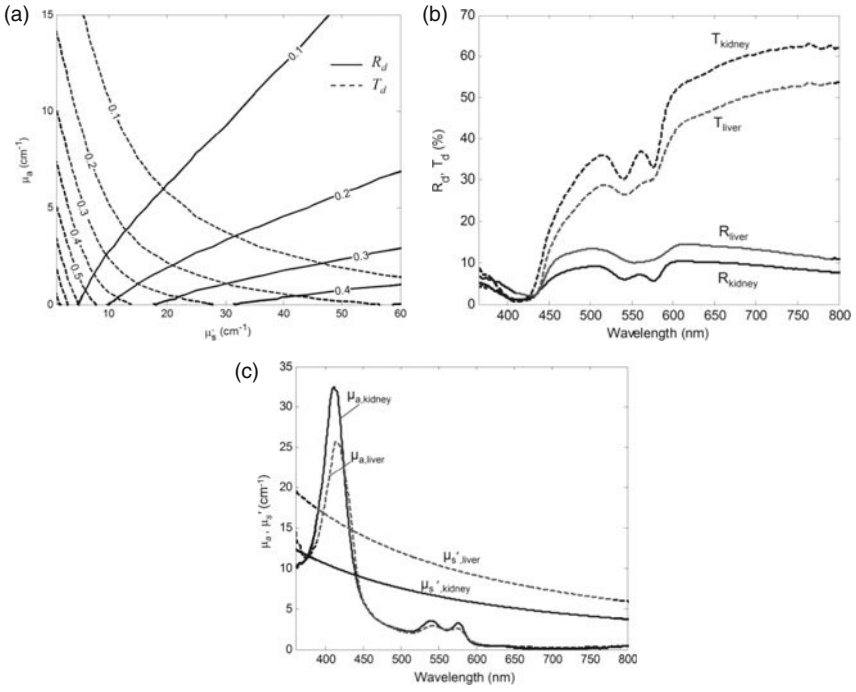


Fig. 8.4 (**a**) Contour plots of constant R_d and T_d as a function of the absorption and reduced scattering coefficients. The contour lines are interpolated from Monte Carlo data for a single integrating sphere set-up. (**b**) Reflectance and transmittance spectra collected from a single integrating sphere in bovine liver (*dashed lines*) and porcine kidney (*solid lines*), using white light irradiation and a spectrometer. (**c**) After solving the inverse problem, the μ_a and μ_s' spectra are extracted (Courtesy M Roy, University of Toronto)

the resulting tables of charts or use a polynomial fit of the data [4, 17], or (b) to use an iterative procedure [5, 6, 16, 26]. The optimum choice of model and calculation method to derive the optical properties from the measurements involves a trade-off between accuracy, speed and flexibility (to describe fully the experimental arrangement).

The system shown in Fig. 8.3a uses a single integrating sphere in two different geometries. In this case the incident light striking the tissue sample is that due to the primary collimated beam. The absolute diffuse reflectance and transmittance can be determined by calibrating the sphere and detector system using reflectance standards. Contour plots of R_d and T_d for a single integrating sphere scheme are shown in Fig. 8.4a. Reflectance and transmittance spectra of bovine liver and porcine kidney (tissue was homogenized by grinding and then diluted to 25% concentration) were measured with this set-up, using a quartz-tungsten-halogen white light source with the emitted light detected by a spectrometer (Fig. 8.4b). The corresponding μ_a and μ_s' spectra for the liver and kidney data were evaluated from the measurements and are shown in Fig. 8.4c. The absorption peaks at 414, 540 and 577 nm are due to blood absorption. The reduced scattering coefficient, μ_s', decreases with wavelength proportional to λ^{-k} where k is 1.49 for liver and 1.50 for kidney.

With a double integrating sphere system (Fig. 8.3b), where the tissue sample is located in a common port of the two spheres, there is "cross talk" between the two spheres. Thus, for example, the total fluence striking one side of the tissue is that from the primary beam, plus a fraction of the diffusely reflected light, plus the light that is transmitted back through the sample from the second sphere. Similarly, the signal from the transmittance sphere has several components. The net effect is to increase the signal in both spheres over the single-sphere case. The powers detected in each sphere then depend on various response factors of the spheres and on both R_d and T_d, which can be calculated from the measured signals if the response factors are calculated or measured experimentally in separate experiments with known standard samples.

With the integrating sphere geometry, it is possible to make a measurement of transmitted collimated light T_c by including an exit port and a collimated detector, or to place the tissue sample in a separate apparatus specifically for measuring T_c. Measurement of R_d, T_d, and T_c is sufficient to resolve μ_a, μ_s, and g. In this case, the sample has to be thin enough to obtain a measurable T_c signal. This reduces the number of scattering events, and the measured R_d and T_d may not correspond to completely diffused light. This has to be accounted for in the modeling. A compromise is also required in the distance of the T_c detector from the sphere: if this is too small, then scattered photons may be detected; but if the incident beam is not perfectly collimated, then the measurable signal decreases as the distance increases.

The light source used for R_d, T_d and T_c measurements can be either a low-power laser or a narrow collimated beam from a lamp. With a white light source and spectrometer (or excitation filters and a monochromator), spectral scans can be made, from which the wavelength dependence of the optical properties can be obtained. With a collimated laser source, interference of the coherent light can occur between the glass slides normally used to hold the tissue sample. This can affect the

reflectance signal significantly, but may be eliminated by using wedge-shaped tissue slices. For either source, the sample diameter should be large compared to the beam diameter in order to minimize light loss through the side of either the sample or, by multiple internal reflections, the slides themselves. At the same time, however, the sample area should be small compared to the surface area of the integrating spheres, otherwise complex, high-order geometrical correction factors are required to determine R_d and T_d accurately.

A significant advantage of using a double integrating sphere is that R_d and T_d can be measured simultaneously, which allows changes in μ_a and μ_s' to be followed, for example, during heating or other modification of the tissue. If T_c is also measured simultaneously, then μ_s and g can also be monitored in the same sample. Alternatively, $T_{c,}$ and hence μ_t, may be measured in a separate, optically-thin sample at the same time as R_d and T_d are measured in an optically-thick sample, utilizing the same source by splitting the incident beam into two separate beams [27].

Considering the number and complexity of the factors which must be taken into account in making integrating sphere measurements of R_d, T_d, and T_c, especially when a simultaneous double-sphere technique is used, it is advisable to perform a complete calibration of the experiment using samples of known absorption and scattering properties over the range of those anticipated for the tissues and wavelengths on interest. The accuracy of the measurements is expected to be poorest if the optical depth or albedo of the sample are either very large or very small. For example, using tissue phantoms comprising different concentrations of a scattering fat emulsion and an absorbing dye, Pickering et al. [23] showed that double-integrating sphere measurements with simultaneous determination of R_d, T_d and T_c gave $\mu_{a,}$ μ_s and g values to an accuracy, under optimal conditions, of typically better than 5% for optical depths between 1 and 10 and albedos between 0.4 and 0.95. For higher albedos the accuracy in μ_a was worse, while for lower albedos the scattering coefficient and anisotropy were less well determined.

8.2.1.3 Indirect Measurements in Bulk Tissue

Added Absorbers

Both the fluence rate distribution in tissue and the diffuse reflectance or transmittance are altered by the addition of a uniformly distributed absorber to the tissue. This “added-absorber” method is equivalent to the poisoned-moderator technique of neutron transport physics [28]. It has been applied in highly scattering tissue by adding increasing concentrations of material of known absorption coefficient to tissue homogenates in a semi-infinite geometry. For broad beam irradiation, the effective attenuation coefficient, μ_{eff}, is measured at each concentration. From diffusion theory

$$\mu_{\text{eff}}^2 = 3\mu_a\left(\mu_a + \mu_{st}'\right) \tag{8.4}$$

where the total absorption coefficient $\mu_a = \mu_{at} + \mu_{ad}$, μ_{at} and μ_{ad} being, respectively, the absorption coefficients of the tissue and the dye, while μ'_{st} is the reduced scattering coefficient of the tissue. If $(\mu_{at} + \mu_{ad}) \ll \mu'_{st}$, then

$$\mu_{\text{eff}}^2 \approx 3\,(\mu_{at} + \mu_{ad})\,\mu'_{st} \tag{8.5}$$

The tissue scattering is assumed to be unchanged by adding the dye. The tissue absorption and transport scattering coefficients are then calculated from the slope and intercept of μ_{eff}^2 versus μ_{ad}. This method has been used with molecular dyes or India ink as the added absorber, to determine μ'_{st} in various soft tissues at visible wavelengths [29]. Care should be taken in using particulate dyes such as India ink, as the scattering coefficient has been shown to be non-negligible due to sub-micron particles suspended in the ink [30].

An alternative is to measure the total diffuse reflectance, R_d as a function of added absorber, where R_d can be expressed as a function of the transport albedo, $a' = \mu'_{st}/(\mu_{at} + \mu_{ad} + \mu'_{st})$, so that the tissue optical properties can be deduced in a similar way from the dependence of R_d on μ_{ad}.

The requirements for adding absorber uniformly to the tissue are most easily met by homogenizing the tissue; however, this can alter the optical properties. Chan et al. performed a study on tissue that was cryopreserved, ground into powder and then thawed to a paste to determine the difference in the tissue optical properties compared to intact tissue [31]. Many adverse effects were cited as potential sources of error due to the cryogenic grinding procedure, such as moisture absorption of thawing specimens, dehydration of thawed specimens, breakdown of cellular and protein structures and air bubbles in the homogenate. Despite this, the μ_a and μ'_s spectra for ground tissues were good approximations of the spectra for intact tissues. One prevailing observation was that ground tissues presented stronger oxy-hemoglobin peaks at 414, 540, and 577 nm than for intact tissues, likely due to the exposure of hemoglobin to air during the grinding process.

Average Refractive Index

A further measurement that can be made in intact bulk tissue is the (average) refractive index, by determining the angle of the emergent cone of light from an optical fiber with the cladding removed and the end aligned with the tissue surface [32]. This is illustrated in Fig. 8.5. Light is launched into the fiber through a diffuser in order to fill the mode volume of the fiber. This, and the use of a large-diameter fiber ($\geq 1000\ \mu$m), allow the application of simple geometric optics to calculate the tissue refractive index, n_t, as

$$n_t^2 = n_q^2 - (n_0 \sin\zeta)^2 \tag{8.6}$$

where n_q and n_o are the refractive indices of the (quartz) fiber core and air, respectively, and ζ is the half-angle of the emergent light cone. The half-angle was

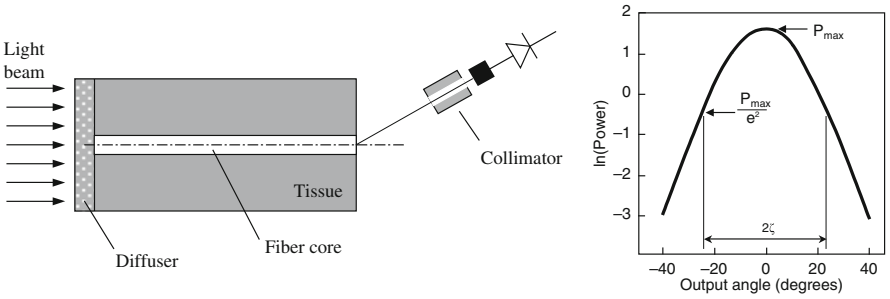


Fig. 8.5 Measurement of average refractive index of tissue using an embedded optical fiber core. The emergent half angle, ζ, is determined from the plot of detected light signal versus angle from the fiber axis. (Adapted with permission from Ref. [32])

empirically determined to span the emergent light cone intensity at a magnitude of $P_{\max}/e^2$, by optimizing the probe against several standards of known refractive index (sucrose solutions). At 632.8 nm (HeNe laser), typical refractive index values for various mammalian tissues have been measured in the range 1.37–1.45, decreasing by around 1% per 100 nm across the visible spectrum.

8.2.2 In Vivo Methods

Determination of the optical properties in vivo is fundamental for the effective use of light in diagnostic and therapeutic applications. Optical cancer diagnostics [33–36], online monitoring of laser interstitial thermal therapy [9], photodynamic therapy dosimetry [37–40], exogenous contrast detection [35, 41] and intrinsic fluorophore calculations [42] are all enabled or improved through accurate determination of μ_a and μ_s'.

Innovation in techniques to determine the tissue optical properties in vivo has been enhanced by recent technological advances in light sources, detection systems, fiber optics technology and computing power. As a result, there are many and varied techniques to measure and calculate μ_a and μ_s' in vivo. These may be categorized in terms of the measurement geometry used, namely reflectance and transmittance. Alternatively, the methods may be classed in terms of measurement quantity, such as intensity, fluence rate or radiance. Measurements can also be defined by the parameter being resolved to provide constraints that allow the separation of absorption and scattering; thus, measurements can be made spatially-, spectrally-, temporally- or angularly-resolved.

This section discusses the in vivo methods of determining optical properties, beginning with a discussion of the total (i.e. spatially-integrated) diffuse reflectance, moving on to a discussion of how μ_a and μ_s' can be retrieved using diffuse reflectance measurements with steady-state and time-dependent illumination.

The optical properties can be extracted from the steady-state diffuse reflectance using spatially-resolved or spectrally-constrained techniques. The time-dependent response of tissue can be measured with time-resolved or frequency-domain diffuse reflectance techniques. The fluence rate- and radiance-based methods will also be discussed, as well as a brief note on diffuse optical tomography. This section includes tables on selected instruments in the literature (Table 8.1), as well as a table on human tissue optical properties measured at various tissue sites in vivo (Table 8.2).

8.2.2.1 Total Steady-State Diffuse Reflectance

The total diffuse reflectance, R_d, is the fraction of the incident flux which is re-emitted through the irradiated surface. For an optically homogeneous, semi-infinite tissue volume irradiated by a collimated beam, R_d can be written as a function of μ_a, μ_s and g and, in the case of surface refractive index mismatch, also n. For high reduced albedo, $a' = \mu_s'/(\mu_a + \mu_s')$, (i.e. ≥ 0.5) the similarity relation for the reduced scattering coefficient, $\mu_s' = \mu_s(1 - g)$, is valid and there is no significant difference in R_d for tissues of different g and μ_s but the same μ_s'. The value of R_d then depends only on the reduced albedo, a', and refractive index, n. Clearly, it is not possible to determine μ_a and μ_s' separately from a measurement of R_d alone at a single wavelength of light.

Diffusion theory can be used to show [2] that

$$R_d = \frac{a'}{1 + 2\kappa\left(1 - a'\right) + \left(1 + \frac{2\kappa}{3}\right)\sqrt{3\left(1 - a'\right)}} \tag{8.7}$$

where $\kappa = \frac{1+r_{id}}{1-r_{id}}$ and r_{id} is the internal reflection coefficient for diffuse light. For index matching at the surface, $r_{id} = 0$ and $\kappa = 1$. For index mismatch, empirical relations between r_{id} and the relative index of refraction, $n_{\text{rel}} = n_{\text{tissue}}/n_{\text{exterior medium}}$, have been derived [43, 44]. For example, for $n_{\text{rel}} = 1.3$–1.5, which encompasses most soft tissues in the visible/near-infrared, and with an air-tissue boundary, $r_{id} = 0.46$–0.60 and $\kappa = 2.6$–4.0. The increase in R_d due to index matching is about $a' \sim 0.9$.

Monte Carlo modeling has been used to check the validity of Eq. (8.7). The agreement is good, except for $g \geq 0.99$ and $a' \leq 0.67$, where the diffusion theory values are significantly higher than the Monte Carlo predictions. Experimental checks in scattering and absorbing media with $a' > 0.95$ have also confirmed the diffusion model [45].

A major application of diffuse reflectance spectrophotometry has been the study of specific chromophores in tissue, such as hemoglobin. The fact that the total diffuse reflectance yields only the reduced albedo and not the separate scattering and absorption contributions limits the usefulness of these measurements, since the true tissue absorption spectrum cannot be obtained; however, assuming that the reduced scattering coefficient is constant, absorption ratios, either measured at different

Table 8.1 Tabulation of selected measurement of tissue optical properties in vivo, grouped by technique. The wavelength range, acquisition time and accuracy associated with each method are listed. The root-mean-square (rms) errors are given for the accuracy statistic when available; otherwise, the maximum error is cited. The variation in validity ranges as well as the use of different standards make a direct comparison of accuracies between techniques difficult, and should be interpreted with caution

Technique	Wavelength range	Acquisition time	Accuracy (Range of validity)	Accuracy standard	Human tissues investigated	References
Spatially-resolved, steady-state diffuse reflectance	543, 633 nm		$\mu_a : \pm 2\ \text{cm}^{-1}$ rms ($\mu_a = 1 - 25\ \text{cm}^{-1}$) $\mu_s' : \pm 3\ \text{cm}^{-1}$ rms ($\mu_s' = 5 - 25\ \text{cm}^{-1}$)	Intralipid-dye phantoms measured using adding-doubling		Pfefer et al. [53]
	600–900 nm	5 s	$\mu_a : \pm 40\%$ max ($\mu_a = 0.1 - 1.5\ \text{cm}^{-1}$)	Intralipid-dye phantoms with μ_a measured by spectrophotometer (s.p.m.)	Skin [52]	Doornbos et al. [52]
	632.8 nm		$\mu_a : \pm 4.7\%$ rms ($\mu_a = 0.2 - 0.6\ \text{cm}^{-1}$) $\mu_s' : \pm 2.4\%$ rms ($\mu_s' = 4 - 10\ \text{cm}^{-1}$)	Microsphere-dye phantoms with μ_a measured by s.p.m.; μ_s' calculated using Mie theory		Lin et al. [51]
	633, 751 nm	1 s	$\mu_a : \pm 14\%$ rms ($\mu_a = 0.02 - 1\ \text{cm}^{-1}$) $\mu_s' : \pm 2.6\%$ rms ($\mu_s' = 5 - 20\ \text{cm}^{-1}$)	Frequency-domain measurements of Intralipid-dye phantoms		Kienle et al. [47]

Table 8.1 (continued)

Technique	Wavelength range	Acquisition time	Accuracy (Range of validity)	Accuracy standard	Human tissues investigated	References
Spectrally-constrained, steady-state diffuse reflectance	500–750 nm	8–60 ms	μ_a : ±10% rms (μ_a = 0 − 8.5 cm^{-1}) μ_s' : ±5% rms (μ_s' = 9.2 − 17.5 cm^{-1})	Spatially-resolved reflectance measurements of TiO_2-dye phantoms, with dye μ_a measured by s.p.m.		Reif et al. [58]
	350–1000 nm	∼20 ms			Breast [86], bronchial mucosa [55]	Amelink et al. [59]
	500–900 nm	200 ms			Esophagus, lung, oral cavity [33]	Bargo et al. [33]
Time-resolved diffuse reflectance	660, 786, 916, 974 nm	20–60 s	μ_a : ±10% max (μ_a = 0 − 0.3 cm^{-1}) μ_s' : ±10% max (μ_s' = 9 − 20 cm^{-1})	Integrating sphere phantom measurements	Prostate [11], breast [66]	Svensson et al. [66]
	610–1010 nm	800 s (10 s per λ)	μ_a : ±10% max (μ_a = 0 − 0.3 cm^{-1}) μ_s' : ±10% max (μ_s' = 9 − 20 cm^{-1})	Integrating sphere phantom measurements	Breast [69]	Torricelli et al. [70]

Table 8.1 (continued)

Technique	Wavelength range	Acquisition time	Accuracy (Range of validity)	Accuracy standard	Human tissues investigated	References
Frequency-domain diffuse reflectance	674, 782, 849, 956 nm		μ_a : ±5% rms ($\mu_a = 0.05 - 0.5$ cm^{-1}) μ_s' : ±3% rms ($\mu_s' = 5 - 20$ cm^{-1})	Intralipid-dye phantoms with dye μ_a measured by s.p.m. and μ_s' estimated using Mie theory [87]	Breast [12], cervix [93]	Pham et al. [76] Tromberg et al. [12]
Spatially-resolved fluence rate	730 nm		μ_a : ±8% rms ($\mu_a = 0.1 - 1$ cm^{-1}) μ_s' : ±18% rms ($\mu_s' = 1.7 - 9.1$ cm^{-1})	Intralipid-dye measurements using a broad-beam, depth-resolved fluence rate measurement set-up	Prostate [88]	Dimofte et al. [88]

Table 8.2 Tabulation of human in vivo tissue optical properties for various tissue types, as well as other physiological properties such as blood volume (BV), hemoglobin concentration ([Hb]) and oxygen saturation (StO_2) that are derived from the absorption coefficient spectrum. The error bars represent standard deviations. Some of the listed optical properties measurements are from very few patients but are included because of the lack of more complete data sets

Tissue	λ (nm)	μ_a (cm^{-1})	μ_s' (cm^{-1})	Physiological properties	N_{patients}	Measurement interface	Technique	References
Esophagus	630	0.80 ± 0.23	7.7 ± 1.5	BV: 1.72 ± 0.93% StO_2: 54 ± 10%	9	Endoscope access, surface measurement	Spectrally-constrained diffuse reflectance	Bargo et al. [33]
	630	0.27 ± 0.14	7.0 ± 2.3		11	Endoscope access, surface measurement	Spatially-resolved diffuse reflectance	Bays et al. [89]
Breast (healthy)	785	0.041 ± 0.021	8.0 ± 2.0	[Hb]: 17 ± 10 μM StO_2: 77 ± 8%	36	Noninvasive surface (skin) measurement	Time resolved reflectance	Svensson et al. [66]
	637 683 785	0.055 ± 0.007 0.042 ± 0.013 0.037 ± 0.013	13.4 ± 2.6 12.9 ± 2.3 11.3 ± 2.1	[Hb]: 15.7 ± 5.1 μM StO_2: 66.4 ± 9.2%	50	Noninvasive surface (skin) measurement	Time resolved transmittance	Spinelli et al. [90]
	750 786 830	0.046 ± 0.024 0.041 ± 0.025 0.046 ± 0.027	8.7 ± 2.2 8.5 ± 2.1 8.3 ± 2.0	[Hb]: 34 ± 9 μM StO_2: 68 ± 8%	52	Noninvasive surface (skin) measurement	Frequency domain transmittance	Durduran et al. [7]

Table 8.2 (continued)

Tissue	λ (nm)	μ_a (cm^{-1})	μ_s' (cm^{-1})	Physiological properties	N_{patients}	Measurement interface	Technique	References
Breast (adipose)	650 1090	0.22 ± 0.18 0.42 ± 0.14	4.0 ± 1.2 2.7 ± 0.5		24	Intraoperative surgical margin measurements	Spatially-resolved diffuse reflectance	van Veen et al. [91]
Breast (tumor)	650 1090	0.32 ± 0.41 0.45 ± 0.17	4.5 ± 1.8 3.0 ± 0.9			As above	As above	As above
Prostate (cancerous, pre-treatment)	762	0.39 ± 0.18	3.37 ± 1.59		22	Interstitial	Spatially-resolved fluence rate	Weersink et al. [41]
	660 786 916	0.5 ± 0.1 0.4 ± 0.1 0.6 ± 0.1	8.7 ± 1.9 7.1 ± 1.6 7.7 ± 1.8	[Hb]: 215 ± 65 μM StO_2: 76 ± 4%	9	Interstitial	Time-resolved diffuse reflectance	Svensson et al. [11]
	732	0.37 ± 0.24	14 ± 11		13	Interstitial	Spatially-resolved fluence rate	Zhu et al. [92]
Cervix	674 811 849 956	0.19 ± 0.04 0.27 ±0.09 0.34 ± 0.10 0.57 ± 0.12	9.05 ± 1.53 5.58 ± 1.40 6.11 ± 1.16 4.98 ± 0.37		10	Non-invasive surface contact	Frequency-domain diffuse reflectance	Hornung et al. [93]
Brain tumor, astrocytoma				BV: 4.96 ± 2.56 mg/mL StO_2: 36 ± 21	6	Open surgical site, tissue contact	Spectrally-constrained diffuse reflectance	Asgari et al. [94]

Table 8.2 (continued)

Tissue	λ (nm)	μ_a (cm^{-1})	μ_s' (cm^{-1})	Physiological properties	N_{patients}	Measurement interface	Technique	References
Brain tumor, glioblastoma				BV:18.40±5.41 mg/mL StO_2: 52 ± 18	7	As above	As above	As above
Brain, white matter	674 849 956	2.5 0.95 0.9	13.5 8.5 7.8		1	Open surgical site, tissue contact	Spatially-resolved diffuse reflectance	Bevilacqua et al. [63]
Brain, temporal lobe cortex	674 811 849 956	0.2 0.2 <0.1 0.25	10 8.2 8.2 8.2		1	As above	As above	As above
Stomach (antrum)	542 600 650	3.6 ± 1.4 1.0 ± 0.7 0.6 ± 0.5	14 ± 3.5 13 ± 2.0 12 ± 1.0	[Hb]: 20 ± 2.6 μM StO_2: 48%	35	Gastroscope access, non-invasive probe contact	Spatially-resolved diffuse reflectance	Thueler et al. [95]

wavelengths or measured dynamically over time at a single wavelength, can be obtained. The total diffuse reflectance can be related to the tissue absorption coefficient by the relation

$$R_d \propto \exp\left(-\mu_a \langle l \rangle\right) \tag{8.8}$$

where $\langle l \rangle$ is the expected mean path length traveled by photons in the tissue. Thus, for two wavelengths λ_1 and λ_2 (or two time points at the same wavelength), if $\langle l_1 \rangle = \langle l_2 \rangle$, then the ratio of the absorption coefficients is given by

$$\frac{\mu_{a1}}{\mu_{a2}} = \frac{\ln R_{d2}}{\ln R_{d1}} \tag{8.9}$$

If one measurement is made at a wavelength where the absorption coefficient is known, then the absolute absorption coefficient at the other wavelength is determined. The use of the water absorption peak at about 970 nm has been proposed for this purpose, since the water content of soft tissues is known. Although μ_a (blood at 970 nm) > μ_a (water at 970 nm), the blood content of most tissues is less than a few percent, so that water is still the dominant chromophore. This method of "normalization" becomes decreasingly accurate for wavelengths less than 970 nm, due to increasing scatter, which alters $\langle l \rangle$.

The total diffuse reflectance can be measured either by using an integrating sphere in contact with the tissue [33], as shown in Fig. 8.6a, or by a detector with wide collection optics offset from the surface (assuming a Lambertian surface), [46] as in Fig. 8.6b. Detection of specular reflection must either be avoided or its contribution to the signal must be subtracted. The diffuse reflectance signal may be calibrated against a reflectance standard for the given source and detector geometry.

It can be seen from Eq. (8.8) that the limitation in determining the true tissue absorption from the total diffuse reflectance measurement is that the mean optical path,$\langle l \rangle$, which depends on the tissue scattering, is unknown a priori. Additional constraints are required to either directly or indirectly resolve this parameter to find μ_a and μ_s'. The following sections demonstrate how the spatial or temporal dependence of the reflectance, as well as spectral constraints, can be used to separate the absorption and scattering.

8.2.2.2 Spatially-Resolved, Steady-State Diffuse Reflectance

The spatially-resolved, steady-state local diffuse reflectance, $R(r)$, measured as a function of the radial distance on the surface, r, from a point or pencil-beam source, may be used to determine both μ_a and μ_s'. The $R(r)$ curve can be measured using fiber optics placed at varying radial distances from a source fiber, or by imaging the diffuse reflectance pattern, as shown in Fig. 8.7.

One method of determining μ_a and μ_s' is by using the spatially-resolved $R(r)$ curve in combination with a measurement of the total diffuse reflectance. If the terminal (logarithmic) slope, S, is defined as

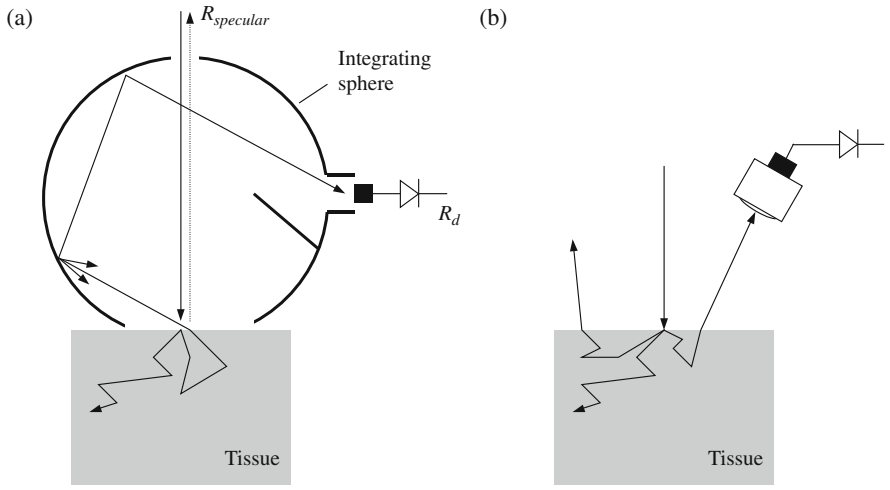


Fig. 8.6 Techniques for measuring total diffuse reflectance from tissue: (**a**) integrating sphere measurement, where the specular reflectance exits through the input port. (**b**) a "distant" detector measures a fraction of the diffusely reflected light over an area of the illuminated surface

$$S = \lim_{r \to \infty} \frac{\partial \left[\ln \left(r^m R_d(r) \right) \right]}{\partial r} \tag{8.10}$$

where m is a constant, then it can be shown [44, 48] that $S \sim \mu_{\text{eff}}$. The value of m has been variously cited as $^1/_2$, 1 or 2, depending on the model used and range of r values over which the measurement is made. For large $r(\gg 1/\mu_s')$, diffusion theory predicts $m = 2$ and this has been confirmed both by Monte Carlo modeling

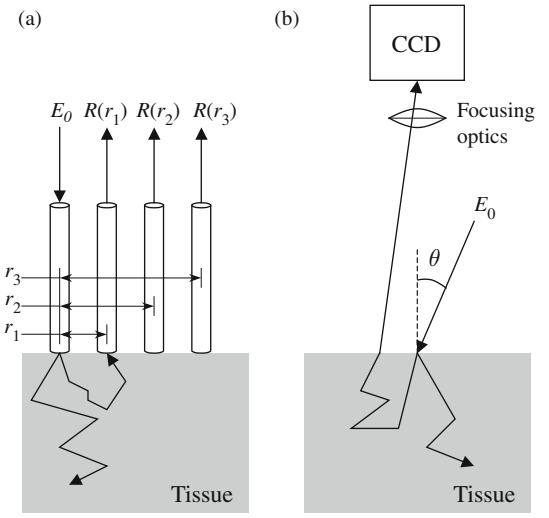


Fig. 8.7 Spatially-resolved diffuse reflectance measurements using (**a**) fiber optics in contact with the tissue surface located at varying radial distances from the source, and (**b**) imaging the diffuse reflectance spread of a single wavelength pencil beam (as in Kienle et al. [47], the light source is canted at a small angle to avoid detecting the specular reflectance)

and by measurements in tissues and phantoms. The measured value of μ_{eff} may then be combined with that of the reduced albedo, a', obtained as above from the total diffuse reflectance in Eq. (8.7), to give both μ_a and μ_s':

$$\mu_a = \mu_{\text{eff}} \left[\frac{1 - a'}{3} \right]^{1/2} \tag{8.11a}$$

$$\mu_s' = \mu_{\text{eff}} a' \left[\frac{1}{3\left(1 - a'\right)} \right]^{1/2} \tag{8.11b}$$

A common method of determining optical properties using spatially-resolved data is by fitting the shape of the measured $R(r)$ curve to an analytical equation derived from diffusion theory, with the optical properties μ_a and μ_s' as free parameters. In this way, a total diffuse reflectance measurement is not required. The correct handling of boundary conditions in modeling $R(r)$ is critical in using the diffusion theory approach. The most commonly used method is to start with the expression for the fluence rate distribution for an isotropic point source (for unit optical power) in an infinite, optically homogeneous medium:

$$\varphi(\rho) = \frac{e^{-\mu_{\text{eff}}\rho}}{4\pi\rho D} \tag{8.12}$$

where ρ is the distance from the source and D is the diffusion constant. A pencil beam incident on a semi-infinite tissue surface (Fig. 8.8a) may then be approximated by an isotropic point source located at depth z_0 below the surface (Fig. 8.8b). The boundary conditions at the surface can be approximated by setting the fluence rate to zero at the physical tissue surface. This is achieved with a second, "negative" point source located as a mirror image of the first real source and summing the fluences, as given by Eq. (8.12), from each, as illustrated in Fig. 8.8c. This source pair is equivalent to a dipole source. A better boundary approximation can be obtained by setting the zero-fluence boundary, not at the tissue surface, but at a certain height, z_b, above the surface (Fig. 8.8d) or by weighting the negative point source located outside the tissue surface. The dipole strength may be obtained in various ways. For example, Allen and McKenzie [49] have equated the fluence at depth from a sheet of such dipoles to that from a broad collimated beam in order to calculate the dipole strength, while Patterson et al. have assumed a fixed dipole separation of $2/\mu_s'$ [50]. Similarly, there are different approximations to obtain the position of the "extrapolated" boundary. The derived values for μ_a and μ_s' were within $\pm 10\%$ of the true values, fitting the data to the form [3]

$$R(r) = \frac{1}{4\pi}\left[z_0 \left(\mu_{\text{eff}} + \frac{1}{\rho_1}\right) \frac{e^{-\mu_{\text{eff}}\rho_1}}{\rho_1^2} + (z_0 + 2z_b)\left(\mu_{\text{eff}} + \frac{1}{\rho_2}\right) \frac{e^{-\mu_{\text{eff}}\rho_2}}{\rho_2^2} \right] \tag{8.13}$$

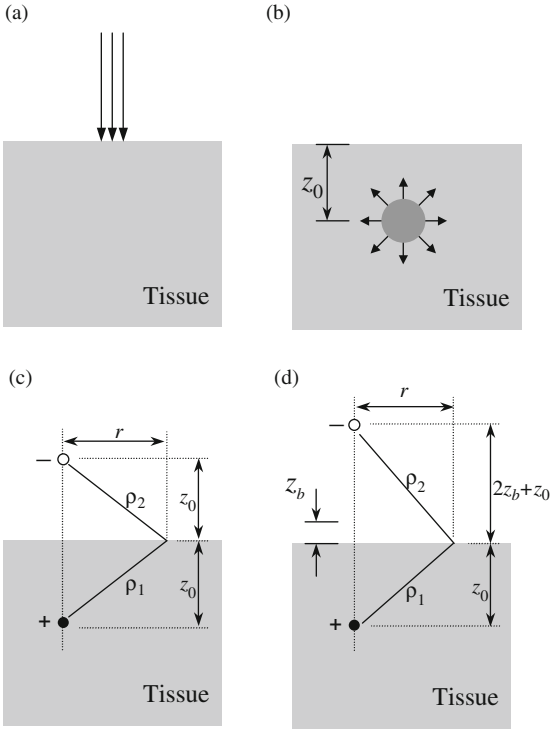


Fig. 8.8 Illustration of the equivalent dipole and extrapolated boundary assumptions used in deriving the diffusion theory expressions for $R(r)$: (**a**) incident collimated beam, (**b**) equivalent first-scatter source comprising a single source at depth z_0, (**c**) summation of positive and negative point sources, or (**d**) extrapolated boundary at distance z_b above the physical tissue surface. μ_a and μ_s' are adjusted until measured reflectance $R(r)$ matches Eq. (8.13)

where $\rho_1^2 = z_0^2 + r^2$ and $\rho_2^2 = (z_0 + 2z_b)^2 + r^2$. The z_b factor depends on μ_a, μ_s' and the internal reflection parameter $\kappa = (1 + r_{id}) / (1 - r_{id})$ (due to index mismatch between tissue and the external medium). The isotropic point source depth may be approximated as $z_0 = 1/(\mu_a + \mu_s')$. The extrapolated boundary distance is given by $z_b = 2\kappa D$, where D is the diffusion constant given by $D = \left[3\left(\mu_a + \mu_s'\right)\right]^{-1}$ [3]. An empirical formulation of r_{id} for index-mismatched boundaries has been widely used, where $r_{id} = -1.44n_{\text{rel}}^{-2} + 0.71n_{\text{rel}}^{-1} + 0.67 + 0.0636n_{\text{rel}}$ [43, 44]. Recall that $\kappa = 1$ for matching internal and external refractive indices. An important practical advantage of using the whole $R(r)$ curve to derive μ_a and μ_s' is that absolute reflectance measurements are not required, since only the *shape* of the curve is determined. This simplifies the measurements considerably. Details and development of Eq. (8.13) are provided in Chapter 6.

Figure 8.9 shows the diffusion approximation for reflectance compared with Monte Carlo simulations for three media, all three having the same scattering properties $\mu_s' = 10\ \text{cm}^{-1}$ and $g = 0.8$ using the Henyey-Greenstein phase function), with μ_a values of 0.1, 1 and 2 cm^{-1}. The Monte Carlo simulation was modeled with matching boundaries, such that the internal reflection parameter was taken as $\kappa = 1$. The extrapolated boundary was, therefore, $z_b = 0.066$, 0.061 and 0.056 cm for the three media with $\mu_a = 0.1$, 1 and 2 cm^{-1}, respectively. Although accurate for r values larger than ~1 mm, the diffusion approximation breaks down

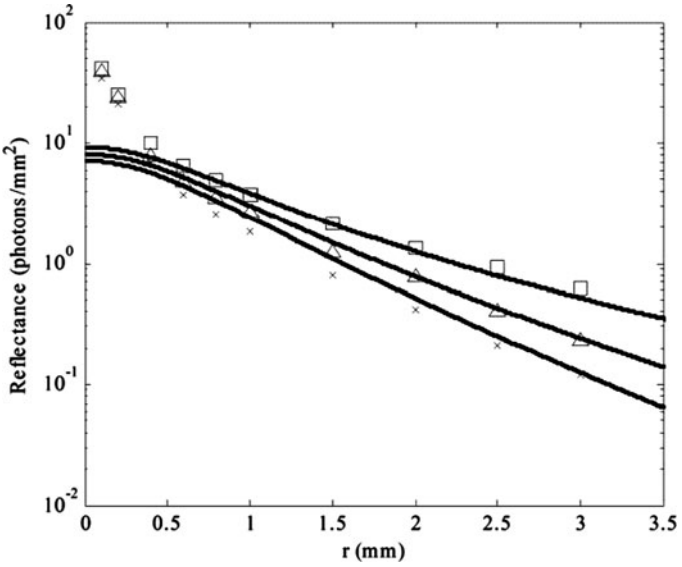


Fig. 8.9 Comparison of diffusion theory predictions of $R(r)$ (*solid lines*) with Monte Carlo simulations (*symbols*). The simulated medium has $\mu_s' = 10\ \text{cm}^{-1}$ and $g = 0.8$ for all three curves, with μ_a varied from 0.1, 1 and 2 cm^{-1}, going from the *top* curve to the *bottom*

with reflectance collected near the source. It is important to note that this diffusion approximation breakdown occurs at shorter source-collector distances with increased scattering events prior to detection, i.e., with increasing μ_s'.

A refinement of this technique is to use an obliquely incident source at an angle, θ_{tissue}, resulting in a diffuse reflectance pattern that is not centered around the point of incidence [51]. The offset of the center of the reflectance pattern, Δx, can be related directly to the diffusion coefficient, D. The effective attenuation coefficient, μ_{eff}, can then be retrieved through least-squares minimization using Eq. (8.13), and the optical properties can then be calculated from D and μ_{eff}.

Doornbos et al. developed a system to measure tissue optical properties by recovering the $R(r)$ curve using nine optical fibers located at varying distances from the source (from 1.75 to 16.79 mm) [52]. Although μ_a and μ_s' were recoverable for low absorption by fitting $R(r)$ to the diffusion model, the method failed at high absorption due to hemoglobin in the UV range. As well, since this technique relied upon absolute measurements of $R(r)$, calibration of the diffusion model to measurements was a challenge. Pfefer et al. also used a multi-fiber probe to determine optical properties by measuring the reflectance at varying distances (from 0.23 to 2.46 mm) from the source [53]. Rather than curve-fitting the $R(r)$ data to a diffusion theory model, the authors used neural network algorithms and a partial least squares technique to recover the optical properties at a single wavelength.

Kienle et al. developed an imaging technique to measure the diffuse reflectance pattern from an incident laser beam [47]. The $R(r)$ curve was recovered by radially binning the image pixels. Imaging techniques offer near-continuous resolution of

diffuse reflectance from the source; also, the non-contact implementation eliminates the problem of optical property distortion from pressure due to contact probes.

Recently, there has been work by Cuccia et al. [10] on using the spatial frequency-domain to determine tissue optical properties. The spatially-resolved reflectance can be considered to be the spatial impulse response of the tissue. The spatial frequency response can be measured by illuminating the tissue with a plane source, E, normally incident on the tissue surface that is spatially modulated in one direction:

$$E\left(f_x\right) = E_0\left[1/2 + M\sin\left(2\pi f_x x\right)\right] \tag{8.14}$$

where f_x is the spatial frequency along the x direction (in the vertical direction along the image in this case, though the direction is arbitrary), M is the modulation amplitude and E_0 is the source strength. Applying this method to a homogenous medium, the tissue optical properties can be fitted to the Fourier-transformed diffusion theory equation [3] for spatially-resolved reflectance, giving estimates of μ_a and μ_s' that show agreement within 10% of the values determined by temporal frequency-domain diffuse reflectance (see Section 8.2.2.4). In principle, the optical properties can be recovered pixel by pixel over the field of view of the image. One of the additional benefits to this method is that the light penetration depth, $1/\mu_{\text{eff}}$, is related to the spatial frequency f_x. This allows for depth-resolved encoding by controlling the spatial frequency. By sampling at various frequencies, tomographic images can be reconstructed based on the optical contrast of buried objects.

8.2.2.3 Spectrally-Constrained, Steady-State Diffuse Reflectance

One of the more popular methods of the past decade to determine optical properties is to use a single fiberoptic source-collector pair to measure the steady-state diffuse reflectance spectrum, as shown in Fig. 8.10. The source fiber delivers broadband white light in the spectral range of interest and the diffuse reflectance spectrum is detected by the collector fiber located at a radial distance, r, from the source.

Since there is only one reflectance measurement per wavelength, solving for $\mu_a(\lambda)$ and $\mu_s'(\lambda)$ relies upon spectral constraint, that is, using a priori knowledge of the shapes of the absorption and scattering coefficient spectra in the forward model. The absorption spectrum can be modeled as a linear combination of the separate chromophore contributions. Here, it is expressed using water content, total hemoglobin concentration and an oxygen saturation term:

$$\mu_a\left(\lambda\right) = c_{\text{water}}\mu_a^{\text{water}}\left(\lambda\right) + c_{Hb}\left[StO_2\mu_a^{\text{oxy}Hb}\left(\lambda\right) + \left(1 - StO_2\right)\mu_a^{\text{deoxy}Hb}\left(\lambda\right)\right] \tag{8.15}$$

where $\mu_a^{\text{water}}\left(\lambda\right)$, $\mu_a^{\text{oxy}Hb}$, and $\mu_a^{\text{deoxy}Hb}$ and are the wavelength-dependent absorption coefficients of water, oxygenated hemoglobin, and deoxygenated hemoglobin, respectively. c_{water} and c_{Hb} are the water and total hemoglobin volume fraction and StO_2 is the oxygenation fraction.

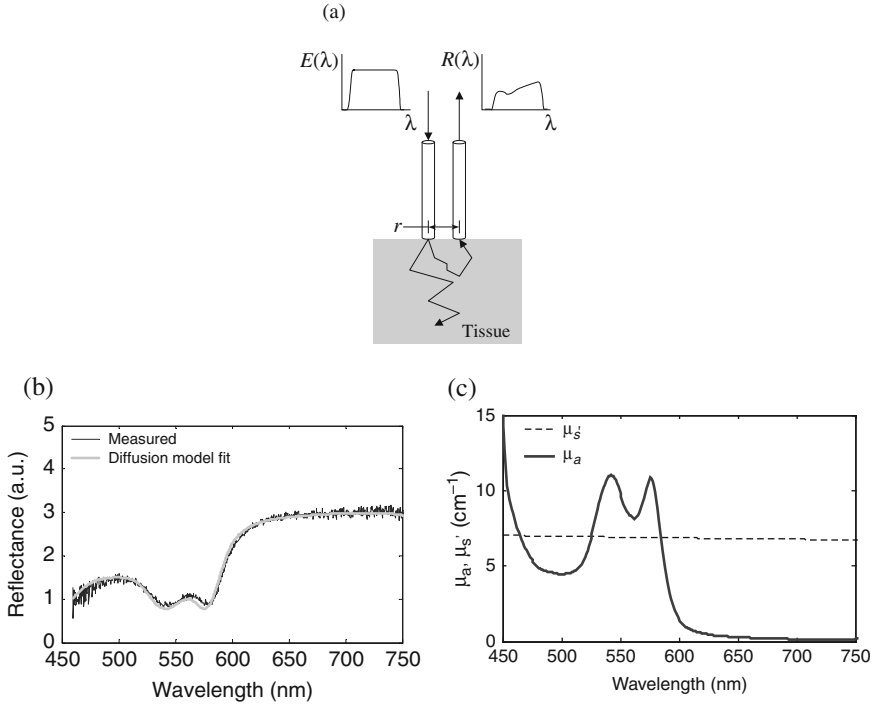


Fig. 8.10 (**a**) Single source-collector pair for measuring the diffuse reflectance spectrum. A broadband white light source is used to illuminate the tissue. A collector fiber connected to a spectrometer records the diffuse reflectance spectrum. (**b**) Diffuse reflectance spectrum measured at the surface of rat muscle in vivo, with a fiber separation distance of 780 μm. The corresponding diffusion theory and spectral constraint models in Eqs. (8.15–8.17) were used to fit the data to solve for the free parameters (c_{Hb}, StO_2, a and b) in the model using a Levenberg-Marquardt least squares algorithm. (**c**) Optical properties spectra of the muscle tissue, calculated from the free parameters extracted from the diffuse reflectance spectrum

Care must be taken to include only the measurable chromophores in order to have an accurate absorption model. Omitting strongly-absorbing chromophores may result in fitting errors. For example, beta carotene is present in measurable amounts in adipose breast tissue and should be included in clinical breast studies if the reflectance spectrum is measured in the range of the 450 nm beta carotene peak [36]. As well, insignificant chromophores should be judiciously excluded to make the model fitting more robust. Water absorption below ~700 nm is negligible compared to that of hemoglobin and may not be needed in model fits; however, in the near-infrared water becomes a formidable component. There have also been some estimations of the absorption of dry, bloodless tissue as a mono-exponential spectral shape [33] that may be used in the absorption model.

The reduced scattering coefficient spectrum from bulk tissue has been shown to fit well to a simple wavelength-dependent power law [52, 57, 58] given by

$$\mu_s'(\lambda) = a\lambda^{-b} \tag{8.16}$$

where a is referred to as the scattering magnitude and b as the scattering power.

The a priori knowledge of the chromophore and scatterer spectra are then typically combined in a forward model of the diffuse reflectance and a curve-fitting algorithm is applied to extract the free parameters. A simple approach to develop a forward model is to use the diffusion theory equation for spatially-resolved, steady-state diffuse reflectance [59, 60]. Here, the radial distance, r, is fixed and the optical properties $\mu_a(\lambda)$ and $\mu_s'(\lambda)$ vary according to wavelength:

$$R(\lambda) = \frac{1}{4\pi}\left[z_0\left(\mu_{\text{eff}} + \frac{1}{\rho_1}\right)\frac{e^{-\mu_{\text{eff}}\rho_1}}{\rho_1^2} + (z_0 + 2z_b)\left(\mu_{\text{eff}} + \frac{1}{\rho_2}\right)\frac{e^{-\mu_{\text{eff}}\rho_2}}{\rho_2^2}\right] \tag{8.17}$$

where z_0, μ_{eff}, ρ_1, ρ_2 and z_b are wavelength-dependent. $\mu_{\text{eff}}(\lambda) = \sqrt{3\mu_a(\lambda)\,[\mu_a(\lambda) + \mu_s'(\lambda)]}$, and $\mu_a(\lambda)$ and $\mu_s'(\lambda)$ are given by Eqs. (8.15) and (8.16). The free parameters are, therefore, the absorber concentrations, oxygen saturation and scattering parameters. This approach has been used successfully for probes with large source-collector distances, where the diffusion approximation is accurate. The minimum limit on r has been variously cited by different authors. Farrell et al. [3] found that diffusion theory agreed with Monte Carlo simulations with source-collector separations greater than 0.5 mm, and with experimental phantoms for distances greater than 1 mm. A similar Monte Carlo comparison to the diffusion equation was carried out by Sun et al. [60], who tested source-collector distances between 0.5–3 mm and found the tissue optical properties to be resolvable within this range, as confirmed by experimental phantoms. An example using this spectrally-constrained, diffusion theory-based technique is shown in Fig. 8.10b and c, with a measurement taken in vivo in rat muscle using a 780 μm fiber separation. Here, water was considered to have negligible absorption in the range 450–750 nm. The free parameter values extracted for this sample were $c_{Hb} = 6.0$ g/L, $StO_2 = 71.3\%$, $a = 11.9$ and $b = 0.087$.

For close-packed fiber probes with small source-collector distances, an alternative approach for a forward model is required, since diffusion theory predictions breaks down under these conditions. The Monte Carlo technique is suitable for modeling probes with close source-collector distances. Even with the accuracy of Monte Carlo simulations there may be problems with violation of the similarity relation governing the reduced scattering coefficient, i.e. for the same μ_s', reflectance measurements may vary significantly for different scattering phase functions. Monte Carlo simulations have been used to check the validity of the μ_s' approximation for the reflectance geometry. With $r = 250\,\mu\text{m}$ and at low μ_s' values, a change in anisotropy can result in significant change in the reflectance: e.g. at $\mu_s' = 5\ \text{cm}^{-1}$, the change in reflectance from $g = 0.2$ to 0.99 is close to 100% [58]. Figure 8.11 shows Monte Carlo data on the breakdown of the μ_s' approximation using a simulated medium of $\mu_s' = 10\ \text{cm}^{-1}$, $\mu_a = 0.1\ \text{cm}^{-1}$, with varying g values given by the Henyey-Greenstein phase function.

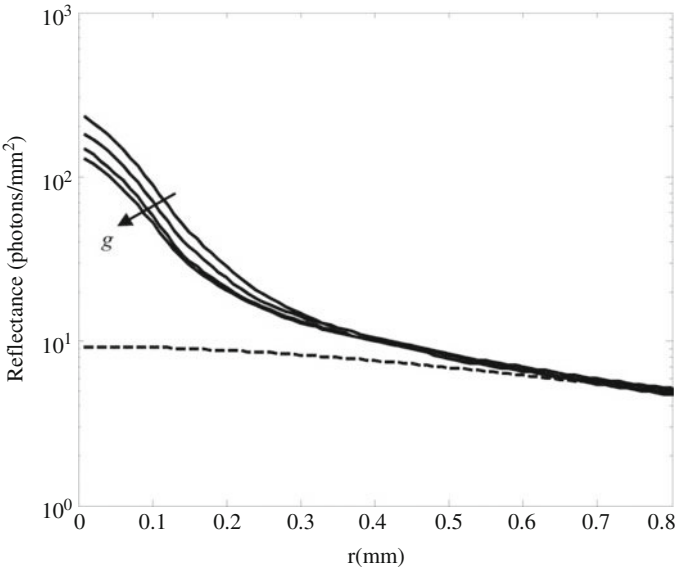


Fig. 8.11 Monte Carlo simulations demonstrating the breakdown of the Similarity Principle at close fiber distances. In a turbid medium ($\mu_s' = 10\ \text{cm}^{-1},\ \mu_a = 0.1\ \text{cm}^{-1}$), the reduced scattering coefficient, $\mu_s' = \mu_s(1-g)$, was held constant, while g had values of 0.4, 0.6, 0.8, and 0.9. The results are plotted versus the fiber separation distance, r, with the *arrow* indicating the direction of increasing g. The diffusion approximation (*dotted line*) is also shown for comparison, using the same z_b parameter as in Fig. 8.9

An empirical approach has been developed using phantoms to create data for a forward transport model [33]. Tissue-simulating acrylamide gel phantoms were formulated using Intralipid to provide optical scattering and India ink for absorption. An 8 x 8 set of gel phantoms were prepared, with all combinations of eight values of μ_s' ranging from 1 to 28 cm^{-1} and eight values of μ_a ranging from 0.01 to 6.4 cm^{-1}. Each of the 64 gels was measured with a reflectance probe to produce an empirical forward light transport map relating optical properties to the reflectance measurements. Analytical expressions were developed from this transport map and combined with a priori absorption and scattering spectral shapes in order to extract $\mu_a(\lambda)$ and $\mu_s'(\lambda)$. Although this empirical approach bypasses the need for theoretical modeling (using either Monte Carlo or diffusion theory) in order to estimate the optical properties, there is still a need for theoretical modeling to examine how the model breaks down at either low μ_s' and/or close fiber distances.

The main benefit to the spectrally-constrained, steady-state diffuse reflectance approach is that only a single fiberoptic source-collector pair is needed for measurement. With the widespread availability of commercial fiberoptic spectrometers and white light sources, these systems are technologically the least complex of all the in vivo techniques, with the added capability of obtaining rapid (<1 s) measurements. The main drawback is that the shape of the absorption and scattering spectral curves must be known a priori. This may be an issue for chromophores with ill-defined

spectral shapes, such as indocyanine green, a commonly used contrast dye for which the shape of the absorption spectrum changes at high concentrations [61], or for studies of the skin where melanin can have a variety of different absorption spectra. In addition, this technique does not account for variations in tissue optical properties of layered media, such as in the skin or the mucosal lining of hollow organs.

In the case of a single source-collector pair, spectral constraint is necessary to determine the optical properties; however, a priori models of the absorption and scattering spectra can also be used in other methods to determine optical properties in order to make them more robust, such as in diffuse optical tomography [62], frequency-domain reflectance [7] or spatially-resolved, steady-state reflectance [52].

8.2.2.4 Time-Resolved and Frequency-Domain Diffuse Reflectance

Time-resolved techniques essentially measure the impulse response of tissue, where the photon "time-of-flight" is measured directly following an input pulse of light; the equivalent measurements in the frequency-domain rely on determining the phase shift and amplitude modulation of the detected light signal with reference to an incident high-frequency modulated beam.

The speed of light in tissue is given by $c' = c_0/n$, where c_0 is the speed *in vacuo* and n is the average refractive index, which, for soft tissues in the red/near-IR, is in the narrow range around 1.38–1.41 [32, 54]. A constant refractive index of $n = 1.4$ is commonly used for light propagation calculations [34, 63, 64], giving $c' = 0.21$ mm/ps or 21 cm/ns. It is useful to express the speed in these units since time-domain measurements are usually made on ps or ns time scales. Consider, for example, source and detector fibers on a tissue surface separated by a distance of $r =$ 10 mm. Following a short (∼ps) light pulse, the minimum time delay for photons to reach the detector is about 48 ps. In reality, however, the arrival of photons is significantly spread out in time due to the different optical path lengths taken in traveling from source to detector, with the photons undergoing multiple scattering (see Fig. 8.12). The average time delay is given by:

$$\langle t \rangle = \frac{\langle l \rangle}{c'} \tag{8.18}$$

where $\langle l \rangle$ is the average path length.

For $r \gg 1/\mu_s'$, time-dependent diffusion theory for a semi-infinite, optically homogeneous medium predicts [50] that the local diffuse reflectance signal, $R(r, t)$, is given by

$$R(r, t) = \left(4\pi Dc'\right)^{-3/2} z_0 t^{-5/2} \exp\left(-\mu_a c' t\right) \exp\left[-\frac{\left(r^2 + z_0^2\right)}{4Dc't}\right] \tag{8.19}$$

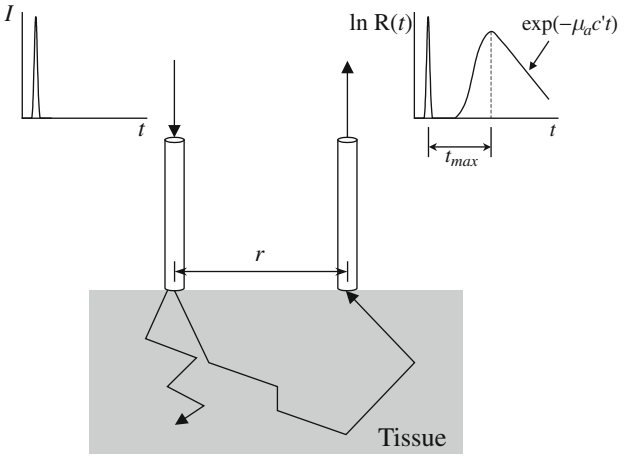


Fig. 8.12 Illustration of the principle of time-resolved reflectance measurements showing temporal pulse broadening due to scattering in tissue

It can be shown from Eq. (8.19) that

$$\mu_a = \left(\frac{-1}{c'}\right) \lim_{t\to\infty} \frac{\partial \left[\ln \left(R\left(r,\, t\right)\right)\right]}{\partial t} \tag{8.20}$$

That is, the absorption coefficient is given directly by the final logarithmic slope of the local time-reflectance curve, independent of r and μ_s'. The transport scattering coefficient can then be calculated from the time-to-maximum of the reflectance curve:

$$\mu_s' = \frac{1}{3r^2}\left(4\mu_a \left(c' t_{\max}\right)^2 + 10 c' t_{\max}\right) - \mu_a \tag{8.21}$$

Tests in phantoms have confirmed these predictions to an accuracy of better than 10% for tissue-like media of high albedo [65].

Figure 8.13 shows an example of experimental time-resolved photon migration exemplified by an interstitial prostate measurement [11] (786 nm, 15 mm fiber separation) and a non-invasive breast reflectance measurement [66] (786 nm, 20 mm fiber separation). While the time-resolved diffusion approximation (Eqs. (8.19)–(8.21)) can be employed for the evaluation of the breast tissue measurement, more refined modeling was required for the low albedo case of prostate tissue. To resolve the prostate optical properties, a Monte Carlo scheme was used as the forward model [56]. An important aspect of evaluating time-resolved measurements is that the instrument response function (IRF) must be measured and deconvolved from the tissue response curve.

The benefits of time-resolved measurements are that, to the extent that the diffusion approximation is accurate, the optical properties can be extracted using

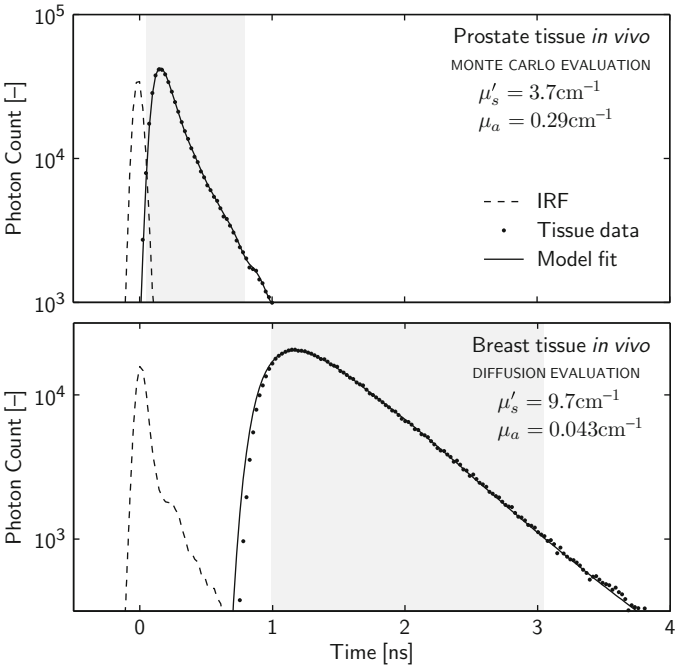


Fig. 8.13 In vivo time-resolved interstitial prostate measurement (*top*) and a non-invasive breast reflectance measurement (*bottom*). The instrument response function (IRF, *dashed lines*), tissue data (*dots*) and the model fits (*solid lines*) are displayed on the same time scale. The grey-shaded areas were used for the curve-fitting. (Courtesy T Svensson, E Alerstam, S Andersson-Engels, Lund University)

direct metrics (final logarithmic slope and time-to-maximum). The main drawback to time-resolved methods are that the technology required to measure and generate light pulses in the range of ps and ns are more costly and complex than steady-state systems.

Many time-resolved studies have used the time-correlated single photon counting (TCSPC) method to measure the $R(t)$ curve [50, 67–69]. Briefly, this employs the fact that, for low intensity pulses, the probability that multiple photons reach the detector is extremely small, so that any signal detected can be considered a single photon. The light source emits several pulses, and the detection system records the time whenever a single photon is detected. The $R(t)$ curve is essentially the histogram of these detection times. The time resolution is constrained by the timing accuracy of the detector electronics; the result is that modern time-resolved reflectance systems using TCSPC have time resolutions down to tens of picoseconds.

Time-resolved systems can be designed to measure a spectral range of optical properties. This can be achieved using either a wavelength-tunable laser (e.g. dye lasers or Ti:Sapphire) or separate lasers emitting at different wavelengths, such as laser diodes. Torricelli et al. have automated the time-resolved technique to a large

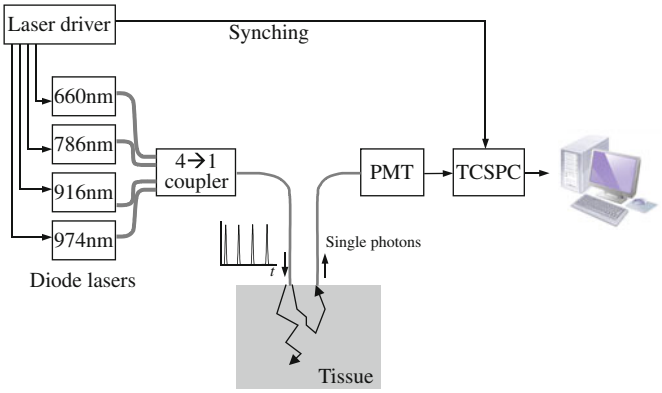


Fig. 8.14 Time-resolved, multi-spectral reflectance instrument. A laser driver controls 4 pulsed diode lasers, with the individual laser pulses separated in time by differing lengths of electrical cabling. The laser diodes send light through 4 optical fibers combined in a 4-into-1 coupler. A source fiber delivers the excitation into the tissue, with a collector fiber coupled to a photomultiplier tube (PMT) detecting single photons and counted using time-correlated single photon counting (TCSPC) synchronized with the laser driver. (Adapted with permission from Ref. [11])

spectral range, 610–1010 nm, with a spectral resolution of 5 nm [70]. Two lasers were used for the excitation source: a synchronously pumped mode-locked dye laser for the 610–695 nm range and an actively mode-locked Ti:sapphire laser for the 700–1010 nm range. Laser diodes can also be used in multi-spectral time-resolved systems. In one system, shown in Fig. 8.14, a laser driver controls four pulsed diode lasers at 660, 786, 916, and 974 nm [11]. These wavelengths were chosen to enable monitoring of the important tissue constituents: hemoglobin, water, and lipids and tissue oxygenation. The present detriments to multi-spectral, time-resolved measurements are the long acquisition times, in the range of ∼1–10 min, as well as the technological complications of providing excitation light and detecting the response at extremely short time slices (∼ps to ns).

Time-resolved measurements have also been used to investigate optical properties of multi-layered media. Since photons that penetrate deeper into tissue have longer flight times than shallower photons, depth-resolved information is encoded into the time-resolved data. Layered media are of particular interest in biomedical optics since several tissues are structured in this way, such as the lining of hollow organs and the skin. A solution to the 2-layer and 3-layer diffusion equation has been presented by Martelli et al. [71]. Experimental work to fit the absorption and scattering coefficients of a 2-layered phantom resulted in good accuracy of the μ_a estimation for both layers, as well as μ_s' of the first layer; however, μ_s' of the second layer and the first layer thickness had large estimation errors [68]. A 2-layer diffusion equation was fitted to 3-layer Monte Carlo simulations in order to investigate the optical properties of muscle beneath layers of subcutaneous fat and skin [72]. Using the multi-layered model and a priori knowledge of the fat layer thickness, the estimation of the muscle absorption coefficient had an error less than 5%; using a

homogeneous model as an approximation, the error was 55%. The thicknesses of the layers in this case were measured by ultrasound imaging.

The analogous frequency-domain expressions for the phase shift, Φ, and amplitude modulation, M, for the local reflectance in the semi-infinite geometry have been obtained by Fourier transformation of the time-domain diffusion theory formula [64]. The dependence of Φ and M on the modulation frequency, f, is shown in Fig. 8.15, for various values of μ_a and μ_s'. The phase shift increases monotonically with frequency and, at any given frequency, increases with increasing scatter or decreasing absorption. The modulation is unity at low frequency and falls to zero over a limited frequency range, depending on the absorption and scatter. Unlike time-resolved reflectance, frequency-domain diffusion expressions do not have closed-form expressions to obtain μ_a and μ_s' from measurable parameters (in this case, Φ and M). Hence, resolving the optical properties relies upon solving an inverse problem, given the measurements of Φ and M, and sometimes a steady-state (i.e. a DC, or zero-frequency) measurement [73, 74].

At low frequency, $2\pi f \ll \mu_a c', M = 1$, and the phase shift is given approximately by

$$\Phi = \frac{\pi\,(\mu_{\text{eff}} r_0)^2 f}{\mu_a c'\,(1 + \mu_{\text{eff}} r_0)} \tag{8.22}$$

where $r_0^2 = r^2 + z_0^2$ and $z_0 = 1/\mu_s'$. Working at low frequency, say less than about 200 MHz, has significant technological advantages, since low-cost modulated sources such as laser diodes or even arc lamps (with an external modulator such as a Pockel's cell) can be used, the latter allowing the possibility of continuous spectral scanning. However, at low frequency, Φ depends on both μ_{eff} and

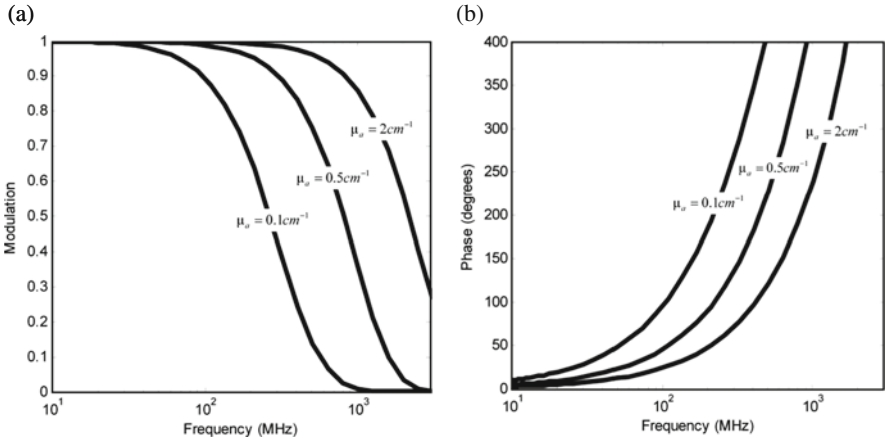


Fig. 8.15 Frequency-domain response for (**a**) modulation and (**b**) phase shift, both in the reflectance geometry and as calculated using the diffusion theory approximation. The scattering coefficient was held at $\mu_s' = 20\ \text{cm}^{-1}$ and the absorption coefficient varied as $\mu_a = 0.1$, 0.5 and $2\ \text{cm}^{-1}$

μ_a, so that a measurement of Φ alone does not allow separation of the absorption and transport scattering coefficients. Even making measurements at more than one (low) frequency does not help, since Φ is proportional to f. A possible solution [75] is to combine a low-frequency phase measurement with a steady-state measurement, such as μ_{eff} determined from the terminal slope of $R(r)$ versus r, as discussed above.

At higher frequencies, a measurement of Φ and $M < 1$ allows both μ_a and μ_s' to be determined. A particular advantage of the frequency-domain is that it is possible to vary f so that boundary effects due to finite tissue volume do not affect the results; however, in tissue this requires high frequencies, typically in the range $\sim$1.5 GHz, depending on the geometry and optical properties (see Fig. 8.15).

Many frequency-domain systems operate over a range of frequencies in order to better resolve the optical properties. Input modulation can be provided by any swept-source electrical device such as a frequency synthesizer [73] or network analyzer [12] that can provide ranges between a few hundred MHz and 1 GHz. As with time-resolved methods, frequency-domain systems can acquire multi-spectral measurements by using a multiplexed array of laser diodes at different wavelengths [12].

Detection methods need to be both phase-sensitive and able to measure the amplitude modulation. Two phase-locked frequency generators have been used, with one generator driving a diode laser delivering light to the tissue, and one generator driving an image intensifier coupled to a CCD camera to detect reflected photons [73]. The two generators modulated the light source and detector at the same frequency at a programmable phase offset to each other. Here, the phase offset between the two frequency generators was varied across 2π, and a fast-Fourier transform was applied to the data to recover Φ and M. A network analyzer has also been used to receive the modulated response measured by a photodiode (coupled to the collector fiber) to measure the phase and modulation amplitude with reference to the input signal (also from the network analyzer), an example is shown schematically in Fig. 8.16 [12, 76].

As with time-resolved methods, the optical properties of multi-layered media have been investigated using the frequency-domain. A 2-layer phantom study was undertaken using a 1.5 GHz frequency-domain reflectance system [76]. Five parameters (μ_a and μ_s' of both layers and the top layer thickness) were fitted to model surfaces using χ^2 minimization, with the number of a priori parameters varied. As with time-resolved techniques, model fitting degraded for more than 3 free parameters. A hybrid steady-state/frequency-domain system [77] was designed to employ spatially-resolved, steady-state reflectance measurements for determining the optical properties of the top layer of a 2-layer phantom and frequency-domain measurements for finding μ_a and μ_s' of the bottom layer. By reducing the number of a priori parameters for the 2-layer frequency-domain model (i.e. μ_a and μ_s' are known for the top layer, given by the steady-state measurements), the optical properties can be recovered robustly.

The techniques applied to reflectance measurements can also be used in the transmittance geometry. This has been primarily used for investigating the optical properties of normal and cancerous breast tissue, since diffuse transmittance

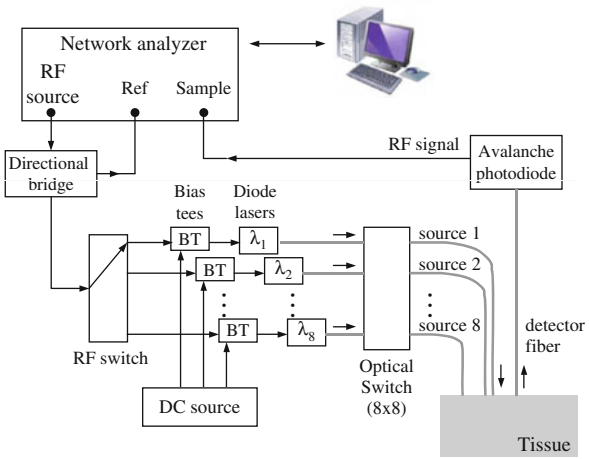


Fig. 8.16 Frequency-domain, multi-spectral reflectance instrument. A network analyzer sends a swept sinusoidal RF signal spanning from 300 kHz to 3 GHz. A directional bridge redirects a small fraction of this signal back to the network analyzer to serve as a reference signal. An RF switch directs the modulated RF signal sequentially to each of 8 diode lasers of differing wavelength. Bias tees and a DC source provide the necessary DC bias for the diode lasers. An 8x8 optical switch routes the laser light into one of 8 source fibers. A detector fiber coupled to an avalanche photodiode measures the tissue response, which is then converted to an RF signal and measured by the network analyzer. The tissue response can be compared to the reference signal to extract M and Φ. (Adapted with permission from Ref. [12])

provides functional information on chromophore and scatterer contributions deep within tissue, and the breast can be compressed to form a well-defined geometry that can be approximated as a semi-infinite slab. Time-resolved and frequency-domain systems have been used for breast studies using transmittance geometry. Thus, Durduran et al. measured the optical properties of breast tissue using frequency-domain measurements for a wide demographic of healthy women to produce baseline values of oxygen saturation (68 $\pm$ 8%) and hemoglobin concentration (34 $\pm$ 9 μM) [7]. Pifferi et al. performed time-resolved transmittance measurements in combination with reflectance measurements in breast for six volunteers aged 23–50 [69]. An interesting result was that the μ_a estimation was relatively insensitive to the measurement geometry for all subjects; however, the estimated μ_s' in the transmission geometry was significantly higher than the reflectance-based estimates. This can partially be explained by the way μ_a and μ_s' are extracted from the time-correlated data. The absorption coefficient is largely dependent upon the falling slope of the time-resolved reflectance curve, whereas μ_s' is primarily sensitive to the rising part, resulting in μ_a having a larger sampling volume than μ_s'. The photon migration paths contributing to the falling slope therefore traverse similar tissue depths in both reflectance and transmittance geometries. Another factor is that the skin layer influences scattering measurements more in reflectance mode, while the deeper core structures within the breast affect the scatter measurements more in transmission mode.

8.2.2.5 Interstitial Fluence Rate and Radiance

Spatially-resolved fluence-rate measurements and angularly-resolved radiance measurements have been investigated for determining the tissue optical properties. Fluence rate [W/cm^2] and radiance [W/cm^2 · *sr*] measurements are interstitial by necessity, and are, therefore, invasive measurement techniques, albeit minimally invasive due to the small size of fiberoptic probes.

One of the benefits of interstitial, spatially-resolved fluence rate measurements is the closed form diffusion theory equation for the fluence rate, ϕ, using the geometry shown in Fig. 8.17. The fluence rate detected by an isotropic detector at a distance r from a point source emitting power P_0 is given by

$$\phi(r) = \frac{P_0 \cdot \mu_{\text{eff}}^2}{4\pi r \mu_a} e^{-\mu_{\text{eff}} \cdot r} \tag{8.23}$$

Conceptually, measurement at two r values is then sufficient to resolve both μ_a and μ_s' using a nonlinear curve-fitting algorithm.

Accurate calibration and placement of the fluence rate probes are vital to resolve μ_a and μ_s'. Given absolute fluence rate measurements, the measurement of the source power term, P_0, needs to be calibrated prior to in vivo measurement. As well, measurement of the distance, r, needs to be accurate, especially close to the source, since a small r shift significantly affects the fluence rate measurement.

Chen et al. used fluence rates from two source-detector distances to measure optical properties during photodynamic therapy (PDT) of canine prostate [38].

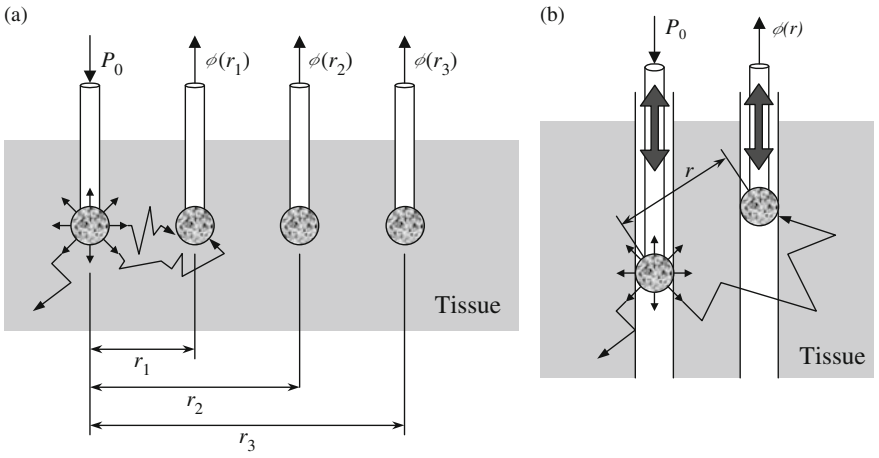


Fig. 8.17 (**a**) Geometry for spatially-resolved fluence rate measurements based on multiple interstitial isotropic probes attached to fiberoptic tips located at varying distances from a point source. (**b**) Set-up that allows for continuous measurement of fluence rate versus source-collector distance, *r*, by computer-controlled motors that adjust the placement of the isotropic probes

Brachytherapy needles (18 Ga.) introduced the source and detector fibers, tipped with isotropic scattering spheres, into the prostate. The needles were retracted during measurement. The authors found a significant increase in μ_a and decrease in μ_s' immediately after PDT using various Photofrin doses. This demonstrated that using the pre-treatment optical properties for fluence rate distribution calculations may be inadequate for calculating fluence rate distributions in the prostate during the entire procedure, and that online monitoring may significantly improve the accuracy of the treatment delivery.

Zhu et al. used interstitial optical detectors for monitoring of optical properties during Motexafin lutetium-mediated photodynamic therapy of the prostate in humans [8, 92]. The spatially-resolved fluence rate was acquired by moving the optical fibers along catheter tracks using computer-controlled stepper motors, as shown in Fig. 8.17b, allowing for near continuous measurements. The optical fiber tips comprised multiply-scattering isotropic spheres. A non-linear fitting algorithm was used to solve for the two free parameters μ_a and μ_s', with redundant fluence rate data provided by measuring over a large range of r. The authors also used a modified version of the effective attenuation coefficient ($\mu_{\text{eff}} = \sqrt{3\mu_a\mu_s'}$, when $\mu_t' \approx \mu_s'$), to allow for a wider range in the optical properties. This study recorded large inter- and intra-patient variability in the μ_a and μ_s' values, further underlining the need for online monitoring of optical properties during PDT. Angell-Petersen et al. [37] used a similar method using fluence rate sensors in parallel tracks to measure the optical properties from the fluence rate in the rat brain. Here, a ruby sphere was utilized as the sensor. This fluoresces under the 632 nm excitation light and, since fluorescence is inherently an isotropic process, the signal is proportional to the fluence rate. At this wavelength, the authors found that tumour tissue has higher absorption and lower scattering than the normal grey matter of the brain.

Weersink et al. used a modified approach of the technique in Fig. 8.17a for PDT of the prostate; rather than a point source, a linear diffusing fiber (20–25 mm in length) was used as the source for optical properties measurement as well as for the treatment light delivery in order to treat a larger tissue volume [41]. Several detector fibers were placed at varying distances from the linear source of length L, which was modeled as a series of N elemental point sources, with the fluence rate at each detector being the sum of the fluence rates from each element:

$$\phi_j = \frac{P_0 3\mu_s'}{4\pi}\left(\sum_{i=1}^{N}\frac{\exp\left(-\mu_{\text{eff}}\rho_{ij}\right)}{\rho_{ij}}\right) \tag{8.23}$$

where ρ_{ij} is the distance from the ith source element to the jth detector, and the effective attenuation coefficient was taken as $\mu_{\text{eff}} = \sqrt{3\mu_a\mu_s'}$. This model was found to closely match Monte Carlo simulations of a linear source (<1% difference) when $L/N = 0.2$ mm. For 22 patients undergoing PDT of the prostate, the average optical property values were $\mu_a = 0.39 \pm 0.18$ and $\mu_s' = 3.37 \pm 1.59$ cm^{-1}.

Dickey et al. provided some of the early demonstrations that relative radiance measurements (i.e. relative measurements at different measurement angles, θ) can

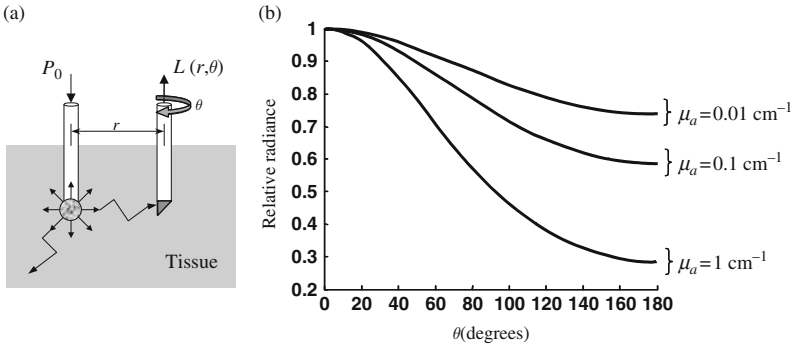


Fig. 8.18 (**a**) Interstitial radiance geometry, from Chin et al. [9]. A point source is located at a distance, where a probe (a micro-prism mounted on an optical fiber) measures the angularly-resolved radiance, $L(r,\theta)$. (**b**) Relative radiance (normalized to the 0° position) plotted against sensor rotation angle, θ, for three absorption coefficients with $\mu_s' = 10\text{cm}^{-1}$, as predicted by the *P3* approximation to the radiative transport equation, with $r = 1$ cm (Courtesy L Chin and B Lloyd, adapted with permission from Ref. [79])

be used to extract the tissue optical properties at a single source-detector distance [78]. The authors performed the measurements in phantoms and ex vivo prostate. The *P3* approximation to the Boltzmann transfer equation was used for the forward model.

Chin et al. applied this concept to develop radiance probes to monitor the development and position of the tissue coagulation front during laser interstitial thermal therapy [9]. *P3* approximations to the radiative transfer equation, diffusion theory (i.e. *P1*) and Monte Carlo techniques were analyzed for their effectiveness in extracting the tissue optical properties, for the angularly-resolved radiance measurement geometry shown in Fig. 8.18 (see Chapter 6 for further detail on the *P1* and *P3* approximations, and Chapter 5 for details on the Monte Carlo technique). One of benefits of radiance [$\text{W}\cdot\text{cm}^{-2}\cdot\text{sr}^{-1}$] measurements is that relative measurements (i.e. relative to the $\theta = 0$ position) contain enough information to recover the optical properties, abrogating the need for complex calibration phantoms and algorithms. The radiance technique is in the early development stages and is currently undergoing characterization and refinement.

8.2.2.6 Diffuse Optical Tomography

With the exception of multi-layered models, the above methods of determining optical properties assume spatially homogeneous conditions. Diffuse optical tomography (DOT) is a rapidly developing technique that can provide estimates of the spatial distribution of μ_a and μ_s' in 2 or 3 dimensions. Near-infrared light is normally used in optical tomography since the relatively low tissue absorption and scattering in this region allows greater depth penetration. There are

several motivations for developing optical tomography systems. As discussed earlier, important functional information is encoded by optical absorbers in tissue, allowing quantification of tissue oxygenation, water content and hemoglobin content. Optical imaging systems are lower in cost than magnetic resonance imaging or X-ray computed tomography systems. As well, optical tomography is well-suited for imaging of molecular probes targeted to disease markers that can provide absorption or fluorescence contrast, an area of significant interest in cancer biology [35].

The technique employs light delivered and collected from multiple projections with known geometries. DOT typically uses a transport model of light in tissue to define the forward problem, often based on the diffusion approximation [62, 81, 82]; an inverse model is then used to extract the spatial map of tissue optical properties. The additional complexity of spatial mapping requires multiple source-collector geometries to constrain the solution; however, it is accepted that the optical tomography problem is inherently ill-posed [80]. To provide a constrained solution, adequate boundary conditions need to be defined. In addition, constraining the solution spectrally with the known spectral shapes of the major absorbers and scatterers in tissue has been shown to produce more robust reconstructions [7, 62]. Measurements have been made in steady-state [82], time-resolved [83] and frequency-domain [84] modes, applying the basic principles outlined earlier.

There are many variations on the measurement geometry, although most systems use the basic configuration of a ring of fiberoptic bundles surrounding the tissue and directed radially inward, with some method of optically coupling the tissue to the fiber ends. Yuan et al. used DOT to image osteoarthritis in finger joints [82]. Source and collector fibers were equally spaced apart around the circumference of a cylindrical chamber at varying depths, with an agar-based gel filling the chamber to provide good optical coupling. McBride et al. used a frequency-domain system for tomographic breast imaging, with a ring of fiberoptic bundles directly contacting the breast and using motorized actuators to move the fibers radially into place [84].

The strength of diffuse optical tomography is that it provides spatial mapping of quantitative, functional information on the chromophores and scatterers in tissue. Experiments in breast tissue-simulating phantoms have yielded mean errors of 7.7% for oxygen saturation and 6.2% for total haemoglobin [62]. There is a significant body of evidence that diseased tissue exhibits distinct functional information (e.g. scattering power, percent oxygenation) from that of healthy tissue [4, 33, 51], providing strong impetus to develop DOT systems for disease detection, particularly in the management of breast cancer. Although DOT can provide valuable quantification of tissue function, the technique has inherently poor spatial resolution and is more prone to instability and artifacts than anatomical imaging systems (i.e. planar X-ray, CT, MRI).

The discussion on diffuse optical tomography is necessarily brief here due to the broadness of the subject, particularly concerning the topic of image reconstruction (See Chapter 19). Further discussion can be found in the reviews by Arridge on reconstruction algorithms [80], Ntziachristos and Chance on breast cancer imaging [35] and Hielscher on small animal imaging [85].

8.3 Photothermal Techniques

Photothermal techniques are based on the measurement of secondary phenomena related to optical-thermal effects following pulsed or modulated light irradiation of the tissue surface. Pulsed photothermal radiometry (PPTR) is based on the time-dependent thermal radiative emission from the tissue surface following an incident pulse of monochromatic light. In the case of photoacoustic spectroscopy (PAS), the energy absorption following a light pulse results in thermal expansion of the tissue, which can be detected at the tissue surface as an acoustic transient (see Chapters 11 and 19).

In both cases, a combined optical-thermal tissue model is necessary to derive the forward model and solve the inverse problem to calculate optical properties. In general, photothermal methods have a greater dynamic range of tissue absorption measurement compared with photometric methods, due to the fact that absorption increases the measured signal in the former, whereas in the latter it subtracts from the signal.

8.3.1 Pulsed Photothermal Radiometry

Pulsed photothermal radiometry involves the thermal response of tissue due to an impulse of monochromatic light, usually delivered using a pulsed laser. An infrared detector measures the rise and fall of the blackbody radiative emission after the light pulse. The time-dependent response measured by the detector is encoded with the tissue μ_a and μ_s' values at the excitation wavelength, along with other optothermal properties.

The general approach to modeling the PPTR effect involves solving the 1-D heat equation for depth into tissue, z, and time, t, given an initial depth-dependent temperature distribution, $T(z, t = 0)$. This temperature distribution is given by the optical fluence distribution, $H(z)$, due to the incident light pulse. Then, using the Stefan-Boltzmann law relating blackbody irradiance to the temperature distribution, an expression for the detected signal can be derived.

PPTR was initially developed for absorbing, non-scattering media. The following theoretical framework is based on the work of Prahl et al. [96]. As with many photometric and photothermal methods, the underlying assumption in PPTR modeling is for homogenous, semi-infinite tissue conditions. The depth-dependent optical fluence distribution for optically non-turbid media irradiated with a broad homogenous beam is given by Beer's law:

$$H(z) = H_0 \exp(-\mu_a z) \tag{8.25}$$

where H_0 is the optical energy that enters the tissue, i.e. the portion of incident light that is not specularly reflected from the tissue surface. The thermal relaxation time is considered to be significantly longer than the light pulse duration, so that the temperature distribution of the tissue at $t = 0$ can be modeled as

$$T(z, t=0) = \frac{\mu_a H_0 \exp(-\mu_a z)}{\rho c} \tag{8.26}$$

where ρ is the density and c is the specific heat capacity of the tissue (not to be confused with the other notation in this chapter for ρ and c).

In order to determine an expression for the blackbody emission from the tissue surface, we require the temperature distribution with respect to depth and time, given by the solution to the 1-D heat equation,

$$\frac{\partial^2 T}{\partial z^2} = \frac{1}{\alpha}\frac{\partial T}{\partial t} \tag{8.27}$$

where α is the thermal diffusivity defined by $\alpha = k/(\rho c)$, with k being the thermal conductivity. Equation (8.26) is the initial condition for this partial differential equation. Assuming that there is no significant energy loss at the tissue surface, the boundary condition is given by $\partial T \big/ \partial z = 0$. This represents a Neumann boundary condition on Eq. (8.27) which has the general solution

$$T(z, t) = \sqrt{\frac{1}{4\alpha\pi t}} \int_0^\infty \left\{ \exp\left[-\frac{(z - z')^2}{4\alpha t} \right] + \exp\left[-\frac{(z + z')^2}{4\alpha t} \right] \right\} T(z', 0)\, dz' \tag{8.28}$$

The Stefan-Boltzmann law has the blackbody emission, E_b, related to the temperature as $E_b = \sigma T^4$ [W/m^2], with the Stefan-Boltzmann constant being $\sigma = 5.6704 \times 10^{-8}\text{W} \cdot \text{m}^{-2} \cdot \text{K}^{-4}$. The infrared emission is obtained from integrating the 4$^{\text{th}}$ power of the resultant temperature distribution over the entire tissue depth, weighted by an escape function. The net emission relative to the blackbody radiation at the native tissue temperature, T_0, is given by

$$S(t) = A\varepsilon\eta\sigma\mu_{\text{IR}} \int_0^\infty \left[T(z, t)^4 - T_0^4 \right] \exp(-\mu_{\text{IR}} z)\, dz \tag{8.29}$$

where A is the effective detection area for the infrared (IR) signal, η is an efficiency constant that takes into account the detector sensitivity and μ_{IR} is the average IR absorption coefficient of the medium in the detector sensitivity bandwidth. The tissue emissivity, ε, is the fraction of energy radiated by the tissue compared with that of blackbody at the same temperature. Typically, $T_0 \gg [T(z, t) - T_0]$, i.e. the native tissue temperature, is far greater than the light-induced temperature rise. Binomial expansion of the 4th power terms in Eq. (8.29) gives

$$T(z,t)^4 - T_0^4 = [T(z,t) - T_0]\left[T_0^3 + T(z,t)\, T_0^2 + T(z,t)^2\, T_0 + T(z,t)^3 \right] \tag{8.30}$$

The terms in the second square brackets with orders of T_0 less than 3 can be dropped since T_0^3 is dominant. This gives the radiometric signal as

$$S(t) = 4A\varepsilon\eta\sigma\mu_{\mathrm{IR}}T_0^3\int_0^\infty [T(z,t) - T_0]\exp(-\mu_{\mathrm{IR}}z)\,dz \tag{8.31}$$

The infrared absorption coefficient, μ_{IR}, is generally high for tissue (a published μ_{IR} value for skin in the 3–5 μm spectral range is 220 cm^{-1}, where the absorption is primarily dominated by water) [97] so the negative exponential term is very small everywhere except extremely close to the surface. Equation (8.31) can be further simplified by removing the temperature term from the integrand, since the surface temperature dominates due to high μ_{IR}. Evaluating the remaining exponential integral results in the radiometric signal expressed as

$$S(t) = 4A\varepsilon\eta\sigma T_0^3[T(0,t) - T_0] \tag{8.32}$$

Solving for the integral at $z = 0$ in Eq. (8.28) given the initial temperature distribution in Eq. (8.26), the solution to the PPTR signal for a homogenous absorber is, therefore [96],

$$S(t) = \frac{4A\varepsilon\eta\sigma T_0^3 H_0}{\rho c}\frac{\mu_{\mathrm{IR}}\mu_a}{{\mu_{\mathrm{IR}}}^2 - \mu_a^2}\left[\mu_{\mathrm{IR}} f\left(\mu_a^2\alpha t\right) - \mu_a f\left({\mu_{\mathrm{IR}}}^2\alpha t\right)\right] \tag{8.33}$$

where $f(x) = \mathrm{erfc}\left(\sqrt{x}\right)\exp(x)$, and erfc is the complementary error function. In the case of an optically-homogeneous tissue with both absorption and scattering, a closed-form expression for $S(t)$ in terms of μ_a and μ_s' has been derived assuming diffusion theory for $H(z)$ [96],

$$H(z) = A\exp(-\mu_{\mathrm{eff}}z) + B\exp\left(-\mu_t' z\right) \tag{8.34}$$

where

$$A = \frac{H_0(9 + 6\kappa)\mu_s' D}{\left(1 + \kappa\sqrt{4\mu_a D}\right)(1 - 9\mu_a D)} \text{ and } B = \frac{-2H_0}{1 - 9\mu_a D} \tag{8.35}$$

with $\mu_t' = \mu_a + \mu_s'$, D is the diffusion constant and κ is the internal reflection parameter (refer to section 8.2.2.1). With the assumption that $\mu_{\mathrm{IR}} \gg \mu_{\mathrm{eff}}$ and $\mu_{\mathrm{IR}} \gg \mu_t'$, the resulting PPTR signal is

$$S(t) = \frac{4A\varepsilon\eta\sigma T_0^3}{\rho c}\left[Af\left(\mu_{\mathrm{eff}}^2\alpha t\right) - Bf\left(\left(\mu_t'\right)^2\alpha t\right)\right] \tag{8.36}$$

Note that in the absence of scattering, Eq. (8.36) reduces to Eq. (8.33). The form of the PPTR signal suggests that the two optical coefficients could be determined from the shape of the PPTR time curve, without requiring the determination of the absolute $S(t = 0)$ value. In practice, the shape is not a very sensitive function of the separate μ_a and μ_s' values, but this limitation can be overcome by measuring the absolute PPTR signal. For this, the tissue emissivity, detector area, detector

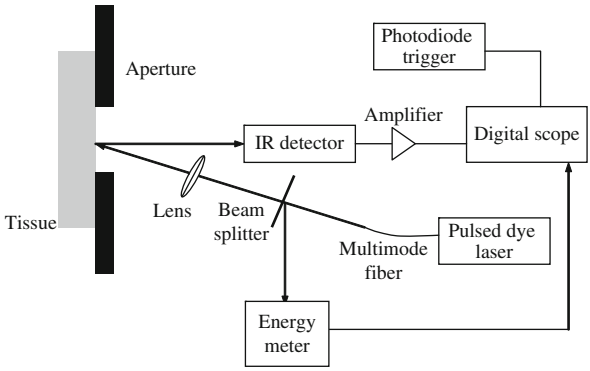


Fig. 8.19 Illustration of a pulsed photothermal radiometry system. (Adapted with permission from Ref. [96])

efficiency, thermal diffusivity, tissue density, specific heat and ambient temperature must be known, and the refractive index mismatch at the surface is important to determine the fraction of light that enters the tissue (i.e. that is not specularly reflected) [98].

Figure 8.19 shows the elements of a typical PPTR system [96]. A dye laser provides ~1 μs pulses at 627 nm which are incident on the sample over a beam diameter of a few mm. A liquid-nitrogen cooled HgCdTe detector is placed at the focal point of a germanium lens to detect the 3–12 μm IR emission from 1 mm^2 of the surface. The output from the detector is amplified and recorded as a time trace on a digital oscilloscope.

The analogous frequency-domain technique has also been explored [99]. An example of this approach used an argon laser that was frequency-modulated by an acousto-optic tunable filter, while the blackbody radiation was collected by a cooled HgCdTe detector. The amplitude and phase response of the tissue has been shown to improve the uniqueness in determining μ_a and μ_s' compared with pulsed time domain systems. Photothermal measurements have also been taken using a fiber optics-based system [100]. An optical fiber transmits and receives optical signals where the fiber contacts the tissue surface. A series of pulses from a diode laser are delivered via the fiber to heat the tissue. A thin polyethylene terephthalate (PET) film is sandwiched between the tissue and the fiber optic tip. Due to thermal effects, the PET film changes optical thickness. Another continuous wave light source, from a wavelength-tunable laser diode, is conducted through the fiber to probe the PET film. The change in optical thickness of the film produces a phase shift in the reflected laser light, effectively acting as an interferometer-based temperature sensor.

PPTR may be used in combination with the total diffuse reflectance to provide a means for estimating both μ_a and μ_s' in highly turbid media. Chen et al. [101] used the configuration shown in Fig. 8.20 to split a 585 nm laser pulse between PPTR detection optoelectronics and an integrating sphere to measure the total diffuse reflectance (refer to Section 8.2.2.1 for details and governing equations). Scattering

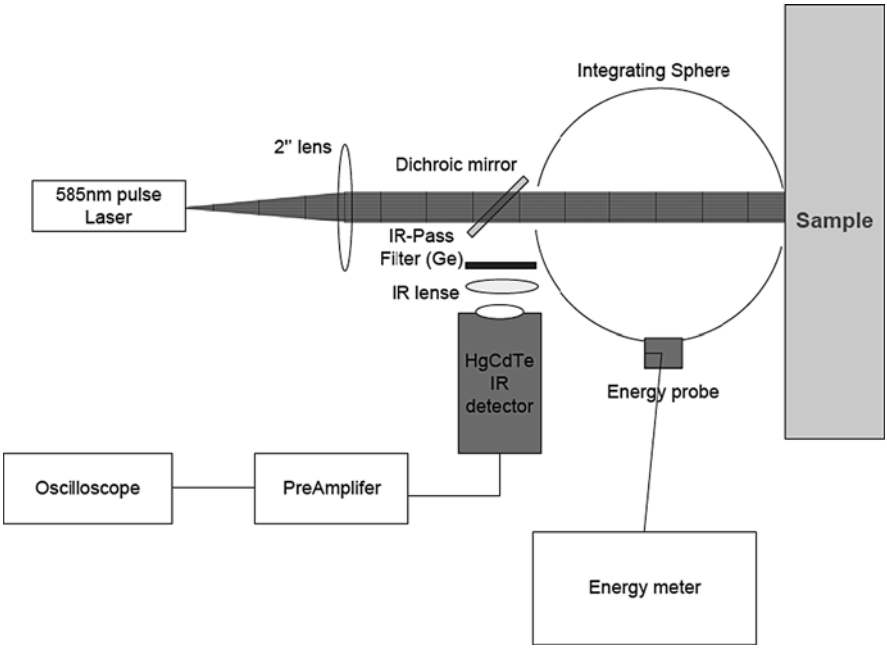


Fig. 8.20 Experimental setup for measuring PPTR together with the total diffuse reflectance [101] (Courtesy B Chen and AJ Welch, adapted with permission from Ref. [101])

media exhibits a stronger fluence rate just beneath the surface, resulting in a temperature jump, T_s, that is higher than T_c, the temperature jump in non-scattering media of the same absorption. The temperature jump ratio was noted to have the linear form $T_s/T_c = 1 + 2\kappa R_d$, where κ is defined in Eq. (8.7), and R_d is the total diffuse reflectance. The temperature jump in non-scattering media was empirically shown to be linearly proportional to the absorption coefficient (range of $\mu_a = 0.2–60\ \text{cm}^{-1}$): together with knowledge of the temperature jump ratio, T_s/T_c, this allows μ_a to be calculated in turbid media. With the total diffuse reflectance measurement, μ_s' can also be recovered using Eq. (8.7).

PPTR has also been used to study layered media. For a subsurface layer that is strongly absorbing compared to the topmost layer, the PPTR signal shows a delayed maximum (at $t > 0$) as heat flows from the lower layer to the surface. This time delay is sensitive to the depth of the absorbing layer. Examples of potential applications are to characterize arterial tissues and to measure the depth of the absorbing blood layer in the dermis. The latter has been performed non-invasively in human skin [13, 102] and has been applied to the study of port wine stain lesions [14, 103].

8.3.2 Photoacoustic Spectroscopy

The photoacoustic effect has been used to determine optical properties in an analogous manner to pulsed photothermal radiometry [104–108]. As in PPTR, an

incident pulse of light into tissue results in localized heating that is spatially dependent on the light fluence distribution. The rapid thermal expansion causes a pressure wave to propagate through the tissue. This is detected at the tissue surface by an acoustic transducer, with the time-dependent signal related to the tissue depth by the speed of sound.

The pressure increase, ΔP, at depth, z, is given by

$$\Delta P(z) = \Gamma \mu_a H(z) \tag{8.37}$$

The Grüneisen parameter, Γ, is the conversion efficiency from thermal to mechanical energy. Γ is non-dimensional and temperature-dependent, and is expressed as

$$\Gamma = \frac{\beta c_s^2}{C_p} \tag{8.38}$$

where β is the coefficient of thermal expansion, c_s is the speed of sound in the medium and C_p is the heat capacity at constant pressure.

The basic assumption here is that the heat deposition by the light source is fast compared to the thermal expansion of the tissue, analogous to the assumption in PPTR modeling that the laser heating time interval is much shorter than the tissue thermal relaxation. The depth-dependent fluence distribution, $H(z)$, may be determined using Beer's Law in the case of an optically clear medium or using a diffusion theory model in a turbid medium, such as the expressions used in the above discussions on PPTR [Eqs. (8.25) and (8.34), respectively].

The temporal profile received by an acoustic transducer measuring the pressure waves at the surface can be correlated to depth as $t = z/C_s$, with t being the time after the light pulse. The pressure is maximum at the surface, i.e. the signal is highest at the time corresponding to $z = 0$. This maximum amplitude is directly proportional to the absorption coefficient:

$$\Delta P(z = 0) = \Delta P_{\max} = \Gamma \mu_a H_0 \tag{8.39}$$

As in PPTR, H_0 is the optical energy that is delivered into the tissue, i.e. the energy that is not specularly reflected. The reduced scattering coefficient may be extracted from the profile of the time-dependent acoustic profile. As mentioned above, either a diffusion theory or Monte Carlo tissue model may be used to determine the form of the fluence distribution, $H(z)$, in order to extract μ_s' based on the shape of the acoustic transient signal, given that μ_a is known from the maximum amplitude of the acoustic signal.

An example of a photoacoustic measurement setup is shown in Fig. 8.21. Since the transducer measures the acoustic transient in transmission geometry, an initial compression (positive) wave is detected first, followed by a tensile (negative) wave (the acoustic transient can also be measured in reflectance geometry). The peak signal is given by Eq. (8.39). The shape of the decaying transient signal may be used

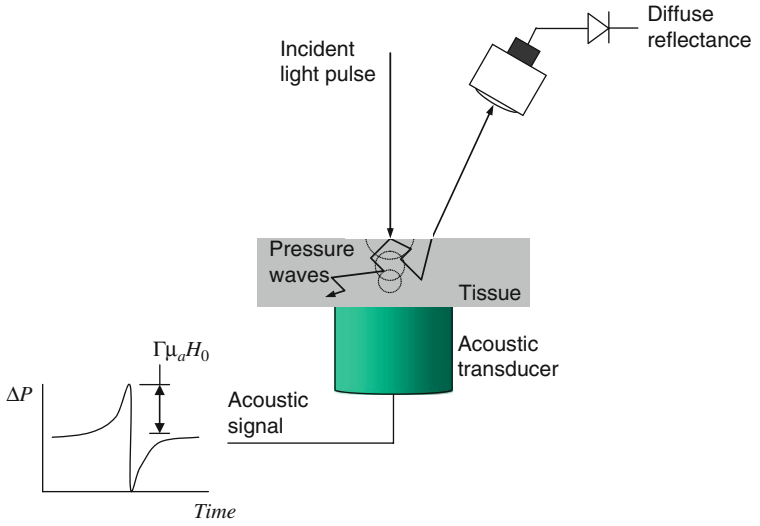


Fig. 8.21 Transmission geometry for measurement of the photoacoustic effect to determine optical properties, as in Oraevsky et al. [109]. The diffuse reflectance may also be measured using a "distant detector" to contribute to the estimation of μ_s' [110]

to estimate μ_s'. An alternative is to measure the total diffuse reflectance using the "distant detector" approach shown in Fig. 8.21. The total diffuse reflectance can be incorporated into the calculation of μ_s' using Eq. (8.7) (refer to Section 8.2.2.1) [110]. Spectrally-resolved photoacoustic measurements may be made using multiple lasers or a white light source coupled to a monochromator. An Nd:YAG laser with 2nd- and 3rd-harmonic converters has also been used to provide 355, 532 and 1064 nm [109].

Substantial progress has been made recently in photoacoustic imaging systems [111–115], described in detail in Chapter 19. The acoustic detector can be scanned linearly above a sample in reflection mode such that the acquired $\Delta P(t)$ signals may be reconstructed into a depth-resolved image slice, analogous to an ultrasound B-scan. A 3-D map may be constructed by point-by-point scanning over the tissue surface, a technique commonly referred to as photoacoustic microscopy [112]. Photoacoustic scans may also be taken circumferentially around a sample, and the image tomographically reconstructed using the photoacoustic equations and solving an inverse problem [113].

Utilizing two or more wavelengths in the VIS-NIR range allows for quantitative mapping of blood oxygenation and hemoglobin changes, taking advantage of the functional information contained in the absorption contrast and the spatial resolution of ultrasound. Using photoacoustic tomography to image the vasculature in rat cerebral cortex, Wang et al. [113, 114] demonstrated marked oxygen saturation differences in the vasculature between hyperoxic, normoxic and hypoxic conditions, imaging with dual wavelengths in the 580–600 nm range. Zhang et al. [115] used photoacoustic microscopy to image the subcutaneous vasculature in rats in vivo. The

system employed multiple wavelengths to classify veins and arteries based on StO_2 maps of the vasculature network. The photoacoustic imaging technique has also been used in rat models to visualize brain lesions and to detect functional changes in the brain from whisker stimulation [113].

While spatial mapping of the absolute values of μ_a and μ_s' are, in principle, possible for certain geometries using photoacoustic imaging techniques, generally the work to date involves mapping relative metrics such as oxygen saturation and blood volume changes. Photoacoustic imaging has shown to be particularly adept at quantitatively resolving oxygen saturation (a "relative" metric since it is based on the ratio of signals at different wavelengths), with ex vivo experiments using bovine blood yielding agreement with spectrophotometry within ~4%. [115]

8.4 Summary and Conclusions

The past decade-and-a-half since the 1st edition of this book [116] has seen significant development in the techniques to measure the optical properties of tissue. This growth has been enabled by improvements in optical technologies (light sources, fiber optics, detectors) and has been motivated by the proliferation of optically-based diagnostic and therapeutic procedures that need information on tissue optical properties. There have been major advances in in vivo photometric techniques in particular; ex vivo photometric techniques are more established technologies, while photothermal, and especially photoacoustic, methods are evolving rapidly.

All the methods described above have potential advantages, depending on the application. Certain techniques are well-suited for specific needs: e.g. spatially-resolved fluence rate measurements can be easily adapted to brachytherapy templates for photodynamic therapy of the prostate; spectrally-constrained diffuse reflectance measurements can be made with small-size fiber optic probes and used in real-time, allowing for measurements in tissue sites whose optical properties are expected to change dynamically in vivo; and time-resolved and frequency-domain measurements have shown potential for investigating the optical properties of layered tissue, such as skin and the mucosal lining of hollow organs. Pulsed photothermal radiometry and photoacoustic methods have a far greater dynamic range in measuring absorption than photometric methods, making these techniques valuable for measurements in the UV-VIS range of high hemoglobin absorption and the IR range of significant water absorption.

On the other hand, each method also has penalties in terms of cost, convenience and performance. A general statement is that the cost, complexity and acquisition time increase from steady-state to frequency-domain to time-resolved measurements; however, the robustness and accuracy also increases to compensate [108]. Spectrally-constrained diffuse reflectance may be a cost-effective and real-time method, but can be disturbed by unexpected deviations from the a priori spectral information. Diffuse optical tomography can provide spatial and spectral information of the tissue optical properties, although the ill-posed nature of the system means that the inverse problem is inherently very complex.

The discussions in this chapter on the strengths and limitations of each method are intended to aid the reader in critically evaluating the best technique(s) for measuring tissue optical properties for their particular biomedical applications.

Acknowledgments The authors would like to thank Alex Vitkin for consultation on photothermal physics. Tomas Svensson, Erik Alerstam, and Stefan Andersson-Engels generously provided data and contributions to the time-resolved diffuse reflectance section. Mathieu Roy performed the integrating sphere measurements on ex vivo tissue. Lee Chin and Brendan Lloyd produced the data for the radiance measurements section. Anthony Kim is supported by NIH Grant R01 NS052274 (USA) and by the Natural Sciences and Engineering Research Council of Canada.

References

1. Gebhart SC, Lin WC, and Mahadevan-Jansen A. *In vitro* determination of normal and neoplastic human brain tissue optical properties using inverse adding-doubling. *Phys. Med. Biol.*, 51:2011–2027 (2006).
2. Prahl SA, van Gemert MJC, and Welch AJ. Determining the optical properties of turbid media by using the adding-doubling method. *Appl. Opt.*, 32:559–568 (1993).
3. Farrell TJ, Patterson MS, and Wilson BC. A diffusion theory model of spatially resolved, steady-state diffuse reluctance for the non-invasive determination of tissue optical properties *in vivo*. *Med. Phys*., 19:879–888 (1992).
4. Peters VG, Wyman DR, Patterson MS, and Frank GL. Optical properties of normal and diseased human breast tissues in the visible and near infrared. *Phys. Med. Biol.*, 35:1317–1334 (1990).
5. Salomatina E, Jiang B, Novak J, and Yaroslavsky AN. Optical properties of normal and cancerous human skin in the visible and near-infrared spectral range. *J. Biomed. Opt.*, 11:064026 (2006).
6. Yaroslavsky IV, Yaroslavsky AN, Goldbach T, and Schwarzmaier HJ. Inverse hybrid technique for determining the optical properties of turbid media from integrating-sphere measurements. *Appl. Opt.*, 35:6797–6809 (1996).
7. Durduran T, Choe R, Culver JP, Zubkov L, Holboke MJ, Giammarco J, Chance B, and Yodh AG. Bulk optical properties of healthy female breast tissue. *Phys. Med. Biol.*, 47:2847–2861 (2002).
8. Zhu TC, Finlay JC, and Hahn SM. Determination of the distribution of light, optical properties, drug concentration, and tissue oxygenation *in-vivo* in human prostate during motexafin lutetium-mediated photodynamic therapy. *J. Photochem. Photobiol. B*, 79:231–241 (2005).
9. Chin LC, Whelan WM, and Vitkin IA. Information content of point radiance measurements in turbid media: Implications for interstitial optical property quantification. *Appl. Opt*., 45:2101–2114 (2006).
10. Cuccia DJ, Bevilacqua F, Durkin AJ, and Tromberg BJ. Modulated imaging: Quantitative analysis and tomography of turbid media in the spatial-frequency domain. *Opt. Lett*., 30:1354–1356 (2005).
11. Svensson T, Andersson-Engels S, Einarsdóttír M, and Svanberg K. *In vivo* optical characterization of human prostate tissue using near-infrared time-resolved spectroscopy. *J. Biomed. Opt.*, 12:014022 (2007).
12. Tromberg BJ, Shah N, Lanning R, Cerussi A, Espinoza J, Pham T, Svaasand L, and Butler J. Non-invasive *in vivo* characterization of breast tumors using photon migration spectroscopy. *Neoplasia* 2:26–40 (2000).
13. Long FH, Anderson RR, and Deutsch TF. Pulsed photothermal radiometry for depth profiling of layered media. *App. Phys. Lett*., 51:2076–2078 (1987).
14. Nelson JS, Jacques SL, and Wright WH. Determination of thermal and physical properties of port wine stain lesions using pulsed photothermal radiometry. *Proc. SPIE* 1643:287–298 (1992).

15. Rosencwaig A. *Photoacoustics and photoacoustic spectroscopy.* Wiley, New York (1980).
16. Friebel M, Roggan A, Müller G, and Meinke MJ. Determination of optical properties of human blood in the spectral range 250 to 1100 nm using Monte Carlo simulations with hematocrit-dependent effective scattering phase functions. *J. Biomed. Opt*., 11:34021 (2006).
17. Jacques SL and Prahl SA. Modeling optical and thermal distributions in tissue during laser irradiation. *Lasers Surg. Med.*, 6:494–503 (1987).
18. Karagiannis JL, Zhang Z, Grossweiner G, and Grossweiner LI. Applications of the 1-D diffusion approximation to the optics of tissues and tissue phantoms. *Appl. Opt.*, 28:2311–2317 (1989).
19. Flock ST, Wilson BC, and Patterson MS. Total attenuation coefficients and scattering phase functions of tissues and phantom materials at 633 nm. *Med. Phys* 14:835–842 (1987).
20. Marchesini R, Bertoni A, Andreola S, Melloni E, and Sichirollo AE. Extinction and absorption coefficients and scattering phase functions of human tissues *in vitro*. *Appl. Opt.*, 28:2318–2324 (1989).
21. Jacques SL, Alter CA, and Prahl SA. Angular dependence of HeNe laser light scattering by human dermis. *Lasers Life Sci.*, 1:309–333 (1987).
22. Vogel A, Diugos G, Nuffer R, and Birngruber R. Optical properties of human sclera, and their consequences for transcleral laser applications. *Lasers Surg. Med.*, 11:331–340 (1991).
23. Pickering JW, Prahl SA, Wieringen N van, Beek JF, Moes CJM, Sterenborg HJCM, and Gemert MJC van. Double integrating sphere system to measure optical properties of tissue. *Appl. Opt.*, 32:399–410 (1993).
24. Star WM, Marijnissen JPA, and van Gemert MJC. Light dosimetry in optical phantoms and in tissues: I. Multiple flux and transport theory. *Phys. Med. Biol.*, 33:437–454 (1988).
25. Yoon G, Welch AJ, Motamedi M, and Gemert MJC van. Development and application of three-dimensional light distribution model for laser irradiated tissue. *IEEE J. Quantum Electron.*, 23:1721–1732 (1987).
26. Dam JS, Dalgaard T, Fabricius PE, and Andersson-Engels S. Multiple polynomial regression method for determination of biomedical optical properties from integrating sphere measurements. *Appl. Opt*., 39:1202–1209 (2000).
27. Splinter R, Cheong W-F, van Gemert MJC, and Welch AJ. *In vitro* optical properties of human and canine brain and urinary bladder tissues at 633 nm. *Lasers Surg. Med.*, 9:37–41 (1989).
28. Profio AE. *Radiation shielding and dosimetry*. Wiley, New York (1979).
29. Profio AE and Sarnaik J. Fluorescence of HpD for tumor detection and dosimetry in photoradiation therapy. In: DR Doiron and CD Gomer (eds) *Porphyrin localization and treatment of tumors*. Liss, New York, pp. 163–175 (1984).
30. Madsen SJ, Patterson MS, and Wilson BC. The use of India ink as an optical absorber in tissue-simulating phantoms. *Phys. Med. Biol.*, 37:985–993 (1992).
31. Chan E, Menovsky T, and Welch AJ. Effects of cryogenic grinding on soft-tissue optical properties. *Appl. Opt.*, 35:4526–4532 (1996).
32. Bolin FP, Preuss LE, Taylor RC, and Ference RJ. Refractive index of some mammalian tissues using a fiber optic cladding method. *Appl. Opt*., 28:2297 (1989).
33. Bargo PR, Prahl SA, Goodell TT, Sleven RA, Koval G, Blair G, and Jacques SL. *In vivo* determination of optical properties of normal and tumour tissue with white light reflectance and an empirical light transport model during endoscopy. *J. Biomed. Opt*., 10:034018 (2005).
34. Fishkin JB, Coquoz O, Anderson ER, Brenner M, and Tromberg BJ. Frequency-domain photon migration measurements of normal and malignant tissue optical properties in a human subject. *Appl. Opt.*, 36:10–20 (1997).
35. Ntziachristos V and Chance B. Probing physiology and molecular function using optical imaging: Applications to breast cancer. *Breast Cancer Res*., 3:41–46 (2001).
36. Palmer GM, Zhu C, Breslin TM, Xu F, Gilchrist KW, and Ramanujam N. Monte Carlo-based inverse model for calculating tissue optical properties. Part II: Application to breast cancer diagnosis. *Appl. Opt.*, 45:1072–1078 (2006).

37. Angell-Petersen E, Hirschberg H, and Madsen SJ. Determination of fluence rate and temperature distributions in the rat brain; implications for photodynamic therapy. *J. Biomed. Opt.*, 12:014003 (2007).
38. Chen Q, Wilson BC, Shetty SD, Patterson MS, Cerny JC, and Hetzel FW. Changes in *in vivo* optical properties and light distributions in normal canine prostate during photodynamic therapy. *Radiat. Res.*, 147:86–91 (1997).
39. Chen Q, Wilson BC, Dereski MO, Patterson MS, Chopp M, and Hetzel FW. Effects of light beam size on fluence distribution and depth of necrosis in superficially applied photodynamic therapy of normal rat brain. *Photochem. Photobiol.*, 56:379–384 (1992).
40. Jankun J, Lilge L, Douplik A, Keck RW, Pestka M, Szkudlarek M, Stevens PJ, Lee RJ, and Selman SH. Optical characteristics of the canine prostate at 665 nm sensitized with tin etiopurpurin dichloride: Need for real-time monitoring of photodynamic therapy. *J. Urol.*, 172:739–743 (2004).
41. Weersink RA, Bogaards A, Gertner M, Davidson SR, Zhang K, Netchev G, Trachtenberg J, and Wilson BC. Techniques for delivery and monitoring of TOOKAD (WST09)-mediated photodynamic therapy of the prostate: Clinical experience and practicalities. *J. Photochem. Photobiol. B.*, 79:211–222 (2005).
42. Zhang Q, Müller MG, Wu J, and Feld MS. Turbidity-free fluorescence spectroscopy of biological tissue. *Opt. Lett.*, 25:1451–1453 (2000).
43. Egan WG and Hilgeman TW. *Optical properties of inhomogeneous materials.* Academic, New York (1979).
44. Groenhuis FAJ, Ferwerda HA, and Ten Bosch JJ. Scattering and absorption of turbid materials derived from reflection coefficients., 1: Theory. *Appl. Opt.*, 22:2456–2462 (1983).
45. Flock ST, Patterson MS, Wilson BC, and Wyman DR. Monte Carlo modeling of light propagation in highly scattering tissue—I: Model predictions and comparison with diffusion theory. *IEEE Trans. Biomed. Eng*., 36:1162–1168 (1989).
46. Jacques SL, Gutsche A. Schwartz JA, Wang L, and Tittle FK. Video reflectometry to specify optical properties of tissue *in vivo*. *Proc. SPIE* IS11:211–226 (1993).
47. Kienle A, Lilge L, Patterson MS, Hibst R, Steiner R, and Wilson BC. Spatially resolved absolute diffuse reflectance measurements for noninvasive determination of the optical scattering and absorption coefficients of biological tissue. *Appl. Opt.*, 35:2304–2314 (1996).
48. Patterson MS, Schwartz E, and Wilson BC. Quantitative reflectance spectrophotometry for the non-invasive measurement of photosensitizer concentration in tissue. *Proc. SPIE* 1065:115–122 (1989).
49. Allen V and McKenzie AL. The modified diffusion dipole model. *Phys. Med. Biol*., 36:1621–1638 (1991).
50. Patterson MS, Chance B, and Wilson BC. Time resolved reflectance and transmittance for the noninvasive measurement of tissue optical properties. *Appl. Opt.*, 28:2331–2336 (1989).
51. Lin SP, Wang L, Jacques SL, and Tittel FK. Measurement of tissue optical properties by the use of oblique-incidence optical fiber reflectometry. *Appl. Opt.*, 36:136–143 (1997).
52. Doornbos RM, Lang R, Aalders MC, Cross FW, and Sterenborg HJ. The determination of *in vivo* human tissue optical properties and absolute chromophore concentrations using spatially resolved steady-state diffuse reflectance spectroscopy. *Phys. Med. Biol.*, 44:967–981 (1999).
53. Pfefer TJ, Matchette LS, Bennett CL, Gall JA, Wilke JN, Durkin AJ, and Ediger MN. Reflectance-based determination of optical properties in highly attenuating tissue. *J. Biomed. Opt.*, 8:206–215 (2003).
54. Arnfield MR, Tulip J, and McPhee MS. Optical propagation in tissue with anisotropic scattering. *IEEE Trans. Biomed. Eng.*, 35:372–381 (1988).
55. Amelink A, Sterenborg HJ, Bard MP, and Burgers SA. *In vivo* measurement of the local optical properties of tissue by use of differential path-length spectroscopy. *Opt. Lett*., 29:1087–1089 (2004).
56. Alerstam E, Andersson-Engels S, and Svensson T. White Monte Carlo for time-resolved photon migration. *J. Biomed. Opt.*, 13:041304 (2008).

57. Mourant JR, Fuselier T, Boyer J, Johnson TM, and Bigio IJ. Predictions and measurements of scattering and absorption over broad wavelength ranges in tissue phantoms. *Appl. Opt.*, 36:949–957 (1997).
58. Reif R, A'amar O, and Bigio IJ. Analytical model of light reflectance for extraction of the optical properties in small volumes of turbid media. *Appl. Opt.*, 46:7317–7328 (2007).
59. Amelink A and Sterenborg HJ. Measurement of the local optical properties of turbid media by differential path-length spectroscopy. *Appl. Opt*., 43:3048–3054 (2004).
60. Sun J, Fu K, Wang A, Lin AWH, Utzinger U, and Drezek R. Influence of fiber optic probe geometry on the applicability of inverse models of tissue reflectance spectroscopy: Computational models and experimental measurements. *Appl. Opt.*, 45:8152–8162 (2006).
61. Landsman MLJ, Kwant G, Mook GA, and Zijlstra WG. Light- absorbing properties, stability, and spectral stabilization of indocyanine green. *J. Appl. Physiol.*, 40:575–583 (1976).
62. Srinivasan S, Pogue BW, Jiang S, Dehghani H, and Paulsen KD. Spectrally constrained chromophore and scattering near-infrared tomography provides quantitative and robust reconstruction. *Appl. Opt*., 44:1858–1869 (2005).
63. Bevilacqua F, Piguet D, Marquet P, Gross JD, Tromberg BJ, and Depeursinge C. *In vivo* local determination of tissue optical properties: Applications to human brain. *Appl. Opt*., 38:4939–4950 (1999).
64. Patterson MS, Moulton JD, Wilson BC, Berndt KW, and Lakowicz JR. Frequency-domain reflectance for the determination of the scattering and absorption properties of tissue. *Appl. Opt.*, 30:4474–4476 (1991).
65. Madsen SJ, Wilson BC, Patterson MS, Park YD, Jacques SL, and Hefetz Y. Experimental tests of a simple diffusion model for the estimation of scattering and absorption coefficients of turbid media from time resolved diffuse reflectance measurements. *Appl. Opt.*, 31:3509–3517 (1992).
66. Svensson T, Swartling J, Taroni P, Torricelli A, Lindblom P, Ingvar C, and Andersson-Engels S. Characterization of normal breast tissue heterogeneity using time-resolved near-infrared spectroscopy. *Phys. Med. Biol*., 50:2559–2571 (2005).
67. Chernomordik V, Hattery DW, Grosenick D, Wabnitz H, Rinneberg H, Moesta KT, Schlag PM, and Gandjbakhche A. Quantification of optical properties of a breast tumor using random walk theory. *J. Biomed. Opt.*, 7:80–87 (2002).
68. Martelli F, Del Bianco S, Zaccanti G, Pifferi A, Torricelli A, Bassi A, Taroni P, and Cubeddu R. Phantom validation and *in vivo* application of an inversion procedure for retrieving the optical properties of diffusive layered media from time-resolved reflectance measurements. *Opt. Lett*., 29:2037–2039 (2004).
69. Pifferi A, Swartling J, Chikoidze E, Torricelli A, Taroni P, Bassi A, Andersson-Engels S, and Cubeddu R. Spectroscopic time-resolved diffuse reflectance and transmittance measurements of the female breast at different interfiber distances. *J. Biomed. Opt.*, 9:1143–1151 (2004).
70. Torricelli A, Pifferi A, Taroni P, Giambattistelli E, and Cubeddu R. *In vivo* optical characterization of human tissues from 610 to 1010 nm by time-resolved reflectance spectroscopy. *Phys. Med. Biol.*, 46:2227–2237 (2001).
71. Martelli F, Sassaroli A, Del Bianco S, Yamada Y, and Zaccanti G. Solution of the time-dependent diffusion equation for layered diffusive media by the eigenfunction method. *Phys. Rev. E*, 67:056623 (2003).
72. Kienle A and Glanzmann T. *In vivo* determination of the optical properties of muscle with time-resolved reflectance using a layered model. *Phys. Med. Biol*., 44:2689–2702 (1999).
73. Gurfinkel M, Pan T, and Sevick-Muraca EM. Determination of optical properties in semi-infinite turbid media using imaging measurements of frequency-domain photon migration obtained with an intensified charge-coupled device. *J. Biomed. Opt.*, 9:1336–1346 (2004).
74. Bevilacqua F, Berger AJ, Cerussi AE, Jakubowski D, and Tromberg BJ. Broadband absorption spectroscopy in turbid media by combined frequency-domain and steady-state methods. *Appl. Opt.*, 39:6498–6507 (2000).

75. Wilson BC, Patterson MS, and Pogue BW. Instrumentation for in-vivo tissue spectroscopy and imaging. *Proc. SPIE* 1892:132–147 (1993).
76. Pham TH, Spott T, Svaasand LO, and Tromberg BJ. Quantifying the properties of two-layer turbid media with frequency-domain diffuse reflectance. *Appl. Opt*., 39:4733–4745 (2000).
77. Fawzi YS, Youssef AB, el-Batanony MH, and Kadah YM. Determination of the optical properties of a two-layer tissue model by detecting photons migrating at progressively increasing depths. *Appl. Opt*., 42:6398–6411 (2003).
78. Dickey DJ, Moore RB, Rayner DC, and Tulip J. Light dosimetry using the P3 approximation. *Phys. Med. Biol.*, 46:2359–2370 (2001).
79. Chin LC, Worthington AE, Whelan WM, and Vitkin IA. Determination of the optical properties of turbid media using relative interstitial radiance measurements: Monte Carlo study, experimental validation and sensitivity analysis. *J. Biomed. Opt.*, 12:064027 (2007).
80. Arridge SR. Optical tomography in medical imaging. *Inverse Problems* 15:R41–R93 (1999).
81. Dehghani H, Pogue BW, Shudong J, Brooksby B, and Paulsen KD. Three-dimensional optical tomography: Resolution in small-object imaging. *Appl. Opt.*, 42:3117–3128 (2003).
82. Yuan Z, Zhang Q, Sobel E, and Jiang H. Three-dimensional diffuse optical tomography of osteoarthritis: Initial results in the finger joints. *J. Biomed. Opt*., 12:034001 (2007).
83. Ntziachristos V, Ma X, and Chance B. Time-correlated single photon counting imager for simultaneous magnetic resonance and near-infrared mammography. *Rev. Sci. Instr.*, 69:4221–4233 (1998).
84. McBride TO, Pogue BW, Jiang S, Osterberg UL, and Paulsen KD. A parallel-detection frequency-domain near-infrared tomography system for hemoglobin imaging of the breast *in vivo*. *Rev. Sci. Instrum*. 72:1817–1824 (2001).
85. Hielscher AH. Optical tomographic imaging of small animals. *Curr. Opin. Biotechnol*., 16:79–88 (2005).
86. van Veen RL, Amelink A, Menke-Pluymers M, van der Pol C, and Sterenborg HJ. Optical biopsy of breast tissue using differential path-length spectroscopy. *Phys. Med. Biol.*, 50:2573–2581 (2005).
87. van Staveren HG, Moes CJM, van Marle J, Prahl SA, and van Gemert MJC. Light scattering in Intralipid-10% in the wavelength range of 400–1100 nanometers. *Appl. Opt.*, 30:4507–4514 (1991).
88. Dimofte A, Finlay JC, and Zhu TC. A method for determination of the absorption and scattering properties interstitially in turbid media. *Phys. Med. Biol.*, 50:2291–311 (2005).
89. Bays R, Wagnieres G, Robert D, Braichotte D, Savary J-F, Monnier P, and van den Bergh H. Clinical determination of tissue optical properties by endoscopic spatially resolved reflectometry. *Appl. Opt.*, 35:1756–1766 (1996).
90. Spinelli L, Torricelli A, Pifferi A, Taroni P, Danesini GM, and Cubeddu R. Bulk optical properties and tissue components in the female breast from multiwavelength time-resolved optical mammography. *J. Biomed. Opt*., 9:1137–1142 (2004).
91. van Veen RLP, Sterenborg HJ, Marinelli, AW, and Menke-Pluymers M. Intraoperatively assessed optical properties of malignant and healthy breast tissue used to determine the optimum wavelength of contrast for optical mammography. J. Biomed. Opt., 9:1129–1136 (2004).
92. Zhu TC, Dimofte A, Finlay JC, Stripp D, Busch T, Miles J, Whittington R, Malkowicz SB, Tochner Z, Glatstein E, and Hahn SM. Optical properties of human prostate at 732 nm measured in mediated photodynamic therapy. *Photochem. Photobiol.*, 81:96–105 (2005).
93. Hornung R, Pham TH, Keefe KA, Berns MW, Tadir Y, and Tromberg BJ. Quantitative near-infrared spectroscopy of cervical dysplasia *in vivo*. *Hum. Reprod*., 14:2908–2916 (1999).
94. Asgari S, Röhrborn HJ, Engelhorn T, and Stolke D. Intra-operative characterization of gliomas by near-infrared spectroscopy: Possible association with prognosis. *Acta Neurochir.(Wien)*, 145:453–459 (2003).
95. Thueler P, Charvet I, Bevilacqua F, St Ghislain M, Ory G, Marquet P, Meda P, Vermeulen B, and Depeursinge C. *In vivo* endoscopic tissue diagnostics based on spectroscopic absorption, scattering, and phase function properties. *J. Biomed. Opt.*, 8:495–503 (2003).
96. Prahl SA, Vitkin IA, Bruggemann U, and Wilson BC. Determination of optical properties of turbid media using pulsed photothermal radiometry. *Phys. Med. Biol*., 37:1203–1217 (1992).

97. Majaron B, Verkruysse W, Tanenbaum BS, Milner TE, and Nelson JS. Spectral variation of the infrared absorption coefficient in pulsed photothermal profiling of biological samples. *Phys. Med. Biol.*, 47:1929–1946 (2002).
98. Anderson RR, Beck H, Bruggemann U, Farinelli W, Jacques SL, and Parrish JA. Pulsed photothermal radiometry in turbid media: Internal reflection of backscattered radiation strongly influences optical dosimetry. *Appl. Opt*., 28:2256–2262 (1989).
99. Nicolaides L, Chen Y, Mandelis A, and Vitkin IA. Theoretical, experimental, and computational aspects of optical property determination of turbid media by using frequency-domain laser infrared photothermal radiometry. *J. Opt. Soc. Am. A Opt. Image. Sci. Vis.*, 18:2548–56 (2001).
100. Laufer JG, Beard PC, Walker SP, and Mills TN. Photothermal determination of optical coefficients of tissue phantoms using an optical fibre probe. *Phys. Med. Biol.*, 46:2515–30 (2001).
101. Chen B. *Experimental and modeling study of thermal response of skin and cornea to infrared wavelengths laser irradiation*. Dissertation, University of Texas, Austin, December 2007.
102. Viator JA Choi B, Peavy GM, Kimel S, and Nelson JS. Spectra from 2.5–15 microm of tissue phantom materials, optical clearing agents and *ex vivo* human skin: Implications for depth profiling of human skin. *Phys. Med. Biol*., 48:N15–N24 (2003).
103. Huang YC, Ringold TL, Nelson JS, and Choi B. Noninvasive blood flow imaging for real-time feedback during laser therapy of port wine stain birthmarks. *Lasers Surg. Med*., 40:167–173 (2008).
104. Bernini U, Reccia R, Russo P, and Scala A. Quantitative photoacoustic spectroscopy of cateractous human lenses. *J. Photochem. Photobiol*. B4:407–417 (1990).
105. Helander P. Theoretical aspects of photoacoustic spectroscopy with light scattering samples. *J. Appl. Phys.*, 54:3410–3414 (1987).
106. Zhao Z and Myllylä R. Measuring the optical parameters of weakly absorbing, highly turbid suspensions by a new technique: Photoacoustic detection of scattered light. *Appl. Opt.*, 44:7845–7852 (2005).
107. Bernini U, Marotta M, Martino G, and Russo P. Spectrophotoacoustic method for quantitative estimation of haem protein content in wet tissue. *Phys. Med. Biol*., 36:391–396 (1991).
108. Gibson AP, Hebden JC, and Arridge SR. Recent advances in diffuse optical imaging. *Phys. Med. Biol.*, 50:R1–R43 (2005).
109. Oraevsky AA, Jacques SL, and Tittel FK. Measurement of tissue optical properties by time-resolved detection of laser-induced transient stress. *Appl. Opt.*, 36:402–415 (1997).
110. Esenaliev RO, Larin KV, Larina IV, Motamedi M, and Oraevsky AA. Optical properties of normal and coagulated tissues: Measurements using combination of optoacoustic and diffuse reflectance techniques. *Proc. SPIE* 3726:560–566 (1999).
111. Yang X and Wang LV. Photoacoustic tomography of a rat cerebral cortex with a ring-based ultrasonic virtual point detector. *J. Biomed. Opt*., 12:060507 (2007).
112. Song KH and Wang LV. Deep reflection-mode photoacoustic imaging of biological tissue. *J. Biomed. Opt.*, 12:060503 (2007).
113. Wang X, Xie X, Ku G, and Wang LV. Noninvasive imaging of hemoglobin concentration and oxygenation in the rat brain using high-resolution photoacoustic tomography. *J. Biomed. Opt*., 11:024015 (2006).
114. Wang X, Pang Y, Ku G, Xie X, Stoica G, and Wang LV. Noninvasive laser-induced photoacoustic tomography for structural and functional *in vivo* imaging of the brain. *Nat. Biotechnol.*, 21:803–806 (2003).
115. Zhang HF, Maslov K, and Sivaramakrishnan M. Imaging of hemoglobin oxygen saturation variations in single vessels *in vivo* using photoacoustic microscopy. *Appl. Phys. Lett*., 90:053901 (2007).
116. Wilson BC. Measurement of tissue optical properties: Methods and theories. In: AJ Welch and MJC van Gemert (eds) *Optical-thermal response of laser irradiated tissue.* Plenum, New York (1995).

Chapter 9
Dynamic Changes in Optical Properties

Jennifer K. Barton

9.1 Introduction

Most biological tissues are highly scattering, and, at resonant wavelengths, have strong absorption from chromophores such as oxyhemoglobin and water. Optical properties of tissue can vary considerably with time, due to actions that are either deliberate (e.g. application of "optical clearing" agents) or unintentional (e.g. dehydration of ex vivo tissue samples). During photothermal therapy, dramatic shifts in optical properties can occur due to high temperatures and reversible or irreversible biochemical reactions. An understanding of dynamic tissue optical properties is important for at least three reasons: (1) implementing tissue handling protocols that assure repeatable optical property measurements, (2) developing mathematical models that accurately predict therapeutic outcomes, and (3) manipulating tissue optical properties to enhance diagnostic information or therapeutic efficacy.

This chapter will discuss optical property changes that occur due to shifts in absorption resonances with temperature, change in index of refraction with temperature (thermal lensing effect), biochemical changes in heated oxygenated hemoglobin, denaturation of heated proteins, and change in water content. The effects of optical clearing agents, which have been attributed to a variety of mechanisms, are discussed last. It is beyond the scope of this chapter to discuss changes in optical properties due to benign and pathological changes such as edema, erythema, and neoplasia.

9.2 Shift in Absorption Resonance with Temperature

Important tissue chromophores, including water, oxyhemoglobin (HbO_2) and deoxyhemoglobin (Hb), have sharp absorption resonance peaks. By choosing laser wavelengths at or near the infrared water absorption peaks of approximately 1.94

J.K. Barton (✉)
Department of Biomedical Engineering, University of Arizona, Tucson, AZ, USA
e-mail: barton@u.arizona.edu

A.J. Welch, M.J.C. van Gemert (eds.), *Optical-Thermal Response of Laser-Irradiated Tissue*, 2nd ed., DOI 10.1007/978-90-481-8831-4_9,

and 2.95 μm, depth of penetration into soft tissue is limited and precise ablation can be performed. Depth of ablation can theoretically be controlled by tuning the treatment laser to wavelengths on- or slightly off-peak. At the HbO_2 absorption peaks, particularly the *Q* bands near 540 and 577 nm, the absorption coefficient of blood is much higher than that of surrounding tissue constituents, enabling selective photothermolysis. Indeed, absorption is so large at these wavelengths that even heating of larger blood vessels is difficult. Therefore, lasers with wavelengths on the steep red wing of the *Q* band have been developed (585–600 nm), with room temperature absorption coefficients one-half to one-tenth of those at 577 nm. However, it has been noted that the absorption spectra for water [1–3], Hb, and HbO_2 [4–8] change in both shape and magnitude as temperature increases. Therefore, both laser engineering efforts to develop devices with "optimal" parameters, and modeling to predict treatment outcome, may be misplaced if dynamic changes in optical properties are not considered.

9.2.1 Water

Alterations in water absorption spectra have been ascribed to various mechanisms. Perturbation of the chemical equilibrium between hydrogen bonded and non-hydrogen bonded water will cause spectral shifts. Dynamic changes in the optical properties of water have implications for photocoagulation and photoablation of tissues. For example, the water absorption peak at approximately 1.94 μm is frequently utilized because the penetration depth can be "tuned" by adjusting the laser wavelength to the absorption peak or slightly off-peak, and because standard fiber optics are practical in this wavelength range. However, several authors have reported that actual damage to tissue does not match predictions from simple static optical properties models. Jansen et al. found that the water absorption peak near 1.94 μm shifts to shorter wavelengths with increasing temperature [9]. They developed an optical/thermal model incorporating this dynamic change in absorption coefficient. A summary figure of these data is shown in Fig. 9.1. Schomacker et al. noted that histologically observed damage in a cornea using a tunable (1.8–2.2 μm) Co:MgF2 laser differed from expectations; for wavelengths shorter than 1.94 μm the depth of damage was less than expected, whereas for wavelengths longer than 1.94 the depth of damage was greater than expected [10]. This observation suggests that the treatment parameters of midinfrared lasers used for photorefractive surgery should be calculated with consideration of dynamic optical properties, to assure that the damage zone does not extend to Descemet's membrane. Lange et al. measured the absorption coefficients (μ_a) of water at 2.01 and 2.09 μm and found that, at 100°C, μ_a was only about two-thirds the value measured at 20°C [11]. It can be expected that during laser ablation, temperatures from 100°C to several hundred °C occur. At 2.01 and 2.09 μm, lower peak temperatures (on the order of 15%) and greater penetration depth will be achieved if the temperature rises to 100°C, compared to that expected from fixed room temperature optical properties.

Barton et al. performed a study looking at the effect of wavelength and temperature on the optical properties of albumin (egg white) through the measurement

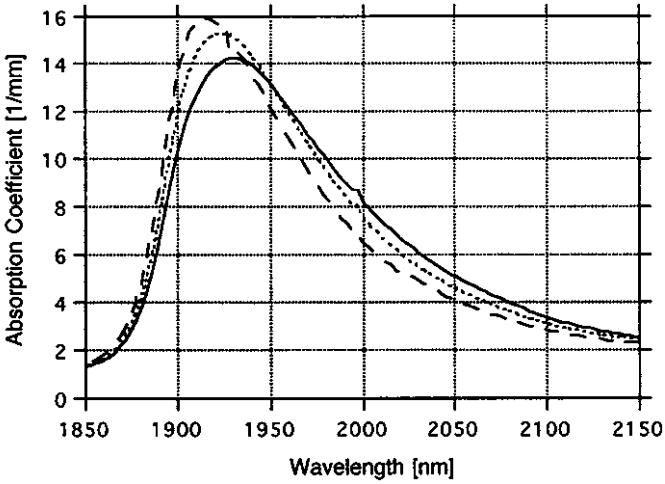


Fig. 9.1 Absorption coefficient of water as a function of temperature, showing a shift of the peak absorption to shorter wavelengths as temperature increases. *Solid line*: 22±1°C, *dotted line*: 49±5°C, *dashed line*: 70±5°C. (Reprinted with permission from [9])

of an action spectrum (Barton JK, Chan EK, Stoneman RC, Rastegar S, Welch AJ (1995) *Temperature and wavelength dependence of the threshold for coagulation in egg white*, "Unpublished"). The endpoint measured was the irradiance necessary to create a microscopically visible denaturation. Since the major chromophore in egg white at midinfrared wavelengths is water (approximately 86% by weight), it was predicted that the coagulation threshold would be a minimum at 1.94 μm at room temperature, and that this minimum would shift to shorter wavelengths with increasing egg white temperature. The output from a tunable (1.88–2.07 μm) Tm:YAG laser was focused to a sub-mm spot on the surface of egg white. The egg white was equilibrated at 21.6 ± 0.1, 28.0 ± 0.5, and 41.6 ± 0.5°C. At each temperature and wavelength setting, the power of the laser was set to a level that produced an easily visible lesion for an exposure duration of 0.5 s, then was incrementally decreased until a lesion was no longer visible. The process was reversed – incrementally increasing from low power to a level where lesions were just visible. The irradiance at which there was a 50% chance of forming a microscopically visible lesion was defined as the threshold.

The results of this study – irradiance threshold for lesion formation with respect to wavelength and temperature – are shown in Fig. 9.2. Bars indicate standard deviations from Probit curves, assuming measurements are from a normal distribution. From this graph it can be seen that, as expected, the threshold for lesion formation at 21.6°C follows an inverted room temperature water absorption curve, with a minimum irradiance needed at 1.94 μm, the peak of the water absorption curve at room temperature. At elevated temperatures, two trends are noticed. First, for a given wavelength, the irradiance necessary to form a lesion decreases. This is expected since denaturation is a rate process dependent upon temperature and

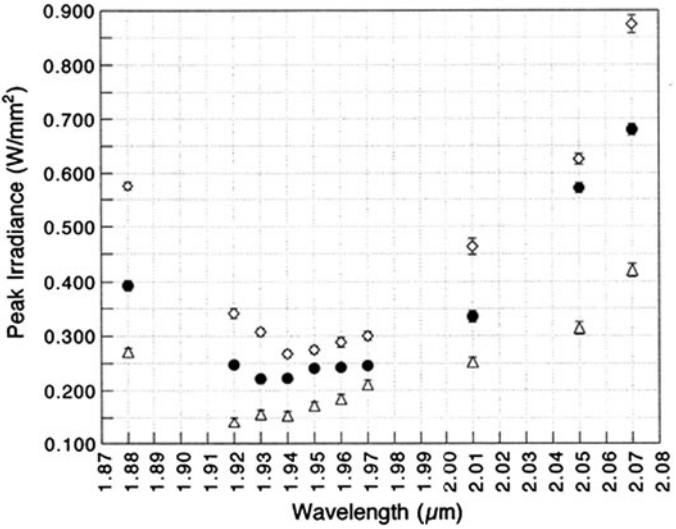


Fig. 9.2 Peak irradiance necessary to achieve a microscopic coagulum in egg white as a function of wavelength and initial temperature, showing a shift to shorter wavelengths with increasing initial temperature. *Diamonds*: 21.6±0.1°C, *filled circles*: 28.0±0.5°C, *triangles*: 41.6±0.5°C

time. That is, less energy is needed to raise the local egg white temperature to the critical temperature needed to form a coagulum with constant irradiation time. As expected, preheating the egg white had the most effect upon the irradiance threshold at wavelengths with lower absorption coefficients (high and low end of the wavelength range employed). Second, the wavelength at which the minimum irradiance is needed for lesion formation shifts to shorter wavelengths. At 41.6°C, a minimum was measured at 1.92 μm. This is in agreement with the data by Jansen et al. (Fig. 9.1) which show that the peak absorption coefficient of water shifts to 1.92 μm at 49°C [9]. The ratio of the threshold irradiance at 41.6°C to the threshold irradiance at 21.6°C is lower at shorter wavelengths (1.88 and 1.93 μm) than it is at 1.94 μm, reflecting the fact that the absorption coefficient is increasing with temperature. These results verify suppositions that absorption in egg white around 1.94 μm is dominated by water, and that the peak of the local absorption coefficient maximum in egg white will shift to shorter wavelengths with increasing temperature. The results also suggest that other soft tissues will behave in a similar fashion, and underscore the importance of considering dynamic changes in optical properties of water when calculating laser dosimetry.

9.2.2 Deoxyhemoglobin and Oxyhemoglobin

Various mechanisms have also been cited for causing reversible changes in the absorption spectra of Hb and HbO_2 with respect to temperature. In Hb, a decrease

in absorption with increasing temperature may be attributable to the change in the population of different spin states of the iron atom [7]. In HbO_2, the bathochromic (red, or longer wavelength) shift in the *Q* band absorption peaks may be due to overlap of vibrationally excited levels of the ground and first electronically excited states at elevated temperatures, leading to transitions with lower energy (longer wavelength) [12]. The bathochromic shift of HbO_2 absorption has been observed in whole blood exposed to air (and thus nearly completely oxygenated). Verkryusse et al. showed that the transmission and reflectance of blood decreased during the initial (pre-coagulation) stages of heating using 0.5 ms, 586 nm laser pulses [13]. Black et al. showed that during the pre-coagulation phase (up to about 76°C of a 10 ms, 532 nm laser pulse), the transmission of probe wavelengths through 200 μm of diluted whole blood at a hematocrit (Ht) of 24.5% had divergent behavior. At 532 nm, the transmission increased about 20%, whereas at 594 nm it decreased approximately 25%. At 633 nm, the transmission decreased slightly and at 675 nm the transmission was unchanged. This spectroscopic data is strongly supportive of a bathochromic shift. Looking at Fig. 9.3, which shows the room-temperature absorptivity of HbO_2, it is apparent that wavelengths on the blue side of the double-peaked absorption curve should experience a decrease in absorptivity with a bathochromic shift, whereas those on the steep red wing should experience an increase. Wavelengths where the slope of the absorptivity curve is relatively flat should show little change. This is exactly what is noted. This same trend was found

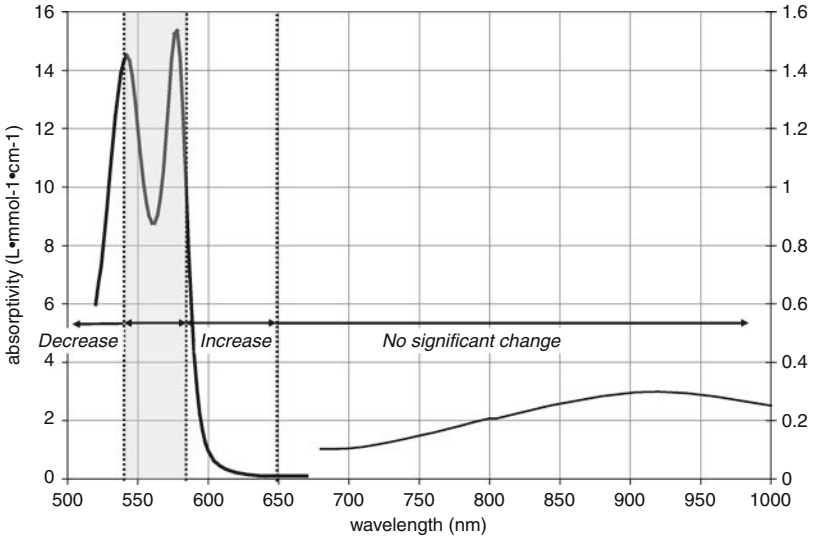


Fig. 9.3 Room temperature absorptivity of HbO_2. Data for wavelengths above 680 nm have been multiplied by 10x to show spectral detail. Marked are regions where the absorptivity is expected to decrease, increase, or have no significant change as a function of increasing temperature. In the *shaded region* (from approximately 540–580 nm) the absorptivity may increase or decrease depending on the magnitude of the temperature increase. Data from Zijlstra et al. [15]

when the pump was a 10 ms, 1064 nm laser, suggesting that the bathochromic shift is a photothermal effect [12]. Therefore, transmittance through blood during heating (in the absence of denaturation) can be described by the Beer-Lambert law with a temperature dependent absorption coefficient. Because the anisotropy of blood is so high (generally given as 0.99–0.995), scattering losses are neglected.

$$\frac{E}{E_o} = \exp\left(-\left[\mu_{a_{RT}} + \Delta\mu_a(T) \bullet \Delta T\right] z\right) \tag{9.1}$$

where $\mu_{a_{RT}}$ is the room temperature value of absorption coefficient (22°C) commonly provided in the literature, $\Delta\mu_a(T)$ is the temperature dependent change in absorption coefficient per °C during heating, and ΔT is the temperature change in °C. Since the spectrum may change shape as well as shift, an absorption spectrum at 76°C cannot be completely determined from a few discrete wavelengths. A complete spectrum at this temperature has not been measured in a standard scanning spectrophotometer because during slow heating, coagulation will occur above about 46°C. Once coagulation occurs, other optical property changes, including an increase in scattering and conversion of HbO_2 to methemoglobin (metHb), happen. From the available data in [14] and [12] it appears that in the temperature range of 22–76°C (76°C is the temperature at which denaturation occurs for the 10 ms laser pulse used in these studies), the change in absorption coefficient at a given probe wavelength is approximately linear and as shown in Table 9.1. At 675 nm and above, no significant change in μ_a was noted. Room temperature absorption coefficients are from [15], assuming fully oxygenated blood and 82 g Hb/l in the diluted blood (24.5% Ht).

The change in absorption spectra of HbO_2 has significant implications for the development of therapeutic vascular lasers. To date it has been assumed generally that the penetration depths of light into the vessel could be predicted by consideration of only the room-temperature HbO_2 absorption spectrum. The argument is frequently made that 585, 595 and 600 nm lasers penetrate more deeply into the

Table 9.1 Dynamic absorption coefficient of diluted blood (24.5% Ht) heated from 22 to 76°C, assuming fully oxygenated blood with 18 g Hb/l. $\Delta\mu_a$ is the % change in the absorption coefficient over this temperature range. μ_{aRT} is the room temperature coefficient of blood. $\Delta\mu_a(T)$ is the change in absorption coefficient per °C over this temperature range

Wavelength [nm]	Δ Transmittance 22–76°C [%]	Δ μ_a 22–76°C [%]	μ_{aRT}[cm^{-1}]	$\Delta\mu_a(T)$ [$cm^{-1}1/°C$]
532	20.0	−4.0	123.00	−0.093
594	−50.0	28.8	23.90	0.118
612	−17.9	15.5	3.74	0.0113
633	−5.3	9.6	1.39	0.0027
675	0	0	1.10	0.00

vessel than shorter wavelengths and therefore should be more effective at inducing photothermal sclerosis [16, 17]. This assumption may not be correct; only by considering the time-temperature history of the laser pulse, with dynamic optical properties, can the extent of damage be predicted.

9.3 Thermal Lensing

It is well known that the index of refraction in materials can be temperature dependent. The major constituent of soft tissues is water, and water has a negative dependence with temperature. The magnitude of this change δn/δT is about –0.0001°C^{-1} at 20°C and 589 nm.

Gordon et al. made the observation in 1965 that build-up and decay transients occurred when transparent liquid and solid cells were placed into the resonator of a He:Ne laser [18]. They attributed these effects to light absorption and heating of the substance by the Gaussian laser beam, leading to a radially-dependent change in index of refraction. This index of refraction change in turn created a divergent lens, explaining the laser power output changes. Notably, absorption of only 10^{-3} to 10^{-4} parts per cm is required to cause a significant thermal lens. Based on a parabolic approximation of the Gaussian beam, the focal length of the lens was derived to be:

$$F = \frac{k\pi n_o \omega_o^2}{0.24\mu_a Pl\left(dn/dT\right)} \tag{9.2}$$

where k is the thermal conductivity of the material [cal/cm • s • K], n_o is the initial bulk material index of refraction, ω_o is the spot radius [cm], μ_a is the absorption coefficient [cm^{-1}], P is the power [W], l is the length of interaction [cm], and dn/dT is the change in index of refraction [K^{-1}]. For example, in a 1 cm length of water, with a 1 W, 1064 nm laser with a spot radius of 1 mm, the focal length of the created lens is about –4.1 cm.

The heating and lens development has a characteristic time given by:

$$t_c = \frac{\omega_o^2}{4k/\rho c} \tag{9.3}$$

where ρ is the density [g/cm^3] and c is the specific heat [cal/g•K]. For the same parameters as above, the characteristic time is approximately 0.45 s.

Thermal lensing may affect both diagnostic imaging and therapeutic interventions. Since the focal length of the thermal lens is related to the square of the spot size, the very small spots used in scanning diagnostic imaging techniques such as confocal microscopy or optical coherence tomography can create short focal length thermal lenses despite the low powers typically employed. Thus, the potential exists to alter the beam profile in the tissue and affect image quality. However, the beam dwell time with these techniques is usually much shorter than the characteristic time

and the depth of focus is small, so in practice thermal lensing is only likely to be a problem if the device beam is stationary for an extended period of time, allowing heat build-up.

In therapeutic applications, the effects of thermal lensing can vary from mild to dramatic. Venkatesh et al., investigated effects of Q-switched Nd:YAG laser (25–30 ns pulse duration, 18–22 mJ energy, 2 mm spot size) on the cornea [19]. While the laser pulses were short enough that thermal lensing did not affect the first pulse, subsequent laser pulses through the same region of the cornea were qualitatively broadened. The thermal lens persisted with a characteristic time of 2.3 s. Their findings are in agreement with observed phenomena during anterior chamber laser surgery, including a focus shift that was dependent on the number of pulses fired with a Q-switched Nd:YAG laser [20]. The observed 3 mm change in focal distance could have implications for the safety and efficacy of procedures such as trabeculotomy.

Quantification of the effect of thermal lensing in water and soft tissues was provided by Ith et al., who measured the two dimensional beam profile of a Ho:YAG laser through varying thickness of material [21]. The parameters of the laser were 2.12 μm wavelength, 250 μm pulse duration, approximately 190 μm spot size, and pulse energy between 4.4 and 10.2 mJ. They found a pulse energy-dependent broadening of the laser beam. For example, at 10.2 mJ energy, the spot size was broadened by up to 25% after transmission through 300 μm of water. Measurements in cartilage showed additional broadening due to scattering in the sample; however, the thermal lensing effect was similar to that of water, suggesting that the effect in most soft tissues can be approximated as that of water. They conclude that during ablation with a Gaussian beam, the effects of thermal lensing are likely to be (1) a decrease of the peak temperature leading to an increase in the ablation threshold, and (2) a decrease in the axial temperature gradient leading to an increased axial extent of damage.

Motamedi et al. showed the existence of thermal lensing in tissue-mimicking phantoms [22]. In their experiments, a 2.25 W Argon laser with 1.0 mm spot size was incident on a 1 cm cuvette with a dye solution ($\mu_a = 3.3\ \mathrm{cm}^{-1}$) or dye and egg white mixture ($\mu_a = 2.0\ \mathrm{cm}^{-1}$). In the case of the dye solution cuvette, within 233 ms of laser irradiation the transmitted beam expanded to 12 times its original size (at a screen 70 cm distant from the sample) and interference fringes were seen due to spherical aberration of the thermal lens. At later times, convection in the sample caused the beam pattern to become asymmetric. With the dye-egg white material, similar results were seen until coagulation was achieved, which had the effect of increasing attenuation and decreasing beam expansion. The investigators also found an inverse dependence on beam expansion with initial spot size, which is consistent with equation (Eq. (9.2)). This experiment suggests that thermal lensing can be a serious consideration during irradiation of moderately absorbing samples, for example skin. Affonce and Fowler performed Monte Carlo and finite difference simulations to predict the effects of thermal lensing during cutaneous irradiation [23]. They found that the depth of penetration of light is nearly unchanged, but that thermal lensing causes a decrease in epidermal fluence and increase in dermal

fluence- an effect that is beneficial if dermal blood vessels are the therapeutic target. The effect is most pronounced for small diameter beams with Gaussian profiles; up to a 23% reduction in epidermal fluence and 35% increase in dermal fluence was found for a 2 mm diameter beam with 20 J/cm^2 radiant exposure. For large, flat top profile beams, the effect of thermal lensing was small, due to the minimal temperature gradient in the center of the beam.

Vincelette et al. found that for near IR wavelengths (1150–1390 nm), thermal lensing protects the retina from laser induced damage. Thermal lensing increases the minimal retinal spot diameter beyond the value attributed to chromatic dispersion [24].

Overall, the possible effects of thermal lensing should be considered when working with small diameter Gaussian beams and long pulse durations or multiple pulses at a single location. High power and moderate absorption coefficient will also increase the possibility of beam expansion or focal distance shifts from thermal lensing.

9.4 Photothermal Conversion of Oxyhemoglobin to Methemoglobin

Based on the knowledge that oxygen solubility in blood plasma decreases exponentially with temperature, Halldorsson estimated that initially 100% oxygenated blood would be completely deoxygenated (as well as thermally coagulated) at temperatures above 80°C [25] under slow heating. He observed that, at 633 nm, the optical absorption of thermally denatured blood was higher than that of oxygenated blood, consistent with the conversion of HbO_2 to Hb. At 1064 nm however, he obtained the result, deemed "confusing" at the time, that the absorption of thermally denatured blood was also higher than that of oxygenated blood. If the mechanism were thermal deoxygenation, one would have expected the absorption at 1064 nm to decrease compared to oxygenated blood. Increased scattering of the denatured blood was postulated as the reason for this anomalous finding.

However, the brownish-red color of thermally coagulated blood suggests a different chromophore than either HbO_2 or Hb. Using coagula created with 532 nm, 10 ms laser pulses, Barton et al. spectrophotometrically identified this species as methemoglobin (metHb), a form of hemoglobin in which the iron exists in the Fe(III) oxidation state but where it is incapable of exchanging molecular oxygen in tissue [26]. While Hb may occur as a transient intermediate of heated, fully oxygenated blood, there was no evidence for a significant fraction of Hb in the final blood coagulum. They further showed that metHb is created at radiant exposures well below those necessary for permanent blood vessel destruction [26], and occurs at temperatures of approximately 76°C when using a 10 ms, 532 nm laser pulse [27]. Unlike the bathochromic shift in absorption, which occurs prior to any permanent alteration of the HbO_2 molecule (at temperatures lower than 76°C for a 10 ms, 532 nm laser pulse), photothermal conversion to metHb happens concurrent with, and

perhaps because of, thermal denaturation of HbO_2, and is an irreversible process. It has been shown spectrophotometrically that metHb formation also occurs in vivo in port-wine stain (PWS) and telangiectasia patients treated with 0.45 ms 585 nm laser irradiation. Randeberg et al. estimated a final metHb concentration of approximately 21% of Hb in the skin using reflectance models [28]. Since metHb formation has been shown to form at green, orange, as well as NIR wavelengths (10 ms 1064 nm laser irradiation [12]), metHb formation appears to be a thermal process, at least for lasers within the visible-NIR wavelength and millisecond pulse duration ranges.

MetHb has a greater absorptivity than HbO_2 throughout the red and near infrared, for example approximately 30x higher at 632 nm and 3x higher at 1064 nm. The absorptivity of MetHb is shown in Fig. 9.4. This greater absorptivity is of particular interest at the commonly available Nd:YAG laser wavelength 1064 nm. This wavelength has been used for deep/large blood vessel treatment because of weak attenuation in tissue thus deep tissue penetration; however, it is so poorly absorbed in blood that pain and collateral damage can occur during treatment. Barton et al. showed in hamster skin flap window models that a subtherapeutic pulse of green (532 nm) light plus a subtherapeutic pulse of 1064 nm light (delayed 6 ms after the start of the 532 nm pulse) caused efficient and lasting (no blood flow at 24 hours) coagulation of blood vessels [26]. Safety may be enhanced due to the use of lower radiant exposures at both wavelengths. In a two-wavelength treatment protocol, the dynamic optical properties are used to advantage- a sufficient fluence of green laser light may not be available at a great enough depths to permanently damage deep and/or large diameter blood vessels, but sufficient energy may be deposited

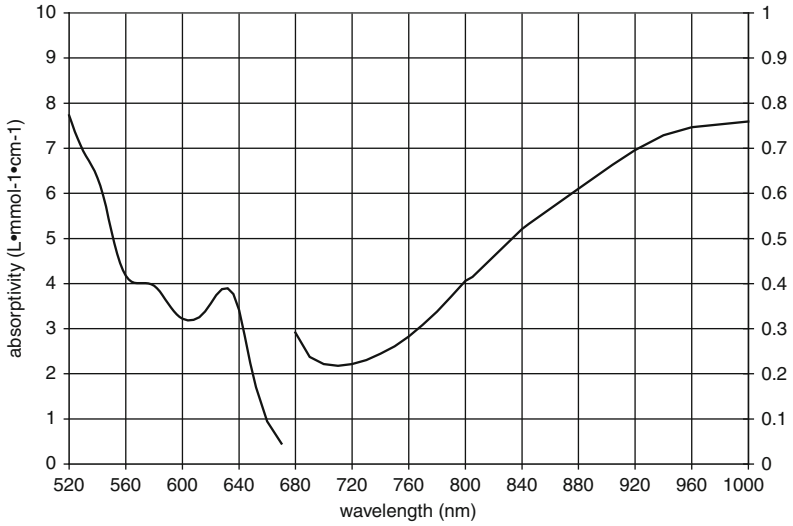


Fig. 9.4 Room temperature absorptivity of MetHb. Data from Zijlstra et al. [15]. Data for wavelengths above 680 nm have been multiplied by 10x to show spectral detail

to raise the temperature of blood high enough for the creation of metHb. A subsequent 1064 nm laser pulse penetrates deeply and is more efficiently absorbed in the metHb-containing blood, causing enough heat generation to damage the vessel.

The formation of metHb also may be advantageous for single wavelength, multiple pulse treatment. Ross et al., Tanghetti et al., and Rohrer et al. have demonstrated successful pulse stacking techniques with dye lasers that imply a significant persistence of metHb from one pulse to the next, even with the effects of perfusion and vascular fenestration (openings in the vessel wall) [29–31]. Mordon et al. attributed metHb formation to the very high efficacy of leg vein clearance using a non-uniform pulse sequence 1064 nm laser [32]. The laser was designed to accommodate the increasing absorption of blood between laser pulses, due to metHb formation. A clearance of 98% of 1–2 mm diameter blood vessels was achieved after 3 sessions. Conversely, it must be noted that if dynamic optical property changes are *not* considered, adverse events may occur. Thermal modeling has shown that there is only a small range of 1064 nm radiant exposures between those needed for a stable coagulation end point and those that can cause cavitation in whole blood. Researchers have noted the potential in vivo for 1064 nm to cause full thickness burns [33, 34], possibly due in part to sudden large increases in the absorption coefficient of blood at 1064 nm when metHb is formed.

9.4.1 Combined Bathochromic Shift and metHb Formation

The two sequential effects of the bathochromic shift and metHb formation can cause very dramatic changes in the absorption of blood during photocoagulation. At 532 nm, both the bathochromic shift and conversion of HbO_2 to metHb will tend to decrease absorption; around 585 nm, the two effects may balance each other, with an increase in absorption due to the bathochromic shift countered by the lower absorptivity of metHb. From 592 nm to about 700 nm, very large increases in absorption coefficient can occur as the two effects combine in an additive manner. With a pump wavelength of 1064 nm, Black et al. showed dramatic decreases in transmission through cuvettes of whole blood (200 μm thickness) at probe wavelengths of 594, 612, and 633 nm [12]. From pre-irradiation transmissions of approximately 9, 28, and 57%, respectively, the transmission decreased to values between 0.1 to a few percent. While increased scattering also played a role, the predominant reason for the transmission drop was a large increase in absorption. Figure 9.5 shows an estimate of the change in the 633 nm absorption coefficient of a 200 μm thickness cuvette of whole blood as a function of time, when irradiated with a 10 ms, 100 J/cm^2, 1064 nm laser, based on data in [12]. Three phases can be seen. The slight increase during the first 4 ms can be attributed to the reversible bathochromic shift of the absorptivity of HbO_2 during heating. The more dramatic increase during the next 6 ms is due to the formation of metHb, which has an absorptivity about 30x greater than HbO_2. Some slow continued development of the absorption coefficient is seen as the blood cools post-laser pulse. The peak temperature of the blood is estimated to be about 76°C at 4 ms into the laser pulse, when metHb creation begins.

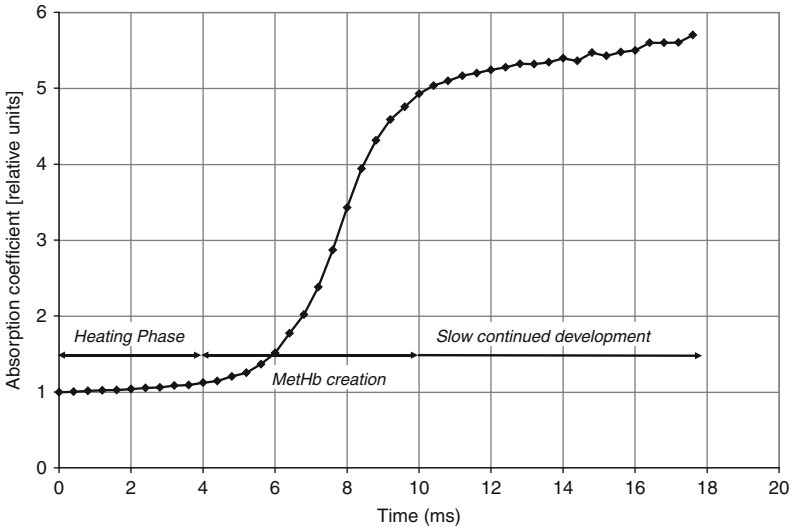


Fig. 9.5 Relative increase in the absorption coefficient of whole blood estimated from integrating sphere measurements at 633 nm, during and immediately after a 10 ms, 1064 nm, 100 J/cm^2 laser pulse [12]. The slight increase during the first 4 ms can be attributed to the bathochromic shift of the absorption resonances of HbO_2. The more dramatic increase during the next 6 ms is due to the formation of metHb, which has an absorptivity about 30x greater than HbO_2. Some slow continued development of the absorption coefficient is seen as the blood cools post-pulse

Kimel et al. showed that the threshold radiant exposure needed for creation of permanent gas bubbles in capillary tubes of blood was independent of wavelength over the range 585–600 nm [35]. This result, quite surprising if only room-temperature optical properties of blood are considered, could be understandable if bathochromic shifts and conversion of Hb to metHb were included in predictive modeling. In vivo, Pikkula et al. showed that for equivalent purpura, 585 and 595 nm produced no statistically significant difference in vascular damage, again suggesting that the bathochromic shift and formation of metHb play a role in creating more parity in vascular damage between the two wavelengths than would be expected based on their respective "native" absorption coefficients alone [36]. Animal experiments demonstrated the increased efficacy of two-wavelength (532 and 1064 nm), multiple-pulse treatments over single-wavelength and single-pulse treatments for permanent coagulation of blood vessels, attributed to factors including the creation of metHb [37]. In sum, these observations reinforce the knowledge that the selection of optimum wavelengths for therapeutic vascular lasers cannot be accomplished through a priori consideration of the room-temperature Hb absorption spectra alone. A complete prediction of outcome requires knowledge over all time of temperature distribution and the resultant concentration and temperature-dependent absorption spectrum of HbO_2, Hb, and metHb. While a complete model of these dynamic optical properties does not yet exist, an estimate of the effects of absorption changes could be made by assuming an initial concentration of 100% HbO_2,

temperature-dependent absorption changes due to the bathochromic shift as sketched in Table 9.1 up to the coagulation temperature, and then 70% conversion within the irradiated volume of HbO_2 to metHb, as noted in Black et al. [14].

9.5 Protein Denaturation

The increase in scattering due to thermal denaturation of proteins is a well-known and ubiquitous phenomenon. For example, it has also been shown that after 300 s in a constant temperature (100°C) water bath, the reduced scattering coefficient μ_s' of aorta increases 10–45% in the visible wavelength range and over 150% in the near infrared range [38].

9.5.1 Tissue Denaturation and Feedback

The concomitant increase in reflected light with tissue denaturation has been used as a feedback mechanism during precision photothermal therapies, especially retinal photocoagulation. Jerath et al. performed numerical modeling and experimentation with albumin/black paint retinal tissue phantoms to determine the relationship between reflectance and laser-induced lesion diameter and thickness using an Argon laser [39]. They found a linear and strong relationship between central reflectance and lesion thickness up to approximately 100 μm. They also noted that the fluence of light reaching the paint layer (corresponding to the retinal pigment epithelium in the eye) changed during lesion development, due to the growth of the overlying highly-scattering coagulated region. The dynamic change in the rate of heat generation means that lesion size cannot be predicted using static optical properties, however depth may be controllable using reflectance feedback. A retinal photocoagulation device based on this principle has been utilized in vivo [40].

Another feedback control option is to stabilize the surface temperature of irradiated tissue, thus controlling the downstream effects of dynamic optical properties. Temperature controlled feedback has been shown to improve the short-term efficacy of laser-welded rat intestinal anastomoses [41]. Temperature feedback has also been used to create solder-enhanced laser welds in rat skin that have comparable strength to sutures, minimal thermal damage, and heal with good cosmesis [42].

9.5.2 Red Blood Cell Membrane Protein Denaturation

In the case of red blood cells, protein denaturation causes changes in membrane conformation leading to changes in the scattering properties of isolated cell suspensions and whole blood. Nilsson et al. noted an abrupt increase then decrease in the reduced scattering coefficient of slowly heated blood at approximately 46°C and 49°C, respectively, due first to a change in shape from biconcave to spheroid, then to fragmentation of the membrane [43]. These effects have been noted for short

pulse laser irradiation as well; they occur at approximately 76 and 80°C, respectively, for blood heated with laser parameters 532 nm wavelength, 10 J/cm^2 radiant exposure, and 10 ms pulse durations [27]. Upon heat diffusion to, and coagulation of, the plasma, a further increase in scattering is noted. These changes in scattering will affect light distribution and energy deposition in the blood, although perhaps not to the same extent as dynamic changes in Hb absorption discussed above [12, 14, 26].

9.5.3 Modeling of Laser Treatment During Denaturation

Dynamic changes in scattering due to thermal denaturation also have implications for laser treatments requiring bulk tissue treatment. A prime example is interstitial laser photocoagulation for lesion formation in the liver. Frequently, large lesions are desired from a single fiber inserted into the tissue. Because energy deposition near the fiber tip is largest, coagulation occurs first at this site. Both the optical and thermal properties of the coagulated tissue vary greatly from the native liver, leading to an alteration in fluence, energy deposition, and temperature throughout the tissue. Due to the typically long exposure times, the cooling effects of blood perfusion and possible surface cooling must also be taken into consideration. Proper prediction of lesion diameter requires an optical-thermal model that takes into consideration changes in optical and thermal properties, and dynamically recalculates temperature field and thermal damage. A flowchart for such a simulation program is shown in Fig. 9.6. After initialization of parameters, the laser pulse is applied in a series of small time steps. After each step, the fluence, energy deposition, temperature distribution, and thermal damage are calculated. If damage has accumulated to the point where coagulation has occurred in any region, new tissue optical and thermal properties are determined. Based on these new properties, the fluence, energy deposition and temperature distribution are recalculated (inner feedback loop in Fig. 9.6). The process is repeated until the temperature field stabilizes. To facilitate convergence, the assumption may be made that fluence and energy deposition do not change and only the temperature distribution is recalculated. The time is incremented and the process is repeated until end of time (end of laser pulse) is reached. While not included in Fig. 9.6, dynamic changes in perfusion can also be incorporated into the model. Several authors have developed dynamic modeling programs including some or all of these factors, including Kim et al. [44] and London et al. [45].

As with all modeling, the results from this dynamic simulation are only as valid as the assumptions made. Various authors have performed modeling of laser interstitial photocoagulation of liver and come to different conclusions about the maximum lesion size possible and the effects of dynamic optical properties. Whelan and Wyman concluded that a lesion nearly 10 mm in diameter is possible without tissue charring, and that a six-fold increase in the reduced scattering coefficient of the tissue will increase the maximum coagulated diameter four-fold [46]. However, Jiang and Zhang found that estimated damage volume did not appreciably change when dynamic optical properties were considered, and that inclusion of dynamic thermal

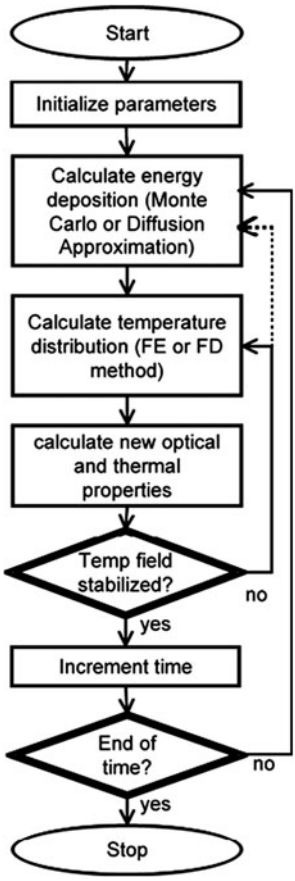


Fig. 9.6 Flow diagram for a dynamic temperature estimation program. In this model, optical and thermal properties are a function of temperature

properties reduced the damage volume somewhat [47]. Zhu et al. concluded that maximal coagulated volume decreases with consideration of dynamic properties [48]. Iizuka et al. used a nonlinear mathematical model to show both theoretically and experimentally that dynamic changes in optical properties have the largest effect on temperatures near the probe tip, that laser power requirements needed to achieve maximal damage volume without tissue charring are very different for static and dynamic cases, and that the maximal volume achieved with dynamic optical properties is considerably lower than with native properties [49]. Clearly, dynamic tissue properties will influence results obtained during actual tissue irradiation; yet the sensitivity of models to assumed parameters, and patient-to-patient variation of tissue properties may preclude the use of modeling alone to completely predict outcome. The inclusion of a feedback mechanism such as MRI may be useful and has been demonstrated clinically for temperature mapping and lesion size mapping [50]. Local temperature feedback additionally may facilitate controlled lesion size [51]. Ultrasound has also been used clinically to monitor lesion size [52].

9.6 Change in Water Content

As the major constituent of soft tissue, water content greatly affects tissue optical and thermal properties. In addition to being a strong absorber in the mid-infrared, water content will affect tissue density, conductivity, and index of refraction. Water exists both tightly bound to proteins and relatively freely moving. When passive and active mechanisms for controlling water content in tissue are interrupted (e.g. by dehydration, cell death, application of an agent, or pressure), dramatic changes in optical properties can occur.

9.6.1 Evaporation

The effect of water loss to the environment has been studied in several tissues. In general, ex vivo tissue undergoing evaporative water loss shows a decrease in thickness, increase in light transmission, increase in absorption coefficient, and relatively constant reduced scattering coefficient. The temperature dependent evaporation increases convective heat loss. However, there are some notable exceptions.

In human aorta samples, a weight loss of approximately 46% due to evaporation led to a thickness reduction of about 20%, a modest increase in transmission (about 10% in the near infrared wavelength range), and a more notable decrease in reflectance (about 20% in the visible). While reduced scattering was essentially unchanged, the absorption coefficient increased by 20–50% for wavelengths where water absorption was weak [38]. In another study of porcine liver, a weight loss of approximately 69% lead to a more dramatic thickness reduction of over 50%. As with aorta, the reduced scattering coefficient was essentially unchanged while the absorption coefficient increased about 146% at 632.8 nm [53]. The study authors postulated that the increase in absorption coefficient was due to shrinkage of tissue causing concentration of the tissue chromophores. In a third study, skin allowed to air-dry showed a 55% decrease in mass, increased transmission and decreased reflectance (the transmission to reflectance ratio increased over 200%) [54]. Absorption and reduced scattering coefficient were not calculated. This same group also exposed rat tail tendon fascicles to air for 2 hours, with a striking effect. Optical coherence tomography images of the tendon showed a 25% reduction in optical thickness and a near-total loss of integrated backscattered light intensity within the fascicle, strongly suggesting a decrease in reduced scattering coefficient. TEM images of the cut cross-sections of fascicles showed that fibril diameter and shape remained fairly constant during evaporation, but that the packing density increased from approximately 0.65 to 0.90. Using a simple heuristic particle-interaction model that predicts an inverted parabolic dependence of reduced scattering coefficient on scattering particle volume fraction (highest scattering at a particle volume fraction of 0.5), it was calculated that the increased fibril packing due to air drying led to a 60% reduction in reduced scattering coefficient. This strong decrease calculated for tendon contrasts with the lack of appreciable change

in reduced scattering coefficient noted in other tissues. Tendon is unique in its organized fibrillar structure and low cellular fraction, which may account for the differences noted.

Water content in the cornea is controlled by a balance of leak into the stroma controlled by the ionic permeability of the endothelium, and pumps in the endothelium which actively move fluid out of the stroma [55]. A specific water concentration is necessary to maintain the hexagonal arrangement of collagen fibrils in the cornea that ensure the transparency of the cornea. Therefore, both corneal edema and corneal dehydration increase light scattering. When the eyes of rabbits in vivo were exposed to air for one hour, a corneal thickness decrease of approximately 12%, and an increase in optical coherence tomography signal (associated with light backscattering) was noticed [56]. Clearly, evaporative water loss causes changes in optical properties, although the magnitude and sign of these changes depends on the tissue type. Care needs to be taken during optical property measurements and imaging or therapy of tissues, especially with ex vivo samples, to maintain in vivo-like levels of hydration.

9.6.2 Pressure

Pressure can be deliberately applied to tissue to enhance diagnostic and therapeutic procedures. For example, applying probe pressure during transscleral cyclophotocoagulation will decrease local water concentration, increase light transmission through the sclera and increase the fluence reaching the target ciliary body. As another example, pressure on human colon tissue increases the depth of imaging of optical coherence tomography, enabling visualization to the level of the submucosa [57]. Chan et al. noted fluid leaking from specimens under compression [58]. They speculated that pressure caused shape changes of cellular components, index matching and increased scatterer and chromophore concentration. With the possible exception of component shape changes, pressure appears to cause similar effects to evaporation. Therefore, it is not surprising that the optical property changes in soft tissue are also similar: a decrease in reflection and an increase in transmission, absorption coefficient and reduced scattering coefficient were generally found for ex vivo samples of human skin, bovine aorta, and bovine and porcine sclera under pressure [58]. In vivo, displacement of water is also postulated as a mechanism for alterations in breast tissue absorption and reduced scattering coefficient during breast compression, measured with a diffuse near-infrared tomography system [59]. Besides water movement, vascular effects need to be considered in vivo; an increase in vascular volume was also expected during sustained application of pressure.

Local vacuum pressure has been applied to skin in an effort to dilate blood vessels and increase efficacy of laser treatment of port wine stains [60]. Visible reflectance spectroscopy in conjunction with the diffusion approximation model was used to estimate changes in optical properties of a normal volunteer's skin placed under 50 kPa of vacuum pressure. A 5-fold increase in blood volume fraction was calculated, as well as an approximately 25% decrease in epidermal absorption coefficient

in the visible wavelength range. This decrease was attributed to the stretching and thinning of the epidermis, which reduced the concentration of the chromophore melanin.

9.6.3 Replacement of Water with Biocompatible Fluid

The strong absorption of mid-infrared light (e.g. Ho:YAG laser at 2.12 μm, Er:YAG laser at 2.94 μm, and free electron laser at 6.45 μm wavelength) in water enables precise tissue ablation. However, unbound, overlying water can quickly attenuate light to a non-therapeutic radiant exposure before it reaches the tissue. One solution to this problem is to place a light-guiding fiber in contact with the tissue. However, in certain applications, such as vitreoretinal surgery, a non-contact approach is desired to avoid unwanted damage to delicate structures. Another option is to exploit the creation of vapor bubbles with Q-switched and microsecond pulsed lasers pulses at these wavelengths. A portion of the laser energy is used to create the bubble, which then creates a low-absorption "channel" for the remainder of the pulse, enabling tissue ablation [61]. However, the strong pressure transients associated with bubble formation and collapse can cause unwanted mechanical damage to the tissue.

A third option that has been explored is to replace the unbound water with a biocompatible fluid that has low absorption in the mid-infrared, e.g. perfluorocarbon. Wesendahl et al. replaced the vitreous humor of enucleated pig eyes with perfluorodecaline, and measured the effects of free-running Er:YAG ablation on the retina [62]. It was possible to maintain a 1 mm separation between the fiber tip and the retinal tissue yet achieve precise tissue ablation with no measureable pressure transients. Mackanos performed a series of experiments in vitro using various types of perfluorocarbons, fiber-target separations and both 2.94 and 6.45 μm wavelength lasers [63]. They found that although absorption was higher in perfluorocarbon at the longer wavelength, precise ablation was possible using high-boiling point compounds (perfluorodecalin) with the fiber in close proximity to the target (about 0.33 mm). An advantage of the 6.45 μm wavelength is that the ablation crater depth was self-limiting because of absorption in the perfluorocarbon; as ablation progressed the distance to the bottom of the ablation crater increased and the radiant exposure at the bottom decreased.

9.7 Optical Clearing Agents

A variety of agents have been shown to have "optical clearing" capability in tissue. Optical clearing has variously been defined as a reduction in the tissue μ_s or μ_s', a reduction in μ_a, an increase in collimated or total transmission, and/or a decrease in the diffuse reflection of light in tissue. Modulated optical properties may enhance visualization of deeper tissue structures, improve tomographic image resolution, and enable deeper penetration of therapeutic levels of laser light. The process can be reversible, with little or no apparent tissue damage. Although understanding of the

mechanisms and applications of optical clearing is incomplete at this time, the below discussion summarizes current thinking.

Three mechanisms have been postulated for the clearing effect: (1) replacement of tissue fluid with an agent of higher refractive index, (2) dehydration, and (3) collagen structure alteration. In most cases these effects work in combination, although the relative contribution of each mechanism may depend upon the type of tissue studied and agent used. A partial list of the substances used as optical clearing agents is given in Table 9.2.

The decrease in scatter caused by increasing the background index of refraction of a suspension is well established. Chance et al. added sugars and electrolytes to lipid and yeast cell suspensions and measured a decrease in scattering factor [64]. Liu et al. showed that addition of mM concentrations of glucose or mannitol decreased the μ_s' of lipid suspensions, while the μ_a remained constant [65]. Tuchin et al. calculated that a 10-fold reduction of the μ_s' of blood could be expected when the refractive index of plasma was increased from the nominal value of 1.340 up to 1.388, assuming the erythrocyte cytoplasm index remained constant at 1.412 [66]. Because the refractive index of interstitial fluid is only slightly higher than water (1.33), whereas that of other tissue constituents is much higher (hydrated collagen n=1.43 [67], melanin n=1.7 [68], and adipose n=1.46 [69]), it can be expected that diffusion of relatively high refractive index agents down their concentration gradient into tissue will cause a decrease in tissue scattering coefficient. However, if the agent is cell-permeable, then increases in both cellular and interstitial index of refraction may occur and negate the initial decrease in scattering [65].

In tissues, the osmolarity of the applied agent can induce water movement and cause a dramatic alteration of tissue physical, optical, and mechanical properties. Hyperosmotic agents can cause both interstitial and intracellular water diffusion out of the tissue. The effect of this dehydration is two-fold: first the decrease in water content causes tissue shrinkage and greater transmission of wavelengths strongly absorbed in water. Second, the attendant increase in packing of scattering centers can decrease μ_s' and increase broadband transmission. Higher packing fraction may decrease the refractive index variations and possibly create larger size scatterers. In practice, index-matching is often caused by a combination of both the mechanisms of dehydration and fluid replacement. Tuchin et al. described two possibilities for refractive index matching: (1) water can leave the interstitial space and exit the sample; (2) the administered fluid can enter the tissue and fill up the interstitial space, and water can partly leave the interstitial space and exit the sample [70]. While the first mechanism is characteristic for ideal osmolytes, the second describes such real osmolytes as polyethylene glycol, Trazograph-60, and glucose. While the osmotically-driven dehydration mechanism is different from passive evaporation, the effect on tissue structure and optical properties has been shown to be qualitatively similar [54]. The action of an ideal osmolyte also can be compared to that caused by mechanical stress, due to the similarity in the mechanisms for inducing fluid flux. Therefore, one also could expect some similarities in the action of optical clearing agents to that caused by pressure.

Water movement caused by hyperosmotic agents can be visualized readily. For example, when a piece of tissue is submerged in glycerol for sufficient time (about

Table 9.2 Partial list of substances investigated as optical clearing agents

Agent	Properties	Osmolality (Osm/kg)	Refractive index	Application reference(s)
Glycerol (Glycerin)	polyalcohol	2.6% v/v isotonic with serum	1.489	[71–74, 76, 77, 81, 83–90, 92, 93, 96]
1,3-Propanediol	polyalcohol		1.433	[64, 76]
1,2-Propanediol (Propylene glycol)	polyalcohol	2.0% v/v isotonic with serum	1.43	[66, 89, 94]
1,4-Butanediol	polyalcohol	26.9	1.44	[73]
1–2-Ethanediol (Ethylene glycol)	polyalcohol	22.6	1.422	[72, 73]
Polyethylene glycol (PEG)	polyalcohol		1.452 (PEG 400)	[70, 79]
Sorbitol	polyalcohol	5.48% w/v isotonic with serum	1.574	[65, 76, 77]
Mannitol	polyalcohol	5.07% w/v isotonic with serum	1.597	[64, 65]
Glucose (Dextrose)	monosaccharide	5.51% w/v isotonic with serum	1.573	[64–66, 70, 77, 83, 85, 94]
Fructose	monosaccharide	5.05% w/v isotonic with serum	1.574	[64, 77]
Sucrose	disaccharide	9.25% w/v isotonic with serum	1.654	[64, 65, 77]
Dextran	polysaccharide		1.651 (MW 504)	[94, 95]
Dimethyl sulfoxide (DMSO)	aprotic solvent	15.2	1.479	[71, 73, 85, 90]
Oleic acid	protic solvent		1.466	[73]
Trazograph-60	x-ray contrast agent		1.46 (76%v/v)	[70, 94]
Synergistic Compounds				
Polypropylene glycol/polyethylene glycol	polyalcohols		1.46/1.452	[91]
Oleic acid/propylene glycol	protic solvent/polyalcohol		1.465/1.43	[79]
DMSO/propylene glycol	aprotic solvent/polyalcohol		1.479/1.43	[79]

20 min), a layer of fluid can be seen surrounding the tissue, distinguishable by its varied index of refraction. It can be assumed that this layer is tissue water. In vivo, injection of a small bleb of glycerine under the skin of a rodent causes the characteristic appearance illustrated in Fig. 9.7. After 30 min, the bleb subsides and the

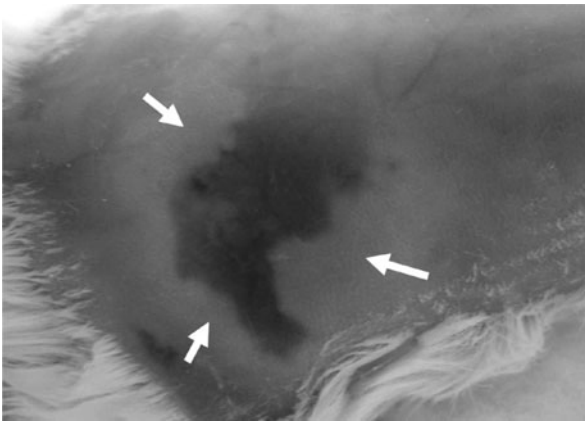


Fig. 9.7 Photograph of dorsal skin of a rodent, under which approximately 0.1 ml of anhydrous glycerol was injected. At 30 min, the tissue at the injection site has taken on a reddish appearance, due to increased transmission and visualization of the underlying muscle layer. Surrounding this site, a whitish, edematous ring is noticed (*arrows*). This appearance is consistent with water movement from the glycerol treated site to the surrounding tissue

overlying skin takes on a dark, reddish appearance as the tissue thins, transmission of light increases, and the underlying muscle layer is visualized. At the periphery of this cleared zone, an edematous ring is noticed, characterized by a whiter, slightly raised appearance. This appearance is consistent with water movement from the treated area to the surrounding skin [71]. Water desorption was studied quantitatively by Xu and Wang using near infrared reflectance spectroscopy [72]. Using excised porcine muscle tissue immersed in glycerol and ethylene glycol, they found that there exist three stages of water desorption, strongly correlated with increases in light penetration measured with optical coherence tomography. Stage one occurs in the first minute and is associated with bulk water displacement by the clearing agent and the most dramatic increase in penetration depth. Stage two occurs during 1–10 min and is associated with a slow linear water loss and penetration depth increase. During stage 3, occurring between 10 and 30 min, water loss equilibrates and attenuation of the tissue increases slightly. Glycerol can cause protein-bound water to desorb and is thus a more effective clearing agent than ethylene glycol.

If replacement of tissue fluid by an agent with higher index of refraction, and dehydration of tissue were the only mechanisms of optical clearing, one would expect there to be some correlation between optical clearing capability and the index of refraction and/or the osmolarity of the agent. However, it has been shown that the optical clearing potential of substances applied to the dermal side of in vitro human skin has no correlation with either osmolality or index of refraction [73]. Therefore, for skin at least, there must be another mechanism of action. In this study, Choi et al. found that hydrophilic hydroxy-terminated agents (such as glycerol) caused the greatest percent reduction in tissue μ_s'. The hydrophyllic organic solvent DMSO performed less well, and lipophilic organic acids such as oleic acid actually increased μ_s', suggesting that these agents did not diffuse through

the dermis [73]. It was shown by Yeh and colleagues that glycerol screens noncovalent intermolecular interactions between low-order collagen structures, resulting in collagen fiber disassembly and reduction in scattering. The clearing effect can be inhibited by formaldehyde fixation, which prevents disassociation [74, 75]. It has been found that the reduction in μ_s' of excised rodent and human skin is correlated with a clearing agent's ability to inhibit collagen self-assembly [76]. Inhibition of fibril assembly was found to increase with sugar alcohol chain length, with sorbitol being twice as effective as glycerol. In studies, sucrose, fructose, and dextrose were also found to have twice the clearing capability as glycerol [77].

In practice, optical clearing agents can be limited by diffusion barriers, perhaps most dramatically by the stratum corneum of the skin. Various approaches have been investigated to increase availability of agents. A synergistic effect can be achieved by the combination of hydrophilic agents such as glycerol or propylene glycol with a dipolar aprotic solvent such as DMSO [78, 79] or a lipophilic fatty acid such as oleic acid [79]. Other methods rely on mechanically disrupting the stratum corneum, and include the use of a Er:YAG laser ablation [80], 980 nm diode laser irradiation in conjunction with a topically-applied absorbing layer [81], flashlamp irradiation [82], gentle abrasion of the skin with sandpaper [83], and tape stripping [84]. Stumpp et al. concluded that sandpaper rubbing was the most effective mechanical method for increasing the penetration of optical clearing agents, but was associated with limited adverse effects including scabbing and erythema [85].

Optical clearing agents may improve optical diagnostic and therapeutic performance. Vargas et al. found that optical clearing of ex vivo and in vivo skin exposed to subdermal DMSO, glucose, and glycerol increased the signal from a fluorescence target placed on the subdermal side of the skin, when illumination and collection were performed from the epidermal side. The signal increase was greatest with DMSO; an approximately 250% increase was seen after 20 min [85]. He and Wang investigated low-light level imaging performance with Monte Carlo simulations and in vitro. They found that optical clearing with 50% glycerol caused a stronger peak signal intensity (factor of ~5) and a narrower full-width half-maximum (FWHM) (factor of ~3) of the detected spot from point-like luminol-hydrogen peroxide-horseradish peroxidase chemoluminescence emitting target (425 nm center wavelength) under 1–3 mm of porcine skin tissue [86]. Jansen et al. examined a diffusing fiber tip connected to a filtered tungsten lamp (wavelengths 550–650 nm) placed under mouse skin, and also found that treatment with 50–100% glycerol narrows the FWHM of the detected spot (factor of ~2) [87]. However, they found that the detected signal decreased by approximately a factor of 2. They attributed both signal changes to a decrease in detection of multiply scattered light. These findings highlight the importance of considering the source and imaging geometry when predicting how optical clearing agents will affect the detected signal. Optical clearing agents increase total transmission of light and decrease scattered light. While this combination increases resolution, it will not necessary increase detected signal. If the source is isotropic and the detector has a low numerical aperture (NA), which is the case in standard bioluminescence imaging systems, the detected signal may decrease after optical clearing to very low levels of μ_s' because photons that

might normally be scattered into the NA of the imaging system instead propagate out of range. There exists an optimum μ_s' for a given source depth, source emission pattern, and detector NA. An optical clearing agent ideally should decrease μ_s' only enough to reach that optimum condition. The disparate results of the luminescence studies are likely due to differences in the thickness of skin studied and the detector geometry.

Microscopy can benefit from optical clearing as well. Plotnikov et al. showed that application of glycerol to skeletal muscle tissue results in a 2.5-fold increase in second harmonic generation imaging depth. They showed that the periodicity of sarcomere structure is unaltered, and suggest that the mechanism of signal improvement is not index matching but rather a reduction in cytoplasmic protein concentration and decrease in the secondary inner filter effect on the SHG signal [88]. Several studies have shown that the optical coherence tomography signal can be improved after application of optical clearing agents. Both imaging depth and image contrast were improved in images of ex vivo tissue after the application of glycerol and propylene glycol [89, 90].

Optical clearing has been shown to improve the results of laser therapy. Khan et al. presented an example case of tattoo removal performed with or without optical clearing using a mixture of polypropylene glycol and polyethylene glycol (PPG:PEG). The PPG:PEG treated skin showed superior reduction in tattoo ink one month after a single laser treatment with a 532 nm Q-switched laser. Additionally, the treated skin did not experience the blistering seen in control skin. A Monte Carlo simulation suggests that the reason for the improved performance is more photons reaching the target tattoo ink, as well as a lower epidermal heat source term [91]. In a Guinea pig model of tattoo studied by McNichols et al., topical application of glycerol after tape stripping increased the response to green and red laser treatment. The effect was significant in light red tattoos treated with Q-switch 532 nm laser light [84].

Optical clearing is also effective in dramatically reducing the radiant exposure necessary to photocoagulate cutaneous blood vessels in an animal model. This reduction is caused by a combination of optical clearing of the overlying skin tissue (increasing laser fluence at the level of the blood vessels), changes in the optical properties of blood, and stasis of blood flow. Vargas et al. showed that subdermal application of glycerol caused a 4-fold to 10-fold reduction in the radiant exposure needed for a 50% probability of permanent coagulation of hamster skin flap blood vessels [92]. In a previous study, photographs and Doppler OCT images of hamster skin showed that subdermal application of glycerol caused a decrease in the skin thickness, dilation of blood vessels, decreased signal (backscatter) from blood, a darker appearance of venule blood, reduction of shadowing beneath the blood vessel, and stasis in the venules and sometimes arterioles [93]. Photographs of the subdermal blood vessels of a hamster skin flap during glycerol application are shown in Fig. 9.8, which illustrate some of these findings. The reduced signal from the blood suggests a decreased μ_s'. The dark appearance of the blood suggests that amount of Hb in the vessel (and thus μ_a) is not decreased. The reduction in shadowing could be due to a decreased μ_s' or the stasis in blood flow, which eliminates

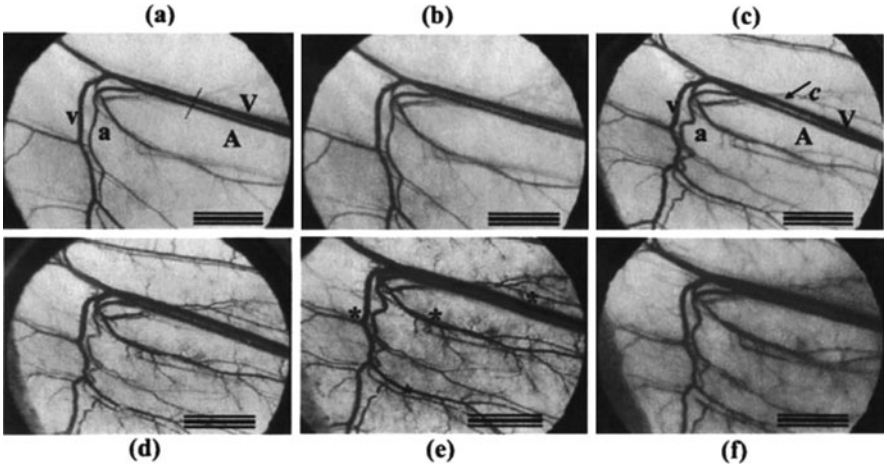


Fig. 9.8 Subdermal blood vessels of the in vivo hamster dorsal skin flap window preparation. (**a**) Native window preparation – a main venule (V) – arteriole (A) pair can be seen branching into smaller vessels (a and v) (*dashed line* represents areas of imaging using Doppler optical coherence tomography), (**b**) Window preparation immediately after the application of glycerol to the subdermal tissue, (**c**) Ten minutes after the application of glycerol, (**d**) Fifteen minutes after the application of glycerol, (**e**) Twenty minutes after the application of glycerol, (**f**) Five minutes after hydration with phosphate-buffered saline. c – focal constriction, * – areas of darkening. Scale bar = 0.25 cm. (Reprinted with permission from [93])

the Doppler shift of forward-scattered photons that may cause them to be shifted out of the bandwidth of the OCT system. Biologically, the osmotic changes caused by glycerol may cause plasma loss secondary to an inflammatory response, causing a change in blood viscosity and slowing of blood flow. If glycerol enters the blood, work by Tuchin indicates that the total attenuation coefficient will decrease due to index matching and possibly due to aggregation and sedimentation of the blood [94]. High molecular weight dextrans can be very effective clearing agents in blood [95]. Depending upon the concentration of glycerol in the blood, hemolysis may also occur, which will also decrease scattering. The changes induced in blood vessels by glycerol means that use of this agent for improved visualization of vasculature (e.g. cerebral blood flow under the intact dura [96]), must be performed with caution.

9.8 Conclusion

As has been shown in this chapter, the optical properties of tissues can be altered by several factors, including temperature (absorption resonance shifts, chemical conversion, denaturation), temperature gradient (thermal lensing), water content (dehydration, pressure-induced water loss, water replacement with a biologically compatible fluid or index matching fluid, osmotically-driven water loss), and shielding of intermolecular noncovalent bonds. For accurate, repeatable optical property

measurements and experimental results, ex vivo tissues should be handled carefully to maintain consistent temperature and water content, as close to the in vivo case as possible. The effects of other alterations, such as interruption of blood flow, metabolic changes, and tissue necrosis cannot be avoided in ex vivo experiments but can be minimized by prompt utilization of the tissue.

During laser therapy, the possibility of optical property changes causing unexpected outcomes should be considered. Each situation will vary, and not all of the changes discussed in this chapter will occur (or have a significant effect) for any one therapeutic protocol. As optical-thermal models become more sophisticated and temperature/damage-dependent optical properties become available, these dynamic effects will be predicted with greater accuracy.

References

1. Collins JR. Change in the infrared absorption spectrum of water with temperature. *Phys Rev*., 26:771–779 (1925).
2. Hale GM, Querry, MR, Rusk AN, and Williams D. Influence of temperature on the spectrum of water. *J. Opt. Soc. Am.*, 62:1103–1108 (1972).
3. Walsh JT Jr and Cummings JP. Effect of the dynamic optical properties of water on midinfrared laser ablation. *Lasers Surg. Med*., 15:295–305 (1994).
4. Cordone L, Cupane A, Leone M, and Vitrano E. Optical absorption spectra of deoxy- and oxyhemoglobin in the temperature range 300–320 K. *Biophys. Chem*., 24:259–275 (1986).
5. San Biagio, PL, Vitrano E, Cupane A, Madonia F, and Palma MU. Temperature induced difference spectra of oxy and deoxy hemoglobin in the near-IR, visible and Soret regions. *Biochem. Biophys. Res. Comm.*, 77:1158–1165 (1977).
6. Steinke JM and Shepherd AP. Effects of temperature on optical absorbance spectra of oxy-, carboxy- and deoxy-hemoglobin. *Clin. Chem*., 38:1360–1364 (1992).
7. Sfareni R, Boffi A, Quaresima V, and Ferrari M. Near infrared absorption spectra of human deoxy- and oxyhaemoglobin in the temperature range 20–40 degrees C. *Biochim Biophys Acta*., 1340(2):165–169 (1997).
8. Gray K and Slade EF. The temperature dependence of the optical absorption spectra of some methaemoglobin derivatives. *Biochem. Biophys. Res. Commun*., 48:1019–24 (1972).
9. Jansen ED, van Leeuwen TG, Motamedi M, Borst C, and Welch AJ. Temperature dependence of the absorption coefficient of water for midinfrared laser radiation. *Lasers Surg. Med*., 14:258–68 (1994).
10. Schomacker KT, Domankevitz Y, Flotte TJ, and Deutsch TF. Co:MgF2 laser ablation of tissue: Effect of wavelength on ablation threshold and thermal damage. *Lasers Surg. Med*., 11:141–151 (1994).
11. Lange BI, Brendel T, and Huttmann G. Temperature dependence of light absorption in water at holmium and thulium laser wavelengths. *Appl. Opt*., 41:5797–803 (2002).
12. Black JF, Wade N, and Barton JK. Mechanistic comparison of blood undergoing laser photocoagulation at 532 and 1,064 nm. *Lasers Surg. Med*., 36:155–165 (2005).
13. Verkryusse W, Nilsson AMK, Milner TE, Beek JF, Lucassen GW, and van Gemert MJC. Optical absorption of blood depends on temperature during a 0.5 ms laser pulse at 586 nm. *Photochem. Photobiol.*, 67:276–281 (1998).
14. Black JF and Barton JK. Chemical and structural changes in blood undergoing laser photocoagulation. *Photochem. Photobiol.*, 80:89–97 (2004).
15. Zijlstra WG, Buursma A, and van Assendelft OW. *Visible and near infrared absorption spectra of human and animal haemoglobin*. VSP, Utrecht (2000).
16. Kienle A and Hibst R. A new optimal wavelength for treatment of port wine stains? *Phys. Med. Biol.*, 40:1559–1576 (1995).

17. van Gemert, MJC, Welch AJ, Pickering JW, Tan OT, and Gijsbers GHM. Wavelengths for laser treatment of port wine stains and telangiectasia. *Lasers Surg. Med*., 16:147–155 (1995).
18. Gordon JP, Leite CC, Moore RS, Porto PS, and Whinnery JR. Long-transient effects in lasers with inserted liquid samples. *J. Appl. Phys*., 36:3–8 (1965).
19. Venkatesh S, Guthrie S, Cruickshank FR, Bailey R, Foulds W, and Lee WR. Thermal lens measurements in the cornea. *Br. J. Ophthalmol*., 69:92–95 (1985).
20. Fankhauser F, Roussel P, Steffen J, Van der Zypen E, and Chrenkova A. Clinical studies on the efficiency of high power laser radiation upon some structures of the anterior segment of the eye. *Int. Ophthalmol.* 3:129 (1981).
21. Ith M, Frenz M, and Weber HP. Scattering and thermal lensing of 2.12-μm laser radiation in biological tissue. *Appl. Opt*., 40:2216–2223 (2001).
22. Motamedi M, Welch AJ, Cheong W, Ghaffari SA, and Tan OT. Thermal lensing in biologic medium. *IEEE J. Quantum Electron*., 24:693–696 (1988).
23. Affonce DA and Fowler AJ. The effect of thermal lensing during selective Photothermolysis. *J. Quant. Spectrosc. Radiat. Transfer*, 73:473–479 (2002).
24. Vincelette RL, Thomas RJ, Rockwell BA, Lund DJ, and Welch AJ. Thermal lensing in ocular media exposed to continuous-wave near-infrared radiation: 1150–1350 nm region. *J. Bioimed Opt*., 13:054005 (2008).
25. Halldorsson, T. Alteration of optical and thermal properties of blood by Nd:YAG laser irradiation. *Proceedings of the 4th congress of the international society for laser surgery*, Tokyo, Japan, Nov. 23–27 (1981).
26. Barton JK, Frangineas G, Pummer H, and Black JF. Cooperative phenomena in two-pulse, two-color laser photocoagulation of cutaneous blood vessels. *Photochem. Photobiol.*, 73:642–50 (2001).
27. Barton JK, Popok D, and Black JF. Thermal analysis of blood undergoing laser photocoagulation. *IEEE J. Sel. Top. Quantum Electron*., 7:936–943 (2001).
28. Randeberg LL, Bonesronning JH, Dalaker M, Nelson JS, and Svaasand LO. Methemoglobin formation during laser induced photothermolysis of vascular skin lesions. *Lasers Surg. Med*., 34:414–419 (2004).
29. Ross EV and Domankevitz Y. Laser leg vein treatments: A brief overview. *J. Cosmetic Laser Ther*., 5:192–197 (2003).
30. Tanghetti EA, Sherr EA, and Alvarado SL. Multipass treatment of photodamage using the pulse dye laser. *Derm. Surg*., 29:686–691 (2003).
31. Rohrer TE, Chatrath V, and Iyengar V. Does pulse stacking improve the results of treatment with variable-pulse pulsed-dye lasers? *Derm. Surg*., 30:163–167 (2004).
32. Mordon S and Brisot D, Fournier N. Using a "non uniform pulse sequence" can improve selective coagulation with a Nd:YAG laser (1.06 micron) thanks to Met-hemoglobin absorption: A clinical study on blue leg veins. *Lasers Surg. Med*., 32:160–170 (2003).
33. Yang MU, Yaroslavsky AN, Farinelli WA, Tsao SS, and Anderson RR. Long-pulsed Nd:YAG laser treatment for port wine stains. *J. Amer. Acad. Derm*., 52(3):480–490 (2005).
34. Zanelli G and Narurkar VA. Full thickness scars from long pulsed 1064 nm lasers. *Lasers Surg. Med*., S16, 80:28 (2004).
35. Kimel S, Choi B, Svaasand LO, Lotfi J, Viator JA, and Nelson JS. Influence of laser wavelength and pulse duration on gas bubble formation in blood filled glass capillaries. *Lasers Surg. Med*., 36:281–288 (2005).
36. Pikkula BM, Chang DW, Nelson JS, and Anvari B. Comparison of 585 and 595 nm laser-induced vascular response of normal in vivo human skin. *Lasers Surg. Med*., 36:117–123 (2005).
37. Jia W, Choi B, Franco W, Lotfi J, Majaron B, Aguilar G, and Nelson JS. Treatment of cutaneous vascular lesions using multiple-intermittent cryogen spurts and two-wavelength laser pulses: Numerical and animal studies. *Lasers Surg. Med*., 39:494–503 (2007).
38. Cilesiz IF and Welch AJ. Light dosimetry: Effects of dehydration and thermal damage on the optical properties of the human aorta. *App. Opt*., 32:477–487 (1993).

39. Jerath MR, Gardner CM, Rylander HG 3rd, and Welch AJ. Dynamic optical property changes: Implications for reflectance feedback control of photocoagulation. *J. Photochem. Photobiol. B.* 16:113–126 (1992).
40. Jerath MR, Chundru R, Barrett SF, Rylander HG 3rd, and Welch AJ. Preliminary results on reflectance feedback control of photocoagulation in vivo. *IEEE Trans. Biomed. Eng.* 41:201–203(1994).
41. Cilesiz I, Thomsen S, and Welch AJ. Controlled temperature tissue fusion: Argon laser welding of rat intestine in vivo, part one. *Lasers Surg. Med*., 21:269–277 (1997).
42. Simhon D, Ravid A, Halpern M, Cilesiz I, Brosh T, Kariv N, Leviav A, and Katzir A. Laser soldering of rat skin, using fiberoptic temperature controlled system. *Lasers Surg. Med*. 29(3):265–273 (2001).
43. Nilsson AMK, Lucassen GW, Verkryusse W, Andersson-Engels S, and van Gemert MJC. Changes in optical properties of human whole blood in vitro due to slow heating. *Photochem. Photobiol*., 65:366–373 (1997).
44. Kim B, Jacques SL, Rastegar S, Thompsen S, and Motamedi M. Nonlinear finite-element analysis of the role of dynamic changes in blood perfusion and optical properties in laser coagulation of tissue. *IEEE Sel. Top. Quantum. Electron*., 2:922–933 (1996).
45. London RA, Flinsky ME, Zimmerman GB, Bailey DS, Eder DC, and Jacques SL. Laser-tissue interaction modeling with LATIS. *Appl. Opt*., 36:9068–9074 (1997).
46. Whelan WM and Wyman DR. Dynamic modeling of interstitial laser photocoagulation: Implications for lesion formation in liver in vivo. *Lasers Surg. Med*. 24:202–208 (1999).
47. Jiang SC and Zhang XX. Effects of dynamic changes of tissue properties during laser-induced interstitial thermotherapy (LITT). *Lasers Med. Sci*., 19:197–202 (2005).
48. Zhu D, Luo Q, Zhu G, and Liu W. Kinetic thermal response and damage in laser coagulation of tissue. *Lasers Surg. Med*., 31:313–321 (2002).
49. Iizuka MN, Vitkin IA, Kolios MC, and Sherar MD. The effects of dynamic optical properties during interstitial laser photocoagulation. *Phys. Med. Biol*., 45:1335–1357 (2000).
50. Wacker FK, Reither K, Ritz JP, Roggan A, Germer CT, and Wolf KJ. MR-guided interstitial laser-induced thermotherapy of hepatic metastasis combined with arterial blood flow reduction: Technique and first clinical results in an open MR system. *J. Magn. Reson. Imaging*., 13:31–36 (2001).
51. Wohlgemuth WA, Wamser G, Reiss T, Wagner T, and Bohndorf K. In vivo laser-induced interstitial thermotherapy of pig liver with a temperature-controlled diode laser and MRI correlation. *Lasers Surg. Med*., 29:374–378 (2001).
52. Giorgio A, Tarantino L, de Stefano G , Farella N, Catalano O, Cusati B, Del Viscovo L, Alaia A, and Caturelli E. Interstitial laser photocoagulation under ultrasound guidance of liver tumors: Results in 104 treated patients. *Eur J. Ultrasound*., 11:181–188 (2000).
53. Zhu D, Luo Q, and Cen J. Effects of dehydration on the optical properties of in vitro porcine liver. *Lasers Surg. Med*., 33:226–231 (2003).
54. Rylander CG, Stumpp OF, Milner TE, Kemp NJ, Mendenhall JM, Diller KR, and Welch AJ. Dehydration mechanism of optical clearing in tissue. *J. Biomed. Opt.*, 11(4):041117 (2006).
55. Fischbarg J and Maurice DM. An update on corneal hydration control. *Experimental Eye Res.*, 78:537–541 (2004).
56. Hosseini K, Kholodnykh, Petrova I, Esenaliev R, Hendrikse F, and Motamedi M. Monitoring of rabbit cornea response to dehydration stress by optical coherence tomography. *Invest. Ophthalmol. Vis. Sci.*, 45:2555–2562 (2004).
57. Sivak MV Jr, Kobayashi K, Izatt JA, Rollins AM, Ung-Runyawee R, Chak A, Wong RC, Isenberg GA, and Willis J. High-resolution endoscopic imaging of the GI tract using optical coherence tomography. *Gastrointest. Endosc*., 51:474–479 (2000).
58. Chan EK, Sorg B, Protsendo D, O'Neil M, Motamedi M, and Welch AJ. Effects of compression on soft tissue optical properties. *IEEE J. Sel. Top. Quantum. Electron*., 2:943–950 (1996).
59. Jiang S, Pogue BW, Paulsen KD, Kogel C, and Poplack SP. In vivo near-infrared spectral detection of pressure-induced changes in breast tissue. *Opt. Lett.*, 28:1212–1214 (2003).

60. Childers MA, Franco W, Nelson JS, and Aguilar G. Laser surgery of port wine stains using local vacuum pressure: Changes in skin morphology and optical properties (part I). *Lasers Surg. Med*., 39:108–117 (2007).
61. Jansen ED, Asshauer T, Frenz M, Motamedi M, Delacretaz G, and Welch AJ, Effect of pulse duration on bubble formation and laser-induced pressure waves during holmium laser ablation. *Lasers Surg. Med*., 18:278–293 (1996).
62. Wesendahl T, Janknecht P, Ott B, and Frenz M. Erbium:YAG laser ablation of retinal tissue under perfluorodecaline: Determination of laser–tissue interaction in pig eyes. *Invest. Ophthalmol. Vis. Sci.*, 41:505–512 (2000).
63. Mackanos MA, Jansen ED, Shaw BL, Sanghera JS, Aggarwal I, and Katzir A, Delivery of midinfrared (6 to 7-microm) laser radiation in a liquid environment using infrared-transmitting optical fibers. *J. Biomed. Opt*., 4:583–593 (2003).
64. Chance B, Liu H, Kitai T, and Zhang Y. Effects of solutes on optical properties of biological materials: Models, cells, and tissues. *Anal. Biochem*., 227:351–362 (1995).
65. Liu H, Beauvoit B, Kimura M, and Chance B. Dependence of tissue optical properties on solute-induced changes in refractive index and osmolarity. *J. Biomed. Opt*., 1:200–211 (1996).
66. Tuchin VV, Xu X, and Wang RK. Dynamic optical coherence tomography in studies of optical clearing, sedimentation, and aggregation of immersed blood. *Appl. Opt*., 41:258–271 (2002).
67. Wang X, Milner TE, Change MC, and Nelson JS. Group refractive index measurement of dry and hydrated type I collagen films using optical low-coherence reflectometry. *J. Biomed. Opt*., 1:212–216 (1996).
68. Vitkin IJW, Wilson B, and Anderson RR. Optical and thermal characterization of natural melanin. *Photochem. Photobiol*., 59:455–462 (1994).
69. Bolin FP, Preuss LE, Taylor RC, and Ference RJ. Refractive index of some mammalian tissues using a fiber optic cladding method. *Appl. Opt*., 28:2297–2302 (1989).
70. Tuchin VV, Maksimova IL, Zimnyakov DA, Kon IL, Mavlutov AH, and Mishin AA. Light propagation in tissues with controlled optical properties, *J. Biomed. Opt*., 2:401–417 (1997).
71. Vargas G, Chan EK, Barton JK, Rylander HG, and Welch AJ. Use of an agent to reduce scattering in skin. *Lasers Surg. Med.*, 24:133–141 (1999).
72. Xu X and Wang RK. The role of water desorption on optical clearing biotissue: Studied with near infrared reflectance spectroscopy. *Med. Phys*., 30:1246–1253 (2003).
73. Choi B, Tsu L, Ishak TS, Iskandar SM, Chess S, and Nelson JS. Determination of chemical agent optical clearing potential using in vitro human skin. *Lasers Surg. Med*., 36:72–75(2005).
74. Yeh AT, Choi B, Nelson JS, and Tromberg BJ. Reversible dissociation of collagen in tissues. *J. Invest. Derm*., 121:1332–1335 (2003).
75. Yeh AT and Hirshburg J. Molecular interactions of exogenous chemical agents with collagen – implications for tissue optical clearing. *J. Biomed. Opt*., 11:014003 (2006).
76. Hirshburg J, Choi B, Nelson JS, and Yeh AT. Collagen solubility correlates with skin optical clearing. *J. Biomed. Opt*., 11:040501 (2006).
77. Hirshburg J, Choi B, Nelson JS, and Yeh A. Correlation between collagen solubility and skin optical clearing using sugars. *Lasers Surg. Med*., 39:140–144 (2007).
78. Xu X and Wang RK. Synergistic effect of hyperosmotic agents of dimethyl sulfoxide and glycerol on optical clearing of gastric tissue studied with near infrared spectroscopy. *Phys. Med. Biol.*, 49:457–468 (2004).
79. Jiang J and Wang RK. Comparing the synergistic effects of oleic acid and dimethyl sulfoxide as vehicles for optical clearing of skin tissue in vitro. *Phys. Med. Biol*., 49:5283–5294 (2004).
80. Stumpp O, Welch AJ, Gill HS, and Prausnitz MR. OCT analysis of microneedle and Er:YAG surface ablation for enhanced transepidermal delivery of hyper-osmotic agents for optical skin clearing. *Proc. SPIE*, 5319:121–129 (2004).
81. Stumpp O, Welch AJ, Milner TE, and Neev J. Enhancement of transepidermal skin clearing agent delivery using a 980 nm diode laser. *Lasers Surg. Med*., 37:278–285 (2005).
82. Tuchin VV, Altshuler GB, Gavrilova AA, Pravdin AB, Tabatadze D, Childs J, and Yaroslavsky IV. Optical clearing of skin using flash lamp-induced enhancement of epidermal permeability. *Lasers Surg. Med.*, 38:824–836 (2006).

83. Stumpp O, Chen B, and Welch AJ. Using sandpaper for noninvasive transepidermal optical skin clearing agent delivery. *J Biomed. Opt*., 11:041118 (2006).
84. McNichols RJ, Fox MA, Gowda A, Tuya S, Bell B, and Motamedi M. Temporary dermal scatter reduction: Quantitative assessment and implications for improved laser tattoo removal. *Lasers Surg. Med.*, 36:289–296 (2005).
85. Vargas G, Chan KF, Thomsen SL, and Welch AJ. Use of somatically active agents to alter optical properties of tissue: Effects on the detected fluorescence signal measured through skin. *Lasers Surg. Med.*, 29:213–220 (2001).
86. He Y and Wang RK. Improvement of low-level light imaging performance using optical clearing method. *Biosens. Bioelectron*., 20:460–467 (2004).
87. Jansen ED, Pickett PM, Machanos MA, and Virostko J. Effect of optical tissue clearing on spatial resolution and sensitivity of bioluminescence imaging. *J. Biomed. Opt*., 11:021119 (2006).
88. Plotnikov S, Juneja V, Isaacson, Mohler WA, and Campahnola PJ. Optical clearing for improved contrast in second harmonic generation imaging of skeletal muscle. *Biophys J*., 90:328–339 (2006).
89. Wang RK, Xu X, Tuchin VV, and Elder JB. Concurrent enhancement of imaging depth and contrast for optical coherence tomography by hyperosmotic agents, *J. Opt. Soc. Am. B*, 18:948–953 (2001).
90. He Y and Wang RK. Dynamic optical clearing effect of tissue impregnated with hyperosmotic agents and studied with optical coherence tomography. *J. Biomed. Opt*., 9:200–206 (2004).
91. Khan MH, Chess S, Choi B, Kelly KM, and Nelson JS. Can topically applied optical clearing agents increase the epidermal damage threshold and enhance therapeutic efficacy? *Lasers Surg. Med*., 35:93–95 (2004).
92. Vargas G, Barton JK, and Welch AJ. The use of hyper-osmotic chemical agent to improve the laser treatment of cutaneous vascular lesions. *J. Biomed. Opt*., 13:021114 (2008).
93. Vargas G, Redinger A, Dozier SS, and Welch AJ. Morphological changes in blood vessels produced by hyperosmotic agents and measured by optical coherence tomography. *Photochem. Photobiol*., 77:541–549 (2003).
94. Tuchin VV, Xu X, and Wang RK. Dynamic optical coherence tomography in studies of optical clearing, sedimentation, and aggregation of immersed blood. *Appl. Opt.*, 41:258–271 (2002).
95. Xu X, Wang RK, Elder JB, and Tuchin VV. Effect of dextran-induced changes and aggregation on optical properties of whole blood. *Phys. Med. Biol*., 48:1205–1221 (2003).
96. Cheng H, Luo Q, Zeng S, Chen S, Luo W, and Gong H. Hyperosmostic chemical agent's effect on in vivo cerebral blood flow revealed by laser speckle. *Appl. Opt*., 43:5772–5777 (2004).

Part II
Thermal Interactions

Chapter 10
Laser Generated Heat Transfer

Kenneth R. Diller

10.1 Background and Rationale

The irradiation of tissue by laser light results in the absorption of energy. Since this is a fully dissipative process, the consequence is that the increment in energy is expressed entirely as a heat transfer absorbed by the tissue. In conjunction with this absorption, there will be an increase in the energy stored locally in the tissue as a function of the geometric pattern of absorption. Two primary mechanisms of energy storage are encountered most frequently during laser irradiation: sensible and latent. Sensible storage results in a change in temperature and latent in a change in phase. The two mechanisms may occur simultaneously or singularly, depending on the initial state of the tissue and the intensity of the irradiation. A local increase in the temperature will cause a diffusion of heat to surrounding areas that are cooler. Therefore, the analysis of heat transfer is an important and relevant component of understanding the process and consequences of laser irradiation of tissue. Phase change and temperature elevation are often a direct source of injury to tissue.

The basic principles of heat transfer constitute a core component of the curriculum for many fields of engineering. The underlying theory and applications are very well known and will be discussed here only in introductory format. Many excellent texts and handbooks are available which provide clear and more comprehensive expositions [1–4]. These sources provide a useful framework in which to understand the thermal aspects of laser irradiation. However, the application of this technology to living tissues tends to produce a number of nonlinear effects that are not typically encountered in inanimate systems. Such effects present unique challenges in modeling the energetic processes and in solving the resulting partial differential equations. This chapter will present an introductory background of heat transfer as it relates to laser irradiation and discuss some of the special effects to be accounted for in

K.R. Diller (✉)
Department of Biomedical Engineering, The University of Texas, Austin, TX, USA
e-mail: kdiller@mail.utexas.edu

A.J. Welch, M.J.C. van Gemert (eds.), *Optical-Thermal Response of Laser-Irradiated Tissue*, 2nd ed., DOI 10.1007/978-90-481-8831-4_10,

the irradiation of living tissues. A capability for thermal analysis is of great practical application in the study of the *Optical-Thermal Response of Laser-Irradiated Tissue*.

10.2 Heat Transfer Fundamentals

10.2.1 Energy Conservation Principles

The starting point for analyzing the thermal aspects of laser tissue interactions is to define and understand the system of interest. In the most general sense, a system is identified as that portion of the universe that is involved directly in a particular process. The remainder of the universe is called the environment. The system interacts with the environment across its boundary. These interactions are directly responsible for changes that occur to the state of the system. The boundary surface provides a locus at which interactions can be identified and accounted for so as to predict resulting changes that will occur to the system.

We will consider a review of basic principles of heat transfer in the context of how they relate to laser generation of heat in tissue. We begin by evaluating laser generated heat in tissue from the perspective of the conservation of energy applied to a tissue system of interest. For this purpose a general system is defined with the skin surface exposed to the atmosphere and laser irradiation introduced to the interior of the tissue via an unspecified mechanism. The system is open so that mass as well as energy can cross the boundary with the environment. The system, along with relevant properties and environmental interactions, is illustrated in Fig. 10.1. For most laser tissue interactions the heating effect does not penetrate to the opposing surface of the affected volume, resulting in a semi-infinite geometric configuration. The specific geometry is important when prescribing the boundary conditions requisite to solving the differential equation that describes the temperature distribution.

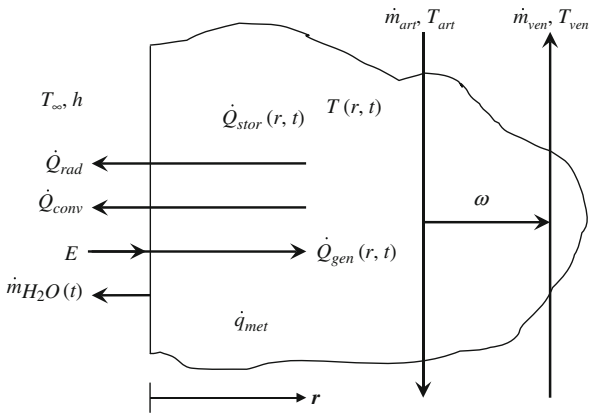


Fig. 10.1 Sources of energy interactions in a tissue system during laser irradiation

The state of the tissue system is characterized by the energy, E, which may vary with both position, $\boldsymbol{r}$, and time, t. The energy is a direct function of the temperature, T, but also depends on the phase of the system, which may variously be solid, liquid or vapor. It also may be necessary to consider more subtle phase changes associated with protein denaturation caused by heating and lipid alterations between gel and lamellar structures. The portion of laser irradiance, E [W/m^2], that is absorbed within tissue creates heat, $\dot{Q}_{\text{gen}}$ [W/m^3], that is treated as a distributed energy source which is often highly transient. $\dot{Q}_{\text{gen}}$ is equivalent to the heat source, S, of Chapter 3. Other environmental interactions that can affect the stored energy include radiation and convection from the skin surface, the loss of vapor phase water from the skin, and convection with blood that is perfused though the vascular network from arterial, $\dot{m}_{\text{art}}$ [kg/s], and venous, $\dot{m}_{\text{ven}}$ [kg/s], sources. This network typically has a very specific geometry that is unique to a tissue or organ and that can affect significantly the capability to exchange heat with the tissue in which it is embedded [5–8]. The rate of blood perfusion between arteries and veins per unit tissue mass is indicated by the symbol ω [ml blood/ml tissues]. The thermal effects of blood perfusion through tissue is discussed in greater detail later in this chapter. The loss of water, $\dot{m}_{\text{H}_2\text{O}}$ [kg/s], involves an energy component associated with phase change from liquid to vapor involving the latent heat of vaporization. Energy is generated internally in the tissue volume by metabolism, $\dot{Q}_{\text{gen}}$ [W/m^3]. Thermal boundary interactions occur over the surface area with the environment and are often characterized as convective, $\dot{Q}_{\text{conv}}$ [W/m^2], and radiative, $\dot{Q}_{\text{rad}}$ [W/m^2], processes.

The laser irradiation process causes the stored energy to be increased from its initial state. A resulting increase in temperature drives the diffusion of heat away from the irradiated area in proportion to temperature gradients that develop in the tissue. Quantitative characterization of the formation of these gradients and the heat flow that they drive comprise the focus of heat transfer analysis. The transient temperature field can be applied to predict the kinetics and accrual of tissue damage, as presented in Chapter 13.

The conservation of energy describes the change in stored energy of a system as a function of the net energy interactions experienced with the environment. For an internal subsystem of elemental dimensions dx, dy, and dz in which there is no work exchange with the environment and through which heat is being conducted, conservation of energy is written as follows.

$$\dot{Q}_{\text{stor}} = \dot{Q}_{\text{cond,in}} - \dot{Q}_{\text{cond,out}} + \dot{Q}_{\text{gen}} + \omega\rho_b c_b\,(T_a - T) \tag{10.1}$$

where ρ_b[kg/m^3] is the density of blood and c_b[J/kg · K] the specific heat of blood. The conductive heat fluxes, $\dot{Q}$, occur internally due to the temperature gradient, as will be described later. The convective and radiative fluxes shown in Fig. 10.1 at the surface are incorporated into the boundary conditions applied to the diffusion equation.

Changes in the energy stored within a tissue can occur as a consequence of several different types of phenomena. Alterations in the temperature are the primary source of energy change, denoting the sensible energy of a system. However, other

sources may also be important under special conditions. Laser heating can produce temperature increases large enough to cause a liquid to boil into a vapor, causing a change in the latent energy of the system. Latent energy can be absorbed or released while the temperature remains constant. Liquid water may undergo boiling internal to the tissue, although it is ultimately removed from the surface. Although water is a primary chemical component of tissue, other molecules can also undergo thermally driven changes in their configuration that will alter the system energy storage. A common outcome of laser heating is the denaturation of tissue proteins, which is likewise a latent energy driven process. There is a dearth of constitutive data for denaturation latent energies in comparison to the very well defined latent energy of phase change for water. Appropriate constitutive relations are needed to represent each term of the energy storage term in Eq. (10.1). If phase change effects are neglected, then the storage term is governed by sensible energy.

$$\dot{Q}_{\text{stor}} = \rho c_p \frac{\partial T}{\partial t} dx \cdot dy \cdot dz \tag{10.2}$$

There are alternate approaches to accounting for phase change effects [9, 10] in the conservation of energy equation. The latent effects generally render the differential equations much more challenging to solve since the locus of the phase change process can migrate progressively within the tissue, giving rise to an internal moving boundary condition [11]. The thermal aspects of tissue ablation are described in Chapter 14.

10.2.2 Conduction Processes

The internal diffusion of heat occurs along temperature gradients established within a tissue. The temperature is a scalar property, and the heat flux is a vector expressed as the flow per cross sectional area normal to the direction of the applied temperature gradient. The relevant constitutive property is the thermal conductivity, k[W/m · K], which is a measure of the ability of the tissue to facilitate the thermal transport process. It is dependent on the chemical composition of the tissue and the state of the system. For this reason, changes in temperature and pressure often can cause significant alterations in the magnitude of heat diffusion that occurs within a tissue.

In order to generate a transport flow with a positive vector, it is necessary for the applied potential gradient to be negative, i.e., with a magnitude decreasing along the coordinate axis. See Fig. 10.2.

For the transport of heat, the equivalent constitutive equation is known as Fourier's law of thermal diffusion. A one dimensional heat flow is given by

$$\dot{Q} = -kA \frac{dT}{dr} \tag{10.3}$$

where $\dot{Q}$ is in units of [W], T is the local temperature within the tissue, and A [m^2] is the area normal to the temperature gradient. This expression is known as Fourier's

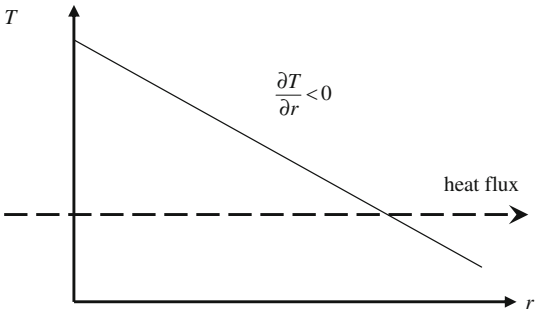


Fig. 10.2 One dimensional heat flux generated in a positive direction by application of a negative gradient in temperature along the direction of flow

law of heat conduction. The equation is written with a negative sign (–) because of requirements imposed by the second law of thermodynamics that the flow occur along a gradient from higher to lower potential.

The various transport mechanisms occur within specific materials as a function of their unique ability to facilitate these processes. In every case, the ability of a type of material to participate in a particular type of transport is related to its molecular structure as may be manifested microscopically and macroscopically. For example, solids, liquids and gases will vary in their ability to enable different types of transport processes. This ability can be quantified and measured in terms of a unique property value associated with each mechanism of transport. These features of a material are called constitutive properties, and they are used in constitutive equations, such as Fourier's law, to calculate the magnitude of a transport flow in a particular type of material subjected to a driving potential. For the diffusion of heat, the constitutive property is the thermal conductivity, k. The value of a k gives a measure of the rate at which a heat transport will occur for a given magnitude of the imposed temperature gradient.

The values for relevant transport properties are available in Chapter 12. These properties are obtained via experiments for which transport processes are caused in specific materials under very tightly controlled conditions, and for which transport models are matched to the temperature and heat flow data. The thermal force applied to a system, the resulting heat flow, and the overall shape geometry are measured as accurately as possible during the experiments. When these measured values are applied in the appropriate constitutive equation, the only unknown is the constitutive property, which is then calculated from the experimental values. This approach to analysis is termed the inverse solution method since it contrasts with the common application of a constitutive equation to calculate a flow from known values of the applied force and the transport property [12]. The values of many transport properties will change with the state of the material, and in particular with the temperature, pressure and water concentration. Therefore, it is important when designing and running an experiment, to measure a transport property to control for the variables, such as temperature and pressure, that define the state of the system. Likewise, when searching for and adopting transport properties to apply for calculations about a process of interest, it is important to ensure that the property values applied are valid

under the state conditions of the transport process of interest. Often during laser irradiation a transport process may occur over a range of temperatures. If available data for the properties indicates that there is a significant variation because of the state property gradient in the system, then it is necessary to either use an average approximation for the magnitude of the property throughout the system or to integrate the local value of the property over the spatial dimensions of the system.

Equation (10.3) may be applied to the diffusion terms in the conservation of energy equation (Eq. (10.1)) to obtain a partial differential equation for the temporal and spatial variations in temperature. This analysis is easiest to follow if we temporarily discount all effects excepting heat diffusion and sensible energy storage. Terms such as a laser induced heating source, convective blood perfusion, metabolism, and others as relevant can be added individually to the resulting differential equation by using the appropriate constitutive expressions.

A microscopic system of dimensions dx, dy, dz is defined in the interior of the tissue as shown in Fig. 10.3. The various properties and boundary flows illustrated represent the individual terms to be accounted for in applying conservation of energy to this elemental volume. The rate of change of energy storage is written as Eq. (10.2).

The heat conduction terms on opposing faces of the volume are expressed via Taylor series expansions, with terms higher than first order neglected. The resulting representation of the conservation of energy law becomes

$$\rho c_p \frac{\partial T}{\partial t} dx \cdot dy \cdot dz = -\frac{\partial \dot{Q}_x}{\partial x} dx - \frac{\partial \dot{Q}_y}{\partial y} dy - \frac{\partial \dot{Q}_z}{\partial z} dz \tag{10.4}$$

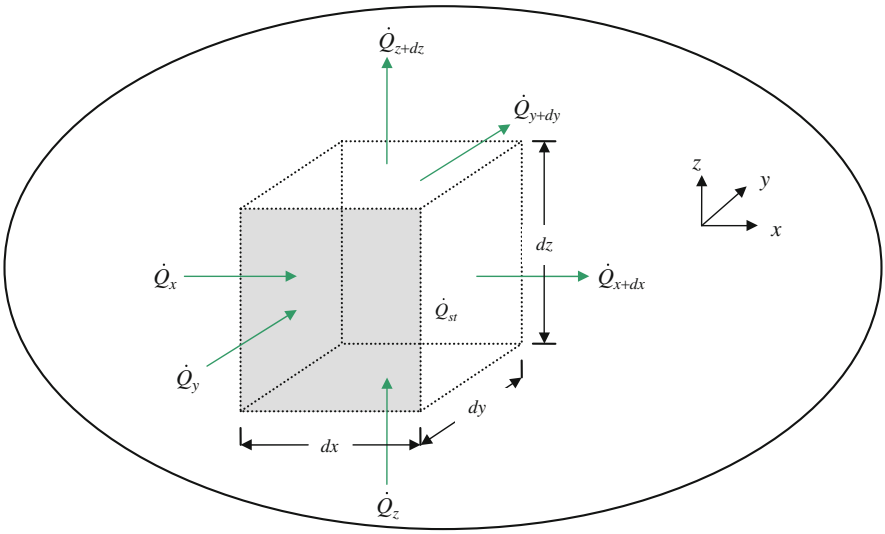


Fig. 10.3 A system consisting of an elemental volume in the interior of a tissue for which changes in stored energy are affected only by local heat diffusion

The individual conduction exchanges across the boundary are described in terms of Fourier's law.

$$\begin{aligned}\dot{Q}_x &= -k \cdot dy \cdot dz \frac{\partial T}{\partial x} \\ \dot{Q}_y &= -k \cdot dx \cdot dz \frac{\partial T}{\partial y} \\ \dot{Q}_z &= -k \cdot dx \cdot dy \frac{\partial T}{\partial z}\end{aligned} \tag{10.5}$$

These terms are substituted into Eq. (10.4), noting that each resulting term contains the system volume, $dx \cdot dy \cdot dz$, which can be divided away.

$$\rho c_p \frac{\partial T}{\partial t} = \frac{\partial}{\partial x}\left(k \frac{\partial T}{\partial x}\right) + \frac{\partial}{\partial y}\left(k \frac{\partial T}{\partial y}\right) + \frac{\partial}{\partial z}\left(k \frac{\partial T}{\partial z}\right) \ [\mathrm{W/m^3}] \tag{10.6}$$

Equation (10.6) is known as the Fourier equation (as distinguished from Fourier's law) or the heat equation in Cartesian coordinates. It can be generalized in expression to be applicable for alternate coordinate systems.

$$\rho c_p \frac{\partial T}{\partial t} = \nabla (k \nabla T) \ [\mathrm{W/m^3}] \tag{10.7}$$

Quite often the heat equation is applied to obtain a solution for the temperature as it varies with space and/or time in a system during a specific process. For this purpose, the foregoing equations may be divided by the product ρc_p to isolate the temperature term on the left side. The resulting thermal property is the thermal diffusivity, $\alpha = k/\rho c$ [$\mathrm{m^2/s}$] (Eq. (10.7)) becomes

$$\frac{\partial T}{\partial t} = \nabla (\alpha \nabla T) \ [\mathrm{K/s}] \tag{10.8}$$

The complete solution of Eq. (10.7) requires the specification of one (initial) boundary condition in time and two spatial boundary conditions for each coordinate along which the temperature may vary independently. These boundary conditions are used to evaluate the constants of integration that result from solution of the partial differential equation. They are determined according to: (a) the geometric shape of the system, including whether there is a composite structure with component volumes having distinct material properties; (b) what the temperature field interior to the system is like at the beginning of the process; (c) the geometry of imposed heat transfer interactions with the environment, such as irradiation from a laser source; and (d) how these environmental interactions may change over time. As an aggregate, these four types of conditions dictate the form and complexity of the mathematical solution to Eq. (10.8), and there are many different outcomes that may be encountered. Mathematical methods for solving this equation have been available

for many years, and some of the most comprehensive and still useful texts are considered classics in the field [13, 14]. The solution can be particularly complicated as terms to account for a laser energy source, blood perfusion, etc. are included in the governing partial differential equation. Governing equations that are intractable to

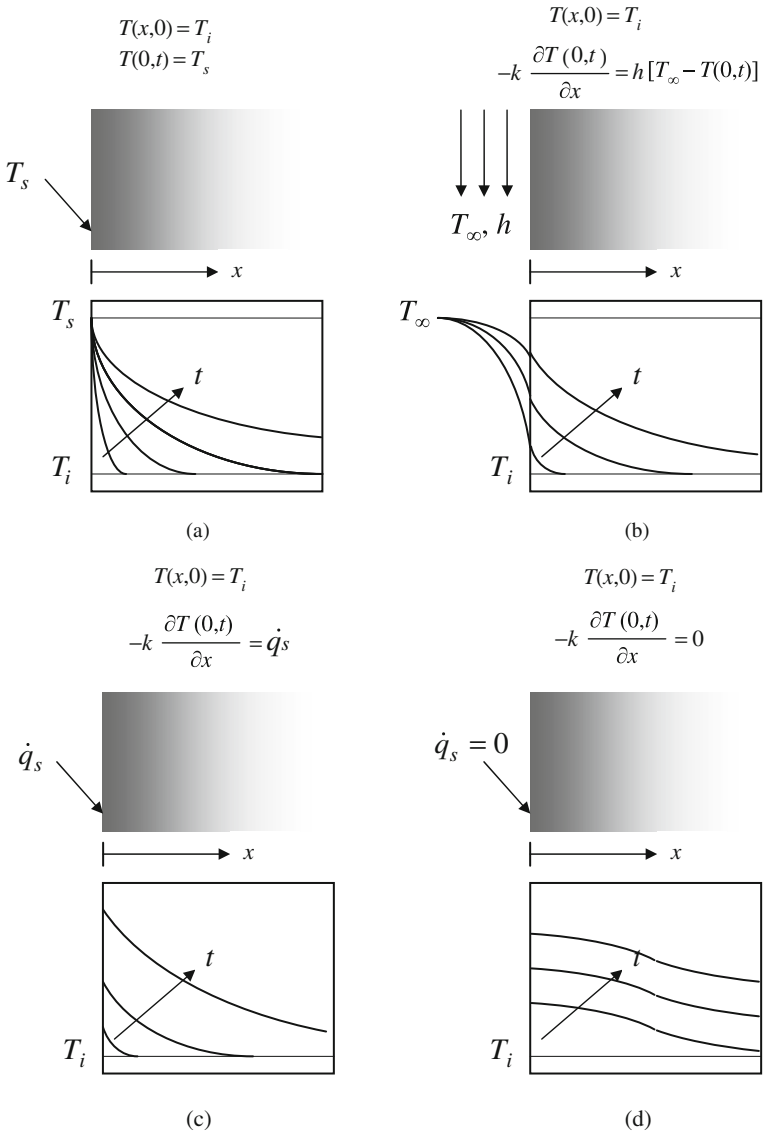


Fig. 10.4 Sample transient temperature distributions for various surface conditions that control the thermal interaction with the environment: (**a**) constant surface temperature, (**b**) convection with a fluid, (**c**) constant surface heat flux, and (**d**) insulation (zero heat flux). The latter case requires an internal heat source to produce a transient temperature field

analytical methods of solution must be addressed by numerical methods of analysis. There is a very wide range of resources available now for this purpose, and they are well beyond the scope of this treatment. The theoretical basis of numerical analysis of heat transfer problems has been presented in a comprehensive treatise [3].

The temporal boundary condition is generally defined in terms of a known temperature distribution within the system at a specific time, usually at the beginning of a process of interest. However, definition of the spatial boundary conditions is not so straight forward. There are three primary classes of spatial boundary conditions that are encountered most frequently. The thermal interaction with the environment at the physical boundary of the system may be described in terms of a defined temperature, heat flux, or convective process. The effects of these types of boundary conditions on the transient development of an internal temperature distribution are illustrated in Fig. 10.4. Further, the discussion of these types of boundary conditions provides a platform from which the principles of convection and radiation heat transfer process may be introduced, as follows.

The simplest specification for each of these conditions is that it remains constant over time, which leads to the most simple mathematical expression for the boundary condition. However, if the actual process of interest precludes using a constant boundary condition, as may be encountered in biological applications, then a more complex specification is required that may require a numerical solution.

A constant temperature boundary is encountered when the environment has a large thermal capacity in comparison to the system of interest. As a consequence, heat can be either received from or supplied to the system without altering the temperature of the environment. Examples occur when the mass of the environment is large and has a capability for distributing energy away from the surface quickly, or an environment that is undergoing a phase change, such as boiling or freezing in which energy storage is latent rather than sensible, resulting in an isothermal process.

Description of the remaining three types of boundary conditions requires that the principles of convection and radiation heat transfer be developed, along with the thermal properties of an insulated surface. The primary focus of the rest of this chapter will be a presentation of these principles.

10.2.3 Convection Processes (Case B)

Convective boundary conditions occur when a solid substrate is in contact with a fluid at a different temperature. The fluid may be in either the liquid or vapor phase. The convective process involves relative motion between the fluid and the substrate. The magnitude of the heat exchange is described in terms of Newton's law of cooling for which the relevant constitutive property of the system is the convective heat transfer coefficient, $h\,[\mathrm{W/m^2K}]$. The primary objective of convection analysis is to determine the value of the convective coefficient, h, to apply in Newton's law of cooling, which describes the convective flow at the surface, $\dot{Q}_s[\mathrm{W}]$ in terms of h,

the interface area, A, between the fluid and solid, and the substrate surface and bulk fluid temperatures, (T_s) and (T_∞).

$$\dot{Q}_s = hA\,(T_s - T_\infty) \tag{10.9}$$

There are four distinguishing characteristics of convective flow that determine the nature and intensity of a convection heat transfer process. It is necessary to evaluate each of these characteristics to calculate the value for the convective heat transfer coefficient, h. These characteristics and the various options they may take are:

1) The source of relative motion between the fluid and solid, resulting in forced (pressure driven) or free (buoyancy driven) convection.
2) The geometry and shape of the boundary layer region of the fluid in which convection occurs, producing internal or external flow. In addition, for free convection the orientation of the fluid/solid interface in the gravitational field is important.
3) The boundary layer flow domain, being laminar or turbulent.
4) The chemical composition and thermodynamic state of the fluid in the boundary layer which dictate numerical values for the constitutive properties relevant to the convective process.

The influence of each of the four principle characteristics must be evaluated individually and collectively, and the value determined for h may vary over many orders of magnitude depending on the combined effects of the characteristics. Table 10.1 presents the range of typical values for h for various combinations of the characteristics as most commonly encountered.

When a new convective transport problem is encountered it will be necessary to address the influence of each of these four characteristics on the process. Together the combined understanding given by the four characteristics provides a specific guide for determining and applying the most appropriate analysis tool for a particular convective problem. Although it is possible to perform a comprehensive analysis of some convection processes based on a full fluid flow analysis of the boundary

Table 10.1 Ranges of typical values for h as encountered for various combinations of convective transport process characteristics

Process characteristics	Range of h (W/m.^{2}K)
Free Convection	
Vapors	3–25
Liquids	20–1000
Forced Convection – Internal/External	
Vapors	10–500
Liquids	100–15,000
Phase Change	
Between liquid and vapor	5,000–100,000

layer, the mathematical procedure is quite tedious and lengthy, and many types of convection are too complex to allow this approach. Alternatively, a very broad range of experimental studies has been conducted to measure convection as a function of the four primary characteristics. The results of these experiments are presented in terms of correlation equations written in terms of dimensionless parameters that enable application for convective processes that differ from the exact physical conditions of the experiment. The standard correlation equations are available in most introductory heat transfer texts [1, 2]. A full compilation of these relationships would be too extensive for the present presentation. Nonetheless, a brief introduction is provided as a guide to applying the information and methods in the convective analysis literature.

The relative motion between a fluid and solid may be caused by differing kinds of energy sources. Perhaps most obviously, an external force can be applied to the fluid or solid to produce the motion (which is termed forced convection). This force is most frequently a mechanical force to move the solid or a pressure gradient on the fluid. However, in the absence of an external motivational force, the heat transfer process itself will cause relative motion. Owing to the constitutive properties of fluids, the existence of a temperature gradient produces a concomitant density gradient. When the fluid is in a force field such as gravity or centrifugation, the density gradient causes internal motion within the fluid by buoyancy effects as the less dense fluid rises and the more dense fluid falls under the action of the force field. This phenomenon is called free convection since no external energy source is applied to cause the motion directly. Any time there is a temperature gradient in a fluid, there is the potential for having free convection heat transfer. As can be anticipated, the fluid flow patterns for forced and free convection are very different, and, therefore, forced and free convection produce quite disparate heat transfer effects. Also, analysis of the fluid flow characteristics in forced and free convection are unique because of the differing patterns of motion. Usually, the magnitude of forced convection effects is much larger than for free convection, as indicated in Table 10.1. Thus, although the potential for free convection will be present whenever a temperature field exists in a fluid, if there is also an imposed forced source of fluid motion, the free convection effects usually will be masked since they are much smaller, and they can be neglected.

The convection process consists of the sum of two separate effects. First, when there is a temperature gradient in a fluid, heat conduction will occur consistent with the thermal conductivity of the chemical species and its thermodynamic state. The conduction effect can be very large, as in a liquid metal, or very small, as in a low density vapor. The conduction occurs via microscopic scale interactions among atoms and molecules, with no net translation of mass. Second, there will be transport of energy associated with the bulk movement of a flowing fluid. The component due to only bulk motion is referred to as advection. Convection involves a net aggregate motion of the fluid, thereby carrying the energy of the molecules from one location to another. These two effects are additive and superimposed.

Under certain circumstances a third effect can contribute to convection, that being when a material undergoes a phase change at the interface surface. In applications

involving laser irradiation of tissue this phenomenon is most often associated with evaporation or boiling of water. The very large decrease in density of the water as it is converted from the liquid to vapor state will result in a localized increase in pressure which can have a profound effect on convective processes. Fluid flow patterns become highly complex, and, if the boiling occurs below the skin surface, the mechanical properties of the overlying tissue, including the elasticity and rupture yield stress, will also play an important role in a specific manifestation of convection at the surface. Needless to say, such processes are not described by generalized methods of convective heat transfer. They require further analysis that addresses the coupled thermal, fluid and mechanical properties of the system and particulars of the imposed laser irradiation process. The influence of evaporation is described in Chapter 13.

A fundamental aspect of convection heat transfer is that the processes involve both velocity and temperature boundary layers in the fluid adjacent to a solid interface. Illustrations of these boundary layers are shown in Fig. 10.5. The velocity boundary layer defines the region wherein viscous drag causes a velocity gradient as the interface is approached. The region outside the boundary layer where the viscous properties do not affect the flow pattern is called the inviscid free stream. The

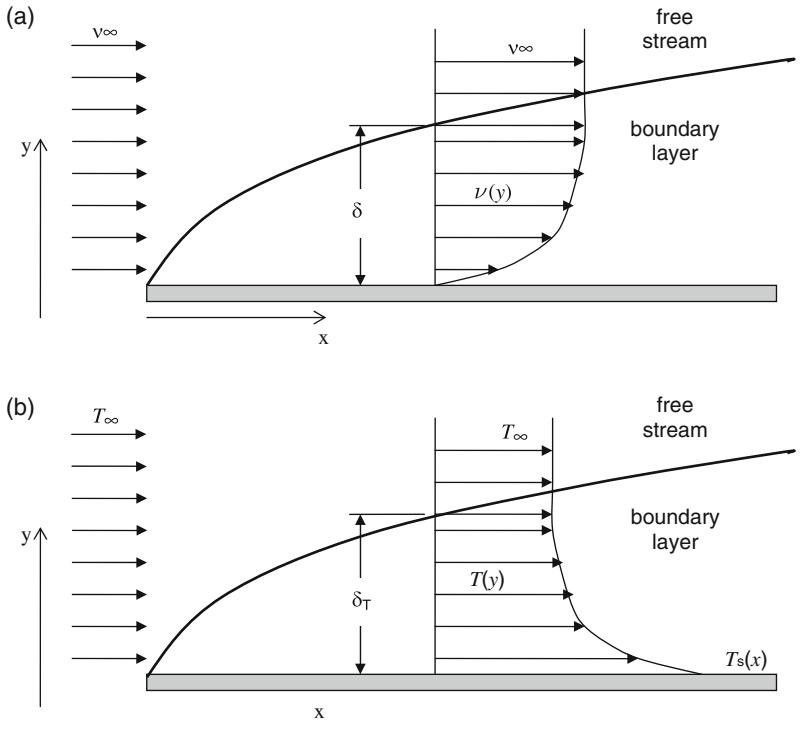


Fig. 10.5 (**a**) Velocity boundary layer (**b**) Temperature boundary layer

fluid velocity outside the boundary layer is designated by v_∞[m/s], and the boundary layer thickness by δ [m], which increases with distance along the interface from the point of initial contact between the fluid and solid. In this case, it is assumed that the interface is planar and the fluid flow is parallel to the interface. In like manner, a thermal boundary layer develops in the flowing fluid as heat transfer occurs between a solid substrate and fluid that are at dissimilar temperatures.

The velocity and temperature boundary layers have similar features which have previously been described when we discussed fluid flow. Both define a layer in the fluid adjacent to a solid in which a property gradient exists. The temperature boundary layer develops because there is a temperature difference between the fluid in the free stream T_∞ and the solid surface T_s. A temperature gradient exists between the free stream and the surface, with the maximum value at the surface and which diminishes to zero at the outer limit of the boundary layer at the free stream. The temperature gradient at the surface defines the thermal boundary condition for conduction in the solid substrate. The boundary condition can be written by applying conservation of energy at the interface. Since the interface has no thickness, it has no mass and is therefore incapable of energy storage. Thus, the conductive inflow is equal to the convective outflow as illustrated in Fig. 10.6.

An important feature of the convection interface is that there is continuity of both temperature and heat flow at the surface, the latter of which is expressed in Eq. (10.10).

$$-k_f \left.\frac{dT}{dy}\right|_{y=0} = h\,(T_s - T_\infty) \tag{10.10}$$

The magnitude of convection heat transfer is directly dependent on the size and flow characteristics within the boundary layer. As a general rule, thicker boundary layers result in a larger resistance to heat transfer and, thus, a smaller value for h. The result is that there can be local variations in convective transport over

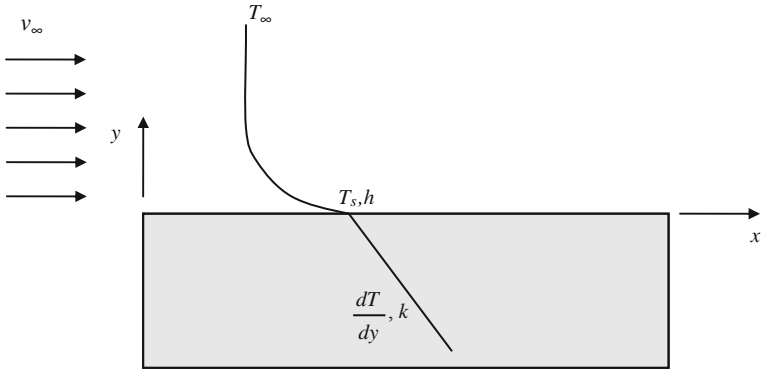


Fig. 10.6 Convective boundary condition at the surface of a conducting tissue

different regions of an interface as a function of the local boundary layer characteristics. In some cases it is necessary to determine these local variations, requiring more detailed calculations. Often it is sufficient to use a single average value over the entire interface, thereby simplifying the analysis. The averaged heat transfer coefficient is denoted by $\bar{h}_L$, where the subscript L defines the convective interface dimension over which the averaging occurs.

Values for the convective heat transfer coefficient appropriate to a given physical system are usually calculated from correlation equations written in terms of dimensionless groups of system properties. The most commonly applied dimensions groups are defined in the following paragraphs. These equations are solved for h which is then placed in Eq. (10.10).

The *Nusselt number* is a dimensionless expression for the convective heat transfer coefficient defined in Eq. (10.10). It can be written in terms of local or averaged (over an entire interface surface) values.

$$Nu_x = \frac{hx}{k_f} \text{ or } \bar{N}u_L = \frac{\bar{h}_L L}{k_f} \tag{10.11}$$

An important distinction to note is that k_f is the thermal conductivity of the fluid, not of the solid substrate. The Nusselt number describes the ratio of total convection effects to pure thermal conduction in the fluid. In principal, its value should always be greater than 1.0 since convection is a combination of both conduction plus advection in the fluid.

Nu represents the ratio of the temperature gradient in the fluid at the interface with the solid to an overall reference temperature gradient based on a physical dimension of the system, L. This dimension has a different meaning, depending on whether the flow geometry is internal or external. For an internal flow in which the boundary layer occupies the volume normal to the interface, the dimension represents the cross-sectional size of the flow passage, such as the diameter, D. For an external flow configuration in which the size of the boundary layer can grow normal to the interface with no physical restriction, the relevant dimension is the distance along the interface from the leading edge at which the fluid initially encounters the solid substrate, such as the length, L. It is important to identify the flow geometry properly in order to use the Nusselt number to calculate a value for the convective heat transfer coefficient h using Eq. (10.11). The value h is then applied to Eq. (10.10).

The Nusselt number is generally determined for a particular convection process as a function of the interface geometry, flow properties of the fluid at the interface, and thermodynamic state of the fluid. These properties are in turn represented by dimensionless ratios defined as the Reynolds, Prandtl and Grashoff numbers, as described below.

The *Reynolds number* is defined by the dimensionless ratio

$$\mathrm{Re}_L = \frac{\rho v L}{\mu} \tag{10.12}$$

where all of the constitutive properties refer to the fluid. ρ is the density [kg/m^3], v is a representative flow velocity [m/s], and μ is the viscosity [N.s/m^2]. An appropriate physical dimension of the interface is indicated by L. The Reynolds number is a primary property applied to describe forced convection processes. It describes the ratio of the inertial and viscous forces associated with the fluid flow. For low Re values the flow is dominated by the viscous resistance resulting in a laminar boundary layer in which the movement of fluid is highly ordered. High values of Re have a much larger inertial component which produces a turbulent boundary layer. The magnitude of convective heat transfer is strongly influenced by whether the boundary layer is laminar or turbulent. The geometry of the interface and boundary layer also plays an important role in the convection process. Accordingly, the Reynolds number can be written in terms of either an effective diameter, D, for an internal flow geometry, or an interface length, L, for an external flow geometry. The transition between the laminar and turbulent regimes is defined in terms of a threshold value for Re, and is very different for internal and external flow geometries. Thus, the *transition values* for Re are given as

$$\mathrm{Re}_{\mathrm{trans,int}} = \frac{\rho \upsilon D}{\mu} = 2300 \tag{10.13}$$

$$\mathrm{Re}_{\mathrm{trans,ext}} = \frac{\rho \upsilon L}{\mu} = 5x10^5 \tag{10.14}$$

The *Prandtl number* is defined by the dimensionless ratio

$$\mathrm{Pr} = \frac{c_p \mu}{k} = \frac{\nu}{\alpha} \tag{10.15}$$

where all of the constitutive properties refer to the fluid. ν is the kinematic viscosity [m^2/s], which is the ratio of the dynamic viscosity and the density. The Prandtl number describes the ratio of momentum diffusivity to thermal diffusivity. It represents a measure of the relative effectiveness of diffusive momentum and heat transport in the velocity and thermal boundary layers. It provides an indication of the relative thickness of these two boundary layers in a convective system. As a general guideline, for the broad range of Prandtl number values that may be encountered, for vapors, $\mathrm{Pr}_v \approx \delta/\delta_T \approx 1$; for liquid hydrocarbons such as oils, $\mathrm{Pr}_{\mathrm{hc}} \approx \delta/\delta_T >> 1$; and for liquid metals, $\mathrm{Pr}_{lm} \approx \delta/\delta_T << 1$.

For forced convection processes the Nusselt number is calculated from a relation of the form

$$Nu_x = f\,(x, \mathrm{Re}_x, \mathrm{Pr}) \;\text{ or }\; \bar{N}u_L = f\,(\mathrm{Re}_L, \mathrm{Pr}) \tag{10.16}$$

In most cases these relations are based on an empirical fit of the equation to experimental data. The range of parameter values for which the correlation is valid must be observed when applying specific convection relations. Examples of these correlation equations are given later (Table 10.2).

Table 10.2 Values for thermal properties of air and water in states encountered often in biological systems. Pressure = 1 atm

Material and state	k (W/m · K)	μ (N · s/m^2)	ρ (kg/m^3)	c (kJ/kg · K)	Pr
Liquid water, 37°C	9.49×10^{-2}	6.95×10^{-4}	993	4.178	4.62
Air, 37°C	2.61×10^{-2}	1.84×10^{-5}	1.17	1.007	0.708
Air, 25°C	2.70×10^{-2}	1.89×10^{-5}	1.13	1.007	0.706

The *Grashof number* is defined by the dimensionless ratio

$$Gr_L = \frac{g\beta\,(T_s - T_\infty)\,L^3}{\nu^2} \tag{10.17}$$

where g is the acceleration of gravity [1/s^2], and β is the volumetric thermal expansion coefficient of the fluid [1/K]. β is given by the relationship

$$\beta = -\frac{1}{\rho}\left(\frac{\partial \rho}{\partial T}\right)_p \tag{10.18}$$

which, for an ideal gas becomes

$$\beta = \frac{1}{T} \tag{10.19}$$

where the temperature is given in absolute units [K]. The Grashof number is a primary property applied to describe free convection processes. It describes the ratio of the buoyant and viscous forces associated with the fluid flow and is the equivalent of the Reynolds number for free convection heat transfer. Accordingly, for free convection processes the Nusselt number is calculated from a relation of the form

$$Nu_x = f\,(x, Gr_x, \mathrm{Pr}) \;\text{ or }\; \bar{N}u_L = f\,(Gr_L, \mathrm{Pr}) \tag{10.20}$$

There are many correlation equations of the type given in Eqs. (10.16) and (10.20) for forced and free convection, respectively. Comprehensive compendia of these relations are available in dedicated books [1, 15, 16]. Only the most generally used relations are presented in this chapter, with a two fold purpose: to illustrate the format of the correlations for various convective domains, and to provide a basic set of correlations that can be applied to the solution of many frequently encountered convection problems.

10.2.3.1 Interior Forced Convection Correlations – For Flow Through a Circular Conduit of Diameter, *D*, and Length, *L*

The conduit length is assumed to be greater than the entrance region at the inlet over which the boundary layers on opposing surfaces grow until they meet at the

centerline. Downstream of this point the entire volume of the conduit is filled with boundary layer flow and is termed fully developed. Fluid properties are evaluated at a mean temperature, T_m, which is an integrated average value for fluid flowing in the boundary layer through the conduit. T_m depends on the velocity and temperature profiles within the flowing fluid, which are quite different for laminar and turbulent boundary layers. In addition, T_m will change along the conduit from the inlet to the outlet as heat is exchanged between the fluid and the wall. Overall, the temperature at which the properties are evaluated should reflect the average value for all of the fluid contained in the conduit at any given time. If v_m is the mean flow velocity over the cross sectional area, A_c, of a conduit, then the mass flow rate, $\dot{m}$ [kg/s], is given by

$$\dot{m} = \rho A_c v_m \tag{10.21}$$

The net convective heat exchange between the fluid and conduit over the entire length equals the change in enthalpy of the fluid between the inlet and outlet.

$$\dot{Q} = H_{\text{out}} - H_{\text{in}} = \dot{m}\left(h_{\text{out}} - h_{\text{in}}\right) = \dot{m}c_p\left(T_{m,\text{out}} - T_{m,\text{in}}\right) \tag{10.22}$$

At any cross section along the length of the conduit the rate of energy flow associated with movement of the fluid (which is the advection rate) is obtained by integrating across the boundary layer.

$$\dot{m}c_p T_m = \int_c \rho v c_p T dA_c \tag{10.23}$$

The velocity change with radius over the cross sectional area in the above integral is substantially different for laminar and turbulent boundary layers. Eliminating the mass flow rate between Eqs. (10.21) and (10.23) yields an expression for the mean temperature over a circular cross sectional area of outer radius r_o, for constant density and specific heat

$$T_m = \frac{2}{v_m r_o^2}\int_0^{r_o} v\left(r\right) T\left(r\right) r dr \tag{10.24}$$

The functions $v(r)$ and $T(r)$ are determined by the profiles of the velocity and temperature boundary layers specific to the flow conditions of interest. They provide a basis for determining the mean temperature for defining the state at which the fluid properties in the following convection correlation relations are evaluated. The following are some of the most commonly applied convection correlation equations with the conditions noted for which they are valid.

Conditions: Fully developed, laminar, uniform temperature of wall surface, T_s.

$$Nu_D = 3.66 \tag{10.25}$$

Conditions: Fully developed, laminar, uniform heat flux at the wall surface, $\dot{q}_s$.

$$Nu_D = 4.36 \tag{10.26}$$

Conditions: Fully developed, turbulent, $\mathrm{Re}_D \geq 10^4$, $L/D \geq 10, 0.6 \leq \mathrm{Pr} \leq 160$, $T_s > T_m$

$$Nu_D = 0.23\mathrm{Re}^{0.8}\,\mathrm{Pr}^{0.4} \tag{10.27}$$

Conditions: Fully developed, turbulent, $\mathrm{Re}_D \geq 10^4$, $L/D \geq 10, 0.6 \leq \mathrm{Pr} \leq 160$, $T_s > T_m$

$$Nu_D = 0.23\mathrm{Re}^{0.8}\,\mathrm{Pr}^{0.3} \tag{10.28}$$

Conditions: Fully developed, turbulent, $3 \times 10^3 \leq \mathrm{Re}_D \leq 5 \times 10^6$, $L/D \geq 10, 0.5 \leq \mathrm{Pr} \leq 2000$

$$Nu_D = \frac{1}{8\,(0.79\ln \mathrm{Re}_D - 1.64)^2}\,\frac{(\mathrm{Re}_D - 1000)\,\mathrm{Pr}}{1 + \frac{12.7\left(\mathrm{Pr}^{2/3} - 1\right)}{8^{1/2}(0.79\ln \mathrm{Re}_D - 1.64)}} \tag{10.29}$$

As illustrated by Eqs. (10.27), (10.28), and (10.29), in some cases alternative correlation relations are available to calculate a value for h under the same conditions.

10.2.3.2 Exterior Forced Convection Correlations

The properties of the fluid are determined for a state defined by the temperature T_f where

$$T_f = \frac{T_s + T_\infty}{2} \tag{10.30}$$

which is the average of the wall and free stream fluid temperatures. The length of the fluid/substrate interface is L.

Conditions: Local convection in laminar region for flow over a flat plate, $0.6 \leq \mathrm{Pr}$

$$Nu_x = 0.332\mathrm{Re}_x^{0.5}\,\mathrm{Pr}^{0.33} \tag{10.31}$$

Conditions: Convection averaged across laminar region, L, for flow over a flat plate, $0.6 \leq \mathrm{Pr}$

$$\bar{N}u_L = 0.664\mathrm{Re}_L^{0.5}\,\mathrm{Pr}^{0.33} \tag{10.32}$$

Conditions: Local convection in turbulent region for flow over a flat plate, $\mathrm{Re}_x \leq 10^8$, $0.6 \leq \mathrm{Pr} \leq 60$

$$Nu_x = 0.0296\mathrm{Re}_x^{0.8}\ \mathrm{Pr}^{0.33} \tag{10.33}$$

Conditions: Convection averaged across the combined laminar and turbulent regions of total length, L, for flow over a flat plate, $0.6 \leq \mathrm{Pr}$

$$\bar{N}u_L = \left(0.037\mathrm{Re}_L^{0.8} - 871\right)\ \mathrm{Pr}^{0.33} \tag{10.34}$$

Conditions: Convection averaged across the entire surface around a cylinder of diameter, D, in perpendicular flow, $0.4, \leq \mathrm{Re}_D \leq 4 \times 10^5,\ 0.7 \leq \mathrm{Pr}$

$$\bar{N}u_D = C\mathrm{Re}_D^n\ \mathrm{Pr}^{0.33} \tag{10.35}$$

where the values of C and n are functions of Re_D as given below.

Re_D	C	n
0.4–4	0.989	0.330
4–40	0.911	0.385
40–4,000	0.683	0.466
4,000–40,000	0.193	0.618
40,000–400,000	0.027	0.805

Conditions: Convection averaged across the entire surface around a sphere of diameter, D, properties based on $T_\infty, 3.5 \leq \mathrm{Re}_D \leq 7.6 \times 10^4, 0.71 \leq \mathrm{Pr} \leq 380$

$$\bar{N}u_D = 2 + \left(0.4\mathrm{Re}_D^{0.5} + 0.06\mathrm{Re}_D^{0.67}\right)\ \mathrm{Pr}^{0.4}\left(\frac{\mu}{\mu_s}\right)^{0.25} \tag{10.36}$$

Conditions: Convection due to impingement of a fluid jet from a round nozzle normal to a solid planar surface with the geometry as defined below; $2,000 \leq \mathrm{Re} \leq 400,000;\ 2 \leq L/D \leq 12;\ 0.004 \leq A_r \leq 0.04.$

$$\bar{N}u = 2\ \mathrm{Pr}^{0.42}A_r^{1/2}\frac{1 - 2.2A_r^{1/2}}{1 + 0.2\left(\mathrm{L}/\mathrm{D} - 6\right)A_r^{1/2}}\mathrm{Re}^{1/2}\left(1 + 0.005\mathrm{Re}^{0.55}\right)^{1/2} \tag{10.37}$$

where the relative nozzle area, A_r, is defined as the ratio of the nozzle exit cross-section to the surface area of the impingement surface.

$$A_r = \frac{A_{\mathrm{nozzle}}}{A_{\mathrm{imping}}} = \frac{D^2}{4r^2} \tag{10.38}$$

Correlation relations for numerous other geometries and conditions are available in various primary sources (Fig. 10.7).

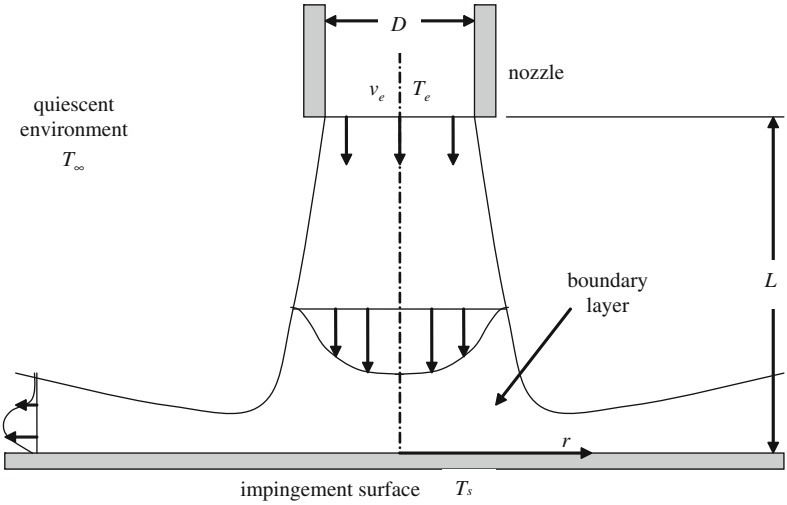


Fig. 10.7 Impingement of a circular fluid jet onto a planar surface

10.2.3.3 Free Convection Correlations

Since free convection processes are driven by buoyant effects, determination of the relevant correlation relations to determine the convection coefficient must start with analysis of the shape and orientation of the fluid/solid interface. This effect is illustrated with the free convection boundary layer adjacent to a vertical cooled flat plate shown in Fig. 10.8. Note that the flow velocity is zero at both the inner and outer extremes of the boundary layer, although the gradient is finite at the solid interface owing to viscous drag of the fluid. The environment is assumed to be quiescent so that there is no viscous shearing action at the outer region of the boundary layer. The following are some of the most commonly applied free convection correlation relations. Many make use of a dimensionless constant, the *Rayleigh number*, $Ra = Gr \cdot \mathrm{Pr}$.

Conditions: Free convection averaged over a vertical plate of length L, including both the laminar and turbulent flow regions over the entire range of Ra

$$\bar{N}u_L = \left\{ 0.825 + \frac{0.387 Ra_L^{1/6}}{\left[1 + \left(0.492/\mathrm{Pr}\right)^{9/16}\right]^{8/27}} \right\}^2 \tag{10.39}$$

Conditions: Free convection averaged over a vertical plate of length L for laminar flow defined by $10^4 \leq Ra_L \leq 10^9$.

$$\bar{N}u_L = 0.59 Ra_L^{1/4} \tag{10.40}$$

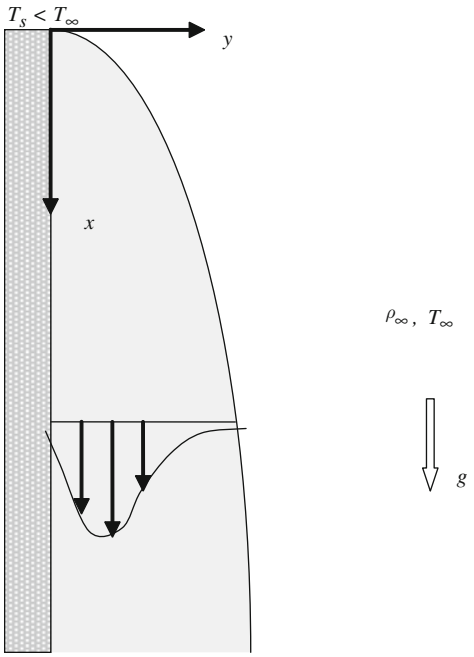


Fig. 10.8 Growth of a free convection boundary layer on a vertical cooled plate

Conditions: Free convection averaged over a vertical plate of length L for turbulent flow defined by $10^9 \leq Ra_L \leq 10^{13}$.

$$\bar{N}u_L = 0.1Ra_L^{1/3} \tag{10.41}$$

Conditions: Free convection averaged over the upper surface of a heated plate or lower surface of a cooled plate having a dimension L; $10^4 \leq Ra_L \leq 10^7$.

$$\bar{N}u_L = 0.54Ra_L^{1/4} \tag{10.42}$$

Conditions: Free convection averaged over the upper surface of a heated plate or lower surface of a cooled plate having a dimension L; $10^7 \leq Ra_L \leq 10^{11}$.

$$\bar{N}u_L = 0.15Ra_L^{1/3} \tag{10.43}$$

Conditions: Free convection averaged over the lower surface of a heated plate or upper surface of a cooled plate having a dimension L; $10^5 \leq Ra_L \leq 10^{10}$.

$$\bar{N}u_L = 0.27Ra_L^{1/4} \tag{10.44}$$

Conditions: Free convection averaged over the entire circumferential surface of a horizontal cylinder having an isothermal surface and a diameter D; $Ra_D \leq 10^{12}$.

$$\bar{N}u_D = \left\{ 0.6 + \frac{0.387Ra_D^{1/6}}{\left[1 + \left(0.559/\mathrm{Pr}\right)^{9/16}\right]^{8/27}} \right\}^2 \tag{10.45}$$

Conditions: Alternatively, free convection averaged over the entire circumferential surface of a horizontal cylinder having an isothermal surface and a diameter D.

$$\bar{N}u_D = CRa_D^n \tag{10.46}$$

where the values of C and n are functions of Ra_D as given in the table below.

Ra_D	C	n
10^{-10}–10^{-2}	0.675	0.058
10^{-2}–10^{2}	1.02	0.148
10^{2}–10^{4}	0.85	1.88
10^{4}–10^{7}	0.48	0.25
10^{7}–10^{12}	0.125	0.333

Conditions: Free convection averaged over the entire circumferential surface of a sphere having an isothermal surface and a diameter D; $Ra_D \leq 10^{11}$; $\mathrm{Pr} \geq 0.7$.

$$\bar{N}u_D = 2 + \frac{0.589Ra_D^{1/4}}{\left[1 + \left(0.469/\mathrm{Pr}\right)^{9/16}\right]^{4/9}} \tag{10.47}$$

Many other significant geometries exist for which free convection correlations exist, such as in the interior of enclosed cavities. Primary sources are readily available from which these correlations can be obtained.

In summary, the foregoing discussion of methods to obtain a value for the convective heat transfer coefficient, h, provides an introduction to the principles that govern convection processes. Although the coverage of convection conditions is far from comprehensive, the methods shown provide a guide for solving a broad range of problems that may be encountered in biomedical applications.

10.2.4 Specified Heat Flux (Radiation Processes) (Case C)

A specified heat flux at the boundary (Fig. 10.4c) is encountered most frequently as thermal radiation exchange with the environment. Alternatively, the heat flux may be fixed at the surface for conditions in which a heating or cooling element is

applied that can be independently controlled from an external source. This situation is encountered in some types of therapeutic applications.

Thermal radiation is primarily a surface phenomenon as it interacts with a conducting medium (except in transparent or translucent fluids, which will be considered at the end of this discussion), and it is to be distinguished from laser irradiation, which comes from a different type of source. Thermal radiation is important in many types of heating, cooling, and drying processes. In the out-of-doors environment solar thermal radiation can have a significant influence on the overall heat load on the skin.

Thermal radiation occurs via the propagation of electromagnetic waves. It does not require the presence of a transmitting material as do conduction and convection. Therefore, thermal radiation can proceed in the absence of matter, such as in the radiation of heat from the sun to earth. All materials are continuously emitting thermal radiation from their surfaces as a function of their temperature and radiative constitutive properties. All surfaces also are continuously receiving thermal energy from their environment. The balance between radiation lost and gained defines the net radiation heat transfer for a body. The wavelengths of thermal radiation extend across a spectrum from about 0.1 μm to 100 μm, embracing the entire visible spectrum. It is for this reason that some thermal radiation can be observed by the human eye, depending on the temperature and properties of the emitting surface.

The foregoing observations indicate that there are three properties of a body (i.e., a system) and its environment that govern the rate of radiation heat transfer: the surface temperature, the surface radiation properties, and the geometric sizes, shapes and configurations of the body surface in relation to the aggregate surfaces in the environment. Each of these three effects can be quantified and expressed in equations used to calculate the magnitude of radiation heat transfer. The objective of this presentation is to introduce and discuss how each of these three factors influences radiation processes and to show how they can be grouped into a single approach to analysis.

The first property to consider is temperature. The relationship between the temperature of a perfect radiating (black) surface and the rate at which thermal radiation is emitted is known as the *Stefan-Boltzmann law*.

$$E_b = \sigma T^4 \tag{10.48}$$

where E_b is the blackbody emissive power [W/m^2], σ is the Stefan-Boltzmann constant which has the numerical value

$$\sigma = 5.678 \times 10^{-8} \quad [\mathrm{W/m^2 \cdot k^4}]$$

Note that the temperature must be expressed in absolute units [K]. E_b is the rate at which energy is emitted diffusely (without directional bias) from a surface at temperature T [K] having perfect radiation properties. It is the summation of radiation emitted at all wavelengths from a surface. A perfect radiating surface is termed black and is characterized by emitting the maximum possible radiation at any given

temperature. The blackbody monochromatic (at a single wavelength, λ) emissive power is calculated from the Planck distribution as

$$E_{\lambda,b}(\lambda, T) = \frac{2\pi h c_o^2}{\lambda^5 \left[\exp\left(\frac{hc_o}{\lambda k T}\right) - 1\right]} \quad [\mathrm{W/m^2 \cdot \mu m}] \tag{10.49}$$

where $h = 6.636 \times 10^{-34}$[J · s] is the Planck constant, $k = 1.381 \times 10^{-23}$[J/K] is the Boltzmann constant, and $c_o = 2.998 \times 10^8$ [m/s] is the speed of light in vacuum. The Planck distribution can be plotted showing $E_{\lambda,b}$ as a function of λ for specific constant values of absolute temperature, T. The result is the nest of spectral emissive power curves in Fig. 10.9. Note that for each temperature there is an intermediate wavelength for which $E_{\lambda,T}$ has a maximum value, and this maximum increases monotonically in magnitude and occurs at shorter wavelengths with increasing temperature. Wien's displacement law, Eq. (10.50), describes the relationship between the absolute temperature and the wavelength at which maximum emission occurs.

$$\lambda_{\max} T = 2898 \quad [\mu\mathrm{m} \cdot \mathrm{K}] \tag{10.50}$$

Equation (10.49) is integrated over the entire emission spectrum to obtain the expression for the total emitted radiation, Eq. (10.48).

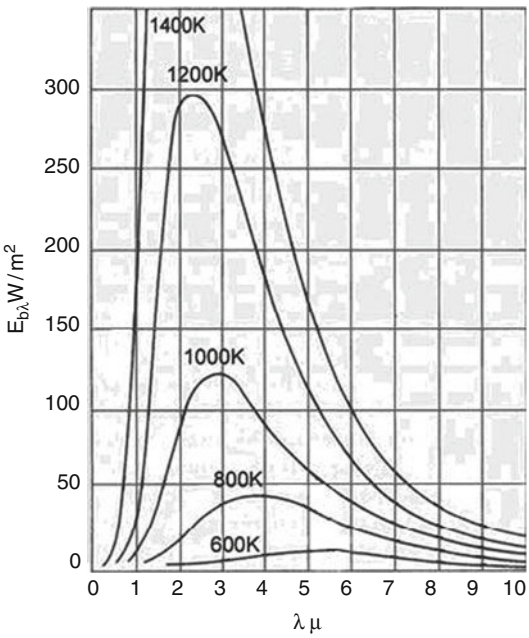


Fig. 10.9 Spectral blackbody emissive power as a function of surface temperature and wavelength

$$E_b(T) = \int_0^\infty \frac{2\pi h c_o^2}{\lambda^5 \left[\exp\left(\frac{hc_o}{\lambda k T}\right) - 1\right]} d\lambda = \sigma T^4 \tag{10.51}$$

Equation (10.51) represents the area under an isothermal curve in Fig. 10.9 depicting the maximum amount of energy that can be emitted from a surface at a specified temperature. This set of equations provides the basis for quantifying the temperature effect on thermal radiation. It applies to idealized, black surfaces.

Next we will consider the effect of real, rather than idealized, surface properties on thermal radiation exchange. Real surfaces emit less than black body radiation at a given temperature. The ratio of real to black radiation levels defines a property called the emissivity, ε. In general, radiation properties are functions of the radiation wavelength and for many practical systems can change significantly over the thermal spectrum. Thus,

$$\varepsilon_\lambda(\lambda, T) = \frac{E_\lambda(\lambda, T)}{E_{\lambda,b}(\lambda, T)} \tag{10.52}$$

An idealized real surface has radiant properties that are wavelength independent and is termed a gray surface. For these conditions,

$$\varepsilon(T) = \frac{E(T)}{E_b(T)} = \frac{E(T)}{\sigma T^4} \tag{10.53}$$

A gray surface has the effect of decreasing the magnitude of the curves in Fig. 10.9 by a constant factor over all wavelengths.

In addition to emission, surfaces are continually bombarded by thermal radiation from their environments. The net radiant flux at a surface is the difference between the energies received and lost. As with emission, the surface radiation properties play an important role in determining the amount of energy absorbed by a surface. A black surface absorbs all incident radiation, whereas real surfaces absorb only a fraction that is less than one. The total radiant flux onto a surface from all sources is called the radiosity and is denoted by the symbol G [W/m^2]. In general the incident radiation will be composed of many wavelengths, denoted by G_λ. A surface can have three modes of response to incident radiation: the radiation may be absorbed, reflected, and/or transmitted. The fractions of incident radiation that undergo each of these responses are determined by three dimensionless properties: the coefficients of absorption, α, reflection, ρ, and transmission, τ. Conservation of energy applied at a surface dictates that the relationship among these three properties must be

$$\alpha + \rho + \tau = 1 \tag{10.54}$$

Figure 10.10 illustrates these phenomena for radiation incident onto a surface that is translucent, allowing some of the radiation to pass through. All three of the properties are wavelength dependent.

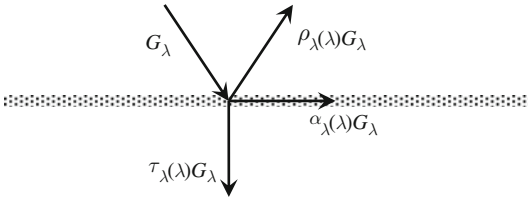


Fig. 10.10 Absorption, reflection and transmission phenomena for a surface irradiated with a multi-wavelength incident radiation, G_λ

The three properties are defined according to the fraction of radiosity that is absorbed, reflected and transmitted.

$$\alpha_\lambda(\lambda) = \frac{G_{\lambda,abs}(\lambda)}{G_\lambda(\lambda)} \tag{10.55}$$

$$\rho_\lambda(\lambda) = \frac{G_{\lambda,ref}(\lambda)}{G_\lambda(\lambda)} \tag{10.56}$$

$$\tau_\lambda(\lambda) = \frac{G_{\lambda,tr}(\lambda)}{G_\lambda(\lambda)} \tag{10.57}$$

It is well known that there is a very strong spectral (wavelength) dependence of these properties. For example, the greenhouse effect occurs because glass has a high transmissivity (τ) at relatively short wavelengths in the visible spectrum that are characteristic of the solar flux. However, the transmissivity is very small in the infrared spectrum in which terrestrial emission occurs. Therefore, heat from the sun readily passes through glass and is absorbed by interior objects. In contrast radiant energy emitted by these interior objects is reflected back to the source. The net result is a warming of the interior of a system that has a glass surface exposed to the sun. The lens of a camera designed to image terrestrial sources of thermal radiation, predominantly in the infrared spectrum, must be fabricated from a material that is transparent at those wavelengths. It is important to verify whether the spectral dependence of material surface properties is important for specific applications involving thermal radiation.

An additional important surface property relationship is defined by Kirchhoff's law, which applies for a surface which is in thermal equilibrium with its environment. Most thermal radiation analyses are performed for processes that are steady state. For the surface of a body n having a surface area A_n, at steady state the radiation gained and lost is balanced so that the net exchange is zero. To illustrate, we may consider a large isothermal enclosure at a temperature, T_s, containing numerous small bodies, each having unique properties and temperature. See Fig. 10.11.

Since the surface areas of the interior bodies are very small in comparison to the enclosure area, their influence on the radiation field is negligible. Also, the radiosity to the interior bodies is a combination of emission and reflection from the enclosure surface. The net effect is that the enclosure acts as a blackbody cavity regardless of

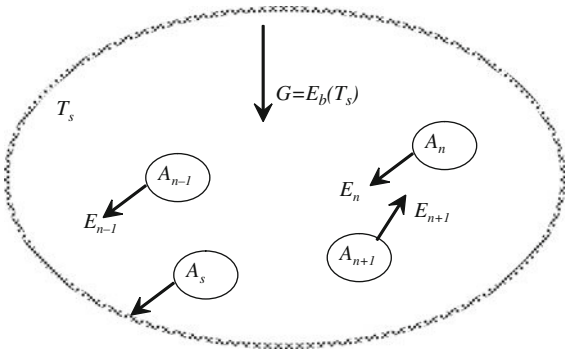


Fig. 10.11 Steady state thermal radiation within a large isothermal enclosure containing multiple small bodies

its surface properties. Therefore the radiosity within the enclosure is expressed as

$$G = E_b\,(T_s) \tag{10.58}$$

For thermal equilibrium within the cavity, the temperatures of all surfaces must be equal $T_{n-1} = T_n = T_{..} = T_s$. A steady state energy balance between absorbed and emitted radiation on one of the interior bodies yields

$$\alpha_n G A_n - E_n\,(T_s)\,(A_n) = 0 \tag{10.59}$$

The term for radiosity may be eliminated between the two foregoing equations.

$$E_b\,(T_s) = \frac{E_n\,(T_s)}{\alpha_n} \tag{10.60}$$

This relationship holds for all of the interior bodies. Comparison of Eqs. (10.53) and (10.60) shows that the emissivity and absorptivity are equal.

$$\alpha = \varepsilon, or, \alpha_\lambda = \varepsilon_\lambda \tag{10.61}$$

The general statement of this relationship is that for a gray surface, the emissivity and absorptivity are equal and independent of spectral conditions.

The third factor influencing thermal radiation transfer is the geometric sizes, shapes and configurations of body surfaces in relation to the aggregate surfaces in the environment. This effect is quantified in terms of a property called the shape factor, which is solely a function of the geometry of a system and its environment. By definition, the shape factor is determined for multiple bodies, and it is related to the size, shape, separation and orientation of the bodies. The shape factor $F_{m\to n}$ is defined between two surfaces, m and n, as the fraction of energy that leaves surface m that is incident onto the surface n. It is very important to note that the shape factor is directional. The shape factor from body m to n is probably not equal to that from body n to m.

Values for shape factors have been compiled for a broad range of combinations of size, shape, separation and orientation and are available as figures, tables and equations [17, 18]. There are a number of simple relations that govern shape factors and that are highly useful in working many types of problems. One is called the reciprocity relation.

$$A_n F_{n\to m} = A_m F_{m\to n} \tag{10.62}$$

This equation is applied for calculating the value of a second shape factor between two bodies if the first is already known.

A second relation is called the summation rule which says the sum of shape factors for the complete environment of a body equals 1.0.

$$\sum_{k=1}^{n} F_{m\to k} = 1.0 \tag{10.63}$$

The summation rule accounts for the entire environment for an object.

A limiting case is shown in Fig. 10.11 in which $A_s >> A_n$. For this geometry, the reciprocity relation dictates that $F_{s\to n}$ be vanishingly small since only a very small fraction of the radiation leaving the large surface s will be incident onto the small surface n. The summation rule then shows that effectively $F_{n\to s} = 1$.

The third geometric relationship states that the shape factors for a surface to each component of its environment are additive. If a surface n is divided into l components, then

$$F_{m\to(n)} = \sum_{j=1}^{l} F_{m\to j} \tag{10.64}$$

where

$$A_n = \sum_{j=1}^{l} A_j.$$

Equation (10.64) is useful for calculating the shape factor for complex geometries that can be subdivided into an assembly of more simple shapes.

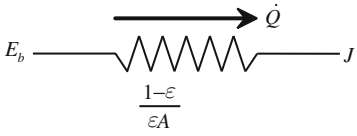


Fig. 10.12 Electrical resistance model for the drop in radiation potential due to a gray surface defined by the property ε

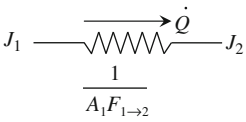


Fig. 10.13 Electrical resistance model for the radiation exchange between two surfaces with a shape factor $F_{1\to 2}$

There exist comprehensive compendia of data for determination of a wide array of shape factors [19]. The reader is directed to such sources for detailed information. Application of geometric data to radiation problems is very straight forward.

Evaluation of the temperature, surface property and geometry effects can be combined to calculate the magnitude of radiation exchange among a system of surfaces. The simplest approach is to represent the radiation process in terms of an equivalent electrical network. For this purpose, two special properties are used: the irradiation, G, which is the total radiation incident onto a surface per unit time and area, and the radiosity, J, which is the total radiation that leaves a surface per unit time and area. Also, for the present time it is assumed that all surfaces are opaque (no radiation is transmitted), and the radiation process is at steady state. Thus, there is no energy storage within any components of the radiating portion of the system.

The radiosity can be written as the sum of radiation emitted and reflected from a surface, which is expressed as

$$J = \varepsilon E_b + \rho G \tag{10.65}$$

The net energy exchanged by a surface is the difference between the radiosity and the irradiation. For a gray surface with $\alpha = \varepsilon$, and therefore $\rho = 1 - \varepsilon$,

$$\frac{\dot{Q}}{A} = \dot{q} = J - G = \varepsilon E_b + (1 - \varepsilon) G - G \tag{10.66}$$

which can also be written as

$$\dot{Q} = \frac{E_b - J}{(1 - \varepsilon)/\varepsilon A} \tag{10.67}$$

The format of Eq. (10.67) is in terms of a flow that equals a difference in potential divided by a resistance. In this case the equation represents the drop in potential from a black to a gray surface associated with a finite surface radiation resistance. The equation can be represented graphically in terms of a steady state resistance.

This resistance applies at every surface within a radiating system which has non-black radiation properties. Note that for a black surface for which $\varepsilon = 1$, the resistance goes to zero.

A second type of radiation resistance is due to the geometric shape factors among multiple radiating bodies. The apparent radiation potential of a surface is the radiosity. For the exchange of radiation between two surfaces A_1 and A_2, the net energy flow equals the sum of the flows in both directions. The radiation leaving surface 1

that is incident on surface 2 is $J_1 A_1 F_{1\to 2}$ and, in like manner, the radiation from 2 to 1 is $J_2 A_2 F_{2\to 1}$.

The net interchange between surfaces 1 and 2 is then the sum of these two flows

$$\dot{Q}_{1\to 2} = J_1 A_1 F_{1\to 2} - J_2 A_2 F_{2\to 1} = \frac{J_1 - J_2}{1/A_1 F_{1\to 2}} \tag{10.68}$$

This process can also be modeled via an electrical network.

These two types of resistance elements can be applied to model the steady state interactions among systems of radiating bodies. The simplest example is of two opaque bodies that exchange radiation only with each other. This problem is characterized by the network shown in Fig. 10.14.

This network can be solved to determine the radiation heat flow in terms of the temperatures (T_1, T_2), surface properties (ε_1, ε_2), and system geometry (A_1, A_2, $F_{1\to 2}$).

$$\dot{Q}_{1\to 2} = \frac{E_{b1} - E_{b2}}{\frac{1-\varepsilon_1}{\varepsilon_1 A_1} + \frac{1}{A_1 F_{1\to 2}} + \frac{1-\varepsilon_2}{\varepsilon_2 A_2}} = \frac{\sigma\left(T_1^4 - T_2^4\right)}{\frac{1-\varepsilon_1}{\varepsilon_1 A_1} + \frac{1}{A_1 F_{1\to 2}} + \frac{1-\varepsilon_2}{\varepsilon_2 A_2}} \tag{10.69}$$

Note that the second term in this equation has a linear differential in the driving potential whereas the third term has a fourth power differential. A major advantage of the electrical circuit analogy is that a radiation problem can be expressed as a simple linear network as compared to the thermal formulation in which temperature must be raised to the fourth power.

Given the network modeling tools, it becomes straightforward to describe radiation exchange among the components of an n bodied system. A three bodied system can be used to illustrate.

In this system each of the three bodies experiences a unique radiant heat flow. For the special case in which one surface, such as 3, is perfectly insulated, meaning that all incident radiation is reradiated, then the diagram is simplified to a combined series/parallel exchange between surfaces 1 and 2.

The radiation heat flow for this system is written as a function of the system properties, with the shape factor reciprocity relation applied for $A_1 F_{2\to 1} = A_2 F_{2\to 1}$, as

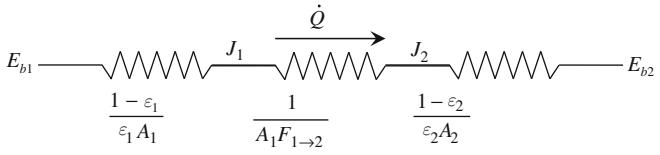


Fig. 10.14 Electrical resistance model for the radiation exchange between two surfaces with a shape factor $F_{1\to 2}$

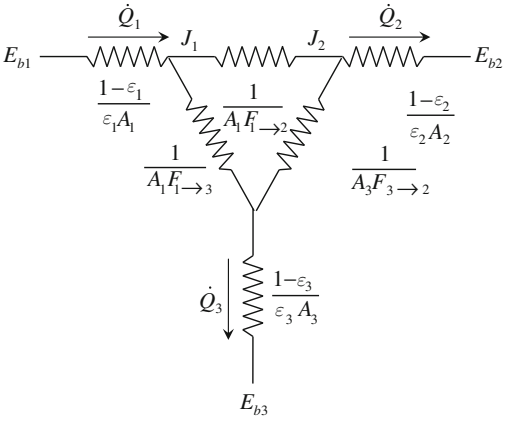


Fig. 10.15 Electrical resistance model for the radiation exchange among the surfaces of a three body system

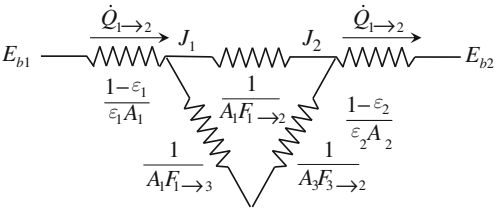


Fig. 10.16 Electrical resistance model for the radiation exchange among the surfaces of a three body system in which surface 3 is perfectly insulated

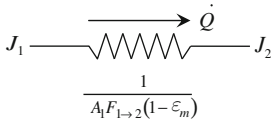


Fig. 10.17 Electrical resistance model for the effect of an absorbing and transmitting medium on the thermal radiation between two surfaces

$$\dot{Q}_{1\to 2} = \frac{\sigma A_1 \left(T_1^4 - T_2^4\right)}{\left(\frac{1}{\varepsilon_1} - 1\right) + \frac{A_1 + A_2 - 2A_1 F_{1\to 2}}{A_2 - A_1 (F_{1\to 2})^2} + \frac{A_1}{A_2}\left(\frac{1}{\varepsilon_2} - 1\right)} \tag{10.70}$$

A final case to be discussed addresses the effect of an absorbing and transmitting medium included in a radiating system. Glasses and gases are examples of this type of media. If the medium is nonreflective, then

$$\alpha_m + \tau_m = \varepsilon_m + \tau_m = 1 \tag{10.71}$$

Radiation that is absorbed by the medium is transmitted and then emitted to its environment. For a system consisting of two surfaces 1 and 2 that see only each other, plus an intervening medium m, the net energy leaving surface 1 that is

transmitted through the medium and arrives at surface 2 is $J_1 A_1 F_{1\to 2} \tau_m$. Likewise, the energy flow in the opposite direction is $J_2 A_2 F_{2\to 1} \tau_m$. The net interchange between surfaces 1 and 2 via transmission through the medium is then the sum of these two flows

$$\dot{Q}_{1\to 2} = J_1 \tau_m A_1 F_{1\to 2} - J_2 \tau_m A_2 F_{2\to 1} = \frac{J_1 - J_2}{1/A_1 F_{1\to 2}\,(1-\varepsilon_m)} \tag{10.72}$$

The effect of the medium can be represented by a radiation network element.

This circuit element can be included in a radiation network model as appropriate to represent the effect of an interstitial medium between radiating surfaces.

Note that in this analysis of radiation all of the equivalent electrical networks contain only resistors, and specifically there are no capacitors. The explicit interpretation of this arangement is that all of the radiation processes considered are at steady state such that no energy storage occurs. For our present analysis the mass of all radiating bodies has been neglected. Since many thermal radiation processes are approximated as surface phenomena, this is a reasonable assumption. Under conditions that demand more comprehensive and sophisticated analysis, this assumption may have to be relaxed, which leads to a significant increase in the complexity of the thermal radiation analysis.

Thermal radiation effects can be distinguished from optical irradiation during laser/tissue interactions. The laser energy is absorbed by the tissue as a spatially and temporally distributed heating source, $\dot{Q}_{\text{gen}}(r,t)$. Thermal radiation is emitted and absorbed from the surface of the tissue as a function of the temperature and radiative properties of the tissue and the environment [9]. These are different physical phenomena, and they likewise are distinguished mathematically when modeling the system behavior.

10.2.5 Insulated Surfaces (Case D)

A special case of a heat flux defined boundary condition occurs for an insulated surface (Fig. 10.4d). For these circumstances, because the surface flux is zero, the temperature gradient at the surface of the underlying substrate is also zero.

$$\dot{Q}_x = kA \left.\frac{dT}{dx}\right|_{x=0} = 0 \tag{10.73}$$

A common application in which this condition is applied occurs when a system has symmetry of both geometry and temperature field. At points, lines or planes of symmetry the temperature gradient is zero. There will be no heat flux crossing that symmetric locus, rendering it the property of being insulated.

10.2.6 Boundary Conditions at Material Interfaces

An important class of boundary conditions is associated with a material interface interior to a system, as illustrated in Fig. 10.18.

Physical constraints demand that two independent conditions occur at the internal material boundary, which can be interpreted directly as the two requisite boundary conditions. One constraint is that the flux be continuous at the interface. Since there is no mass associated with an interface, there is no capacity for storing energy. Conservation of energy then states that the flow into the interface must equal the flow out of the interface.

$$\dot{Q}_A = \dot{Q}_B \tag{10.74}$$

Fourier's law applies for the conduction in both materials, which are anticipated to have different values for their thermal conductivities. Equation (10.74) becomes at the interface

$$k_A \frac{\partial T_A}{\partial x} = k_B \frac{\partial T_B}{\partial x} \tag{10.75}$$

The second constraint is that the temperatures be continuous at the interface.

$$T_A = T_B \tag{10.76}$$

Note, however, that differing thermal conductivities of the interface materials means that the temperature slopes will not be continuous.

When tissue is optically irradiated, the fluence rate is continuous at a tissue interface if the two surfaces have the same index of refraction, n. However, the resulting

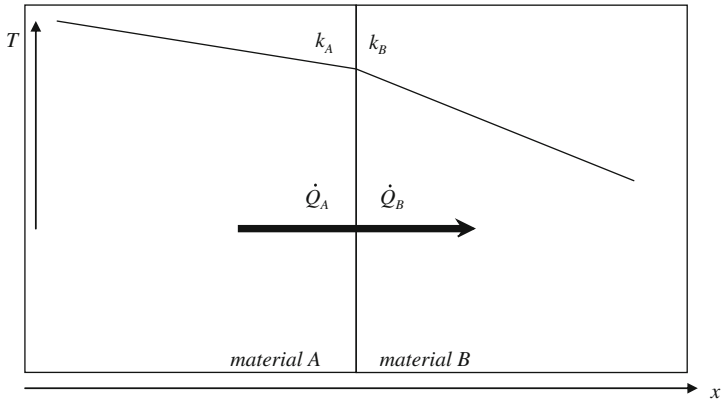


Fig. 10.18 Specification of boundary conditions at a material interface. Note that there is continuity of both potential and flow

heat source, S, will be discontinuous if the absorption coefficients of the two layers are different (see Eq. (3.55)).

10.3 Bioheat Transfer with Blood Perfusion

A unique aspect of heat transfer in biological tissues is the thermal interaction between the local tissue and blood that is perfused through it. Typically, blood coming from the body core is at a temperature different than that in peripheral regions of the body. Blood flowing through the vascular network behaves as a forced convection process. The vasculature is characterized by a large range of diameters, lengths and flow velocities as well as tortuous geometries. Also, blood perfusion within any given tissue or organ will vary over the course of a day and with the type and level of activity. For some physiological conditions arterial and venous blood temperature may be significantly different, and the flows are in a counter current geometry that provides for direct heat exchange. Correspondingly, there is large variation in the convective efficiency among different components of the vasculature. Blood perfusion driven convection is important to many physiological processes such as thermoregulation, muscle exercise, and inflammation.

The effects of blood perfusion on tissue temperature gradient may be modeled by adding specific terms to the energy balance equation. There have been many different approaches proposed to account for the internal convection with blood based on the influence of local thermal equilibration, vessel size, countercurrent flow patterns, and bleed-off perfusion to microvessels [6, 8]. Discussion of pros and cons of these methods of analysis are presented in the cited references. Most of the alternative analyses embody specific physiological geometries in conjunction with complex mathematical formulations. A discussion of their relative merits is beyond the scope of this presentation. The reader is referred to the review by Charney [6] for further details.

The seminal work on the thermal effects of blood perfusion was published in 1948 by Pennes [5], based on the assumption of equilibration between arterial blood that is diverted through the local microvasculature. This effect, along with that of metabolism, are embodied in a modified version of the conduction equation often termed the Pennes bioheat equation.

$$\rho c_p \frac{\partial T}{\partial t} = \nabla (k \nabla T) + \varpi \rho_b c_b (T_a - T) + \dot{Q}_{\text{met}} \tag{10.77}$$

Although there has been extensive discussion in the literature concerning the validity of the Pennes equation, it has had a demonstrated efficacy for describing blood perfusion convection effects in many types of application [7]. It is thought to be especially appropriate to apply to model blood flow effects in the skin. A major advantage of the Pennes equation is a simple mathematical formulation in which the term that accounts for perfusion heat transfer is linear in temperature.

Chato [20] and Chen and Holmes [21] published nearly simultaneously analyses of perfusion heat transfer that provide important insight in the blood convection process in tissues. They modeled the convection process as blood flows through the vascular arcades to identify the point in the vasculature at which the temperature of arterial blood reaches that of the surrounding tissue. This analysis demonstrates that thermal equilibration occurs in the terminal arterioles, which have a diameter of approximately 60 μm. Thus, the blood is essentially equilibrated prior to entering the capillaries. Although mass exchange between blood and tissue occurs primarily in the capillaries, the concept that heat exchange occurs there also is not correct. The implication of this analysis for the Pennes equation is that the diffuse microvascular blood perfusion must initiate at the level of the arterioles and venules from the perspective of convective heat transfer.

The control of blood perfusion is quite sensitive to changes in local and core temperatures. Variation in perfusion is a key component in the functioning of the human thermoregulatory system. A key operational goal of human thermoregulation is to maintain the body core within a narrow range of temperatures. Thus, when the core temperature begins to rise above a threshold, the peripheral vasculature dilates to increase blood flow to the skin where heat convected from the core can be exchanged to the environment. In contrast, when the core temperature is falling, the peripheral vasculature constricts so as to impede the loss of heat from the core to the environment and thereby conserve the temperature of the body core. These vasodilatation and vasoconstriction processes typically occur with time constants on the order of a minute or more, which is considerably longer than the duration of most laser irradiation processes followed by the thermal relaxation. Therefore, vasoactive regulation does not directly affect laser irradiation of tissue.

In contrast, local changes in blood perfusion in response to changes in local tissue temperature can occur very quickly, as fast as one second. When temperatures rise to the levels at which injury can occur, approximately in the mid-40s C, blood flow will stop immediately. This change will affect the distribution of heat in tissue following an irradiation. The loss of blood flow is accommodated in thermal models by dropping the perfusion term from the Pennes equation (Eq. (10.77)). The time constant of this response is within the duration of most laser irradiation and the subsequent thermal relaxation process and should be accounted for in process models.

10.4 Special Analyses of Bioheat Transfer Relating to Laser Irradiation of Tissue

The foregoing discussion of bioheat transfer has presented a brief and general overview of the topic. In the context of the present book it is appropriate to address some more narrowly focused topics on the thermal effects of laser irradiation of tissue.

10.4.1 Spray Cooling of the Skin Surface During Laser Irradiation

A frequent objective of laser irradiation procedures is to deposit thermal energy (heat) into tissue within a targeted subsurface volume. The absorption characteristics of the tissue for the applied wavelength may be such that in order to achieve the necessary temperature elevation in the interior volume, an unacceptable temperature rise occurs in the surface tissue resulting in unwanted damage to that tissue. An option for overcoming this limitation is to provide simultaneous or pre-cooling of the surface in conjunction with the irradiation to prevent the surface tissue temperature from rising to injurious levels [22–24]. This procedure is based on a process known as spray cooling or impingement jet heat transfer in which a stream of cool fluid is directed obliquely onto a substrate in order to produce a local convective cooling effect [25–28]. Jet impingement convection is illustrated in Fig. 10.7. The magnitude of the convection cooling is determined by the velocity, direction and pattern of spray, the thermal properties of the spray fluid, and the temperature of the spray. A correlation equation for conditions of a simplified geometry is given in Eq. (10.37) [26]. Surface cooling by jets of combined liquid and vapor phase has also been evaluated [25], but is not included in this discussion. This process has a long history of many industrial applications, and its adaptation to control the spatial temperature distribution in tissue during laser irradiation is a novel application that has issued in significantly improved outcomes for numerous laser treatments. Thermal analysis of the process may be additionally complicated if the cooling fluid is a liquid cryogen, in which case the liquid evaporates upon contact, issuing in a joint sensible and latent heat transfer interaction with the skin and a concomitant two phase external flow phenomenon.

Holman and colleagues [27, 28] have investigated spray cooling of vertical surfaces by a horizontal stream of subcooled Freon-113. The subcooling has the effect of reducing the complexity of boiling from the impingement heat transfer process, making the analysis and correlation considerably simpler. Nonetheless, unless the extent of subcooling is very large, phase change effects will persist at the cooled surface. The primary motivation for this work was the cooling of electronic equipment, which also has a spatially distributed pattern of internal energy generation. Pressurized spray from a nozzle produced a stream of liquid droplets striking the heated surface. From a mechanistic perspective a large number of physical characteristics of the spray system have been identified that exert a direct influence on the ability to cool a surface. These include the mass flux, spray droplet velocity, droplet diameter, and the distance between the spray nozzle and surface. Some characteristics are that the cooling effect is more effective with: increasing droplet mass flux, degree of subcooling of the cryogen, and magnitude of the Weber number, *We*, which is a dimensionless parameter defined in terms of the ratio of the inertial and surface tension properties of the spray.

$$We = \frac{\rho \upsilon^2 d_d}{\sigma} \tag{10.78}$$

where d_d is the diameter of the spray droplets, and σ is the surface tension between the liquid and vapor phases of the spray fluid. The Weber number characterizes the impact dynamics of the spray droplets with the surface. A large value of *We* describes physical conditions for which the formation of a vapor layer between the droplets and the surface will be minimized, thereby leading to more efficient cooling of the surface. Experimental observations have shown that the mechanism of heat transfer between the spray droplets and the solid surface is by subcooled boiling which is a combination of convection and liquid – vapor phase change. Extensive experimental data for spray cooling with Freon-113 led to a correlation equation relating the cooling heat flux, $\dot{q}(\mathrm{W}\cdot\mathrm{m}^{-2})$, to: *We*, the flow velocity of the spray, the temperature difference between the spray liquid and the surface, ΔT, and the magnitude of the mass flux of spray, which is linearly related to the separation of the spray nozzle and warm surface, x. For most spray nozzle configurations, the cross sectional area of the spray pattern will increase in size with distance from the source, thereby reducing the magnitude of the flux $(\mathrm{kg}\cdot\mathrm{s}^{-1}\cdot\mathrm{m}^{-2})$. The degree of spray liquid subcooling is embodied in the term ΔT, which is the subcooling of the liquid below the saturation temperature, T_{sat}, plus the differential between the saturation and warm surface temperatures, if indeed a coolant is used which boils at a temperature below that of the surface being cooled. With adequate nucleation conditions present, the spray liquid will begin to boil when it is warmed to the saturation temperature. These relationships are expressed in terms of a correlation equation written with dimensionless parameters [27, 28].

$$\frac{\dot{q}x}{\mu_f h_{fg}} = 9.5We^{0.6}\left(\frac{c_p\Delta T}{h_{fg}}\right)^{1.5} \tag{10.79}$$

The thermodynamic properties of the spray are evaluated at the fluid temperature, T_f, Eq. (10.30).

The temperature difference between the spray liquid and the heated surface is normalized to the latent heat of vaporization between the liquid and vapor states, h_{fg}, with the liquid specific heat, c_p. This ratio compares the capability of the spray liquid to undergo sensible and latent heat transfers as it interacts with the warm surface, which may be referred to as the Stefan number, *Ste*. The Reynolds number may be defined in terms of the droplet diameter, and the droplet diameter normalized to the nozzle to surface distance. These dimensionless parameters are given by the relations

$$Ste = \frac{c_p\Delta T}{h_{fg}} \tag{10.80}$$

$$\mathrm{Re}_{d_d} = \frac{\rho\upsilon d_d}{\mu_f} \tag{10.81}$$

$$d_d^* = \frac{d_d}{x} \tag{10.82}$$

so that the spray cooling heat flux can be written as

$$\frac{\dot{q}}{\rho \upsilon h_{fg}} = 9.5 \frac{We^{0.6}}{\mathrm{Re}} d_d^* Ste^{1.5} \tag{10.83}$$

The heat flux at the surface is normalized to the energy of the approaching spray stream in terms of the momentum per unit volume and the latent heat. As with all empirical heat transfer correlations, the above equation is valid over the range of experimental test conditions for which it was derived, which are given in detail in the source reference [27, 28]. Although this correlation has not been tested for the conditions that match typical cryogen spray cooling of the skin, it can provide a generalized basis for evaluation of such processes.

More recent measurements using subcooled R-134a sprayed onto skin demonstrated less sensitivity to droplet size and nozzle to surface distance under conditions for which spray cooling is applied with laser irradiation [24]. Empirical correlations have been applied to identify a peak value for the convective heat transfer coefficient to be on the order of 4000 W/m^2·K [29], noting that the magnitude of h during a spay cooling episode increases rapidly (on the order of milliseconds) from 0 to a maximum value and then decreases rapidly to 0 at the termination [30]. This transient data provides the basis for the thermal boundary conditions at the surface for the cooling process [31].

Spray cooling of a cryogen at subzero temperatures runs the risk of causing ice formation, either in the exterior boundary to the skin by frost formation from water vapor in the ambient air or additionally within the skin itself. From nearly all aspects, ice formation is an undesirable result. A layer of ice will significantly slow the rate of cooling of temperature by sensible heat loss by imposing a layer undergoing phase change at the skin surface. Liberation of the latent heat of fusion from liquid water when it freezes represents a large heat sink to absorb and energy flow that would otherwise be cooling the tissue. Because freezing occurs at constant temperature, the process will impose an isothermal boundary condition of 0°C onto the skin. Freezing of the skin will result in frostbite, with a resultant likelihood of tissue injury and possible scarring. Thus, ice formation during spray cooling is to be avoided, and this effect is neglected in the present analysis. Inclusion of ice formation effects would require a more sophisticated analysis procedure to account for a moving energy absorbing boundary, which is known as a classic Stefan problem in heat transfer.

The spray cooling process in conjunction with laser irradiation of skin has been analyzed from the specific perspective of augmenting laser irradiation of the skin to protect near surface structures from thermal injury [22, 23]. A complete 3-D analysis requires three sets of dependent equations: (1) light propagation of the laser beam, (2) local heating and heat conduction created by the absorption of laser light, and (3) cooling owing to the cryogen spray. These equations may include parameters that are a function of tissue temperature and water content. This process can be simplified for analysis by assuming that the dynamics of the jet impingement and

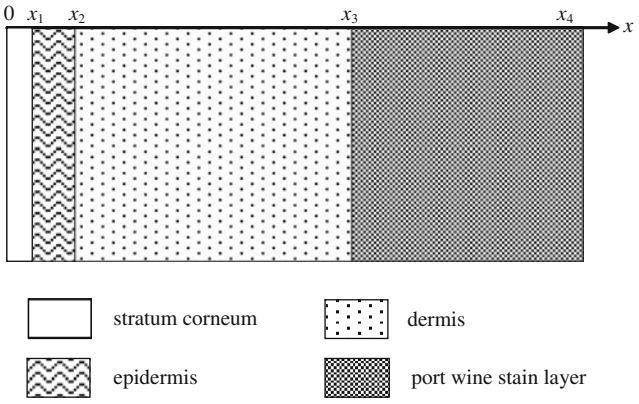


Fig. 10.19 Composite system for analysis of jet impingement cooling of the skin surface with a cryogen spray. A boundary layer consisting of cryogen spray produces convective cooling with heat transfer coefficient h. Absorption of laser irradiation produces a heat source in the epidermis and PWS layer

evaporation result in a uniform material boundary on the skin surface consisting of cryogen. The geometry for this example is simplified to a one dimensional, multi-layered, semi-infinite medium in Cartesian coordinates as shown in Fig. 10.19. Additional insight is provided in Chapter 23.

The transient heat transfer process can be considered in terms of two separate phenomena. One is the spray cooling onto the skin surface which will issue in the penetration of a thermal cooling wave into the underlying tissues. The second is the absorption of laser energy distributed as a heat source on the interior of the tissue as a function of the optical properties of the laser light and the matching spectral absorption properties of the various tissue components. The heating is assumed to occur very rapidly during a brief laser pulse (e.g., 450 μs) that is short in comparison with the subsequent thermal diffusion processes within the skin. Thus, the spatial temperature profile that is produced at the end of the laser pulse, relative to the starting uniform value, T_i, approximates the initial conditions for the one dimensional transient heat transfer that determines whether the Port Wine Stain (PWS) vessels will be destroyed and the epidermis undamaged. The temperature distribution pattern in the skin is characterized based on assuming no heat conduction prior to the end of the laser pulse and heat is generated only in the melanin of the epidermis and in the vasculature of the PWS tissue volume. This assumes the absorption coefficient of the dermis is zero. The concentration of blood in the PWS layer and the absorption coefficient of blood lead to an effective absorption coefficient for the PWS layer of μ_a^{blood}. The resulting initial temperature rise within the skin at the end of the laser pulse is given by

$$T(x,0) - T_i = \Delta T_{0,\text{epid}} = \frac{\phi(x)\, t_0 \mu_a (epi)}{\rho c} \quad \text{for} \quad x_1 \le x \le x_2 \quad (10.84a)$$

$$T(x,0) - T_i = \Delta T_{0,PWS} = \frac{\phi(x)\, t_0 \mu_a^{\text{blood}}}{\rho c} \quad \text{for} \quad x_3 \le x \le x_4 \tag{10.84b}$$

$$T(x,0) - T_i = 0 \quad \text{since} \quad \mu_a = 0 \quad \text{for} \quad x \le x_1, \quad x_2 \le x < x_3, \quad x > x_4 \tag{10.84c}$$

where $\phi(x)$ is the local fluence rate and t_0 is the laser pulse duration. The fluence rate, which is typically larger than the irradiance (see Fig. 3.15 and Chapter 23), can be determined using Monte Carlo (Chapter 5) or Diffusion Theory (Chapter 6). Anvori *et al.* further simplify their analysis by assuming a constant temperature across the epidermis and Beers Law absorption in the blood plexus PWS layer [23]. An initial laser induced temperature using Anvari's et al. assumptions is illustrated in Fig. 10.20. The temperature distribution provides the initial temperature boundary condition for the laser heating portion of the total solution.

It is important to note that the cooling boundary condition can be controlled in time to manipulate the temperature distribution caused by the absorption of laser light within the skin by applying a differential start time for the spraying and irradiation. In particular, pre-cooling will initiate a cold wave to move into the skin. With proper timing, the penetration depth of the cooling can be limited to an intermediate level so that the epidermis is protected from thermal injury whereas the PWS remains fully sensitive to the therapeutic effects of the laser irradiation. This type of cooling strategy will alter the initial thermal conditions for the laser heating process to be more complex than defined in Eq. (10.84), with the superposition of the preliminary cooling pattern.

The cooling at the surface is characterized in terms of a simple convective heat transfer coefficient, h, that is uniform across the boundary area of interest. Thus, the boundary condition is expressed as Eq. (10.10) during active cooling and is assumed to be insulated before and after.

$$-k_f \left.\frac{dT}{dx}\right|_{x=0} = h\,(T_s - T_\infty) \tag{10.84d}$$

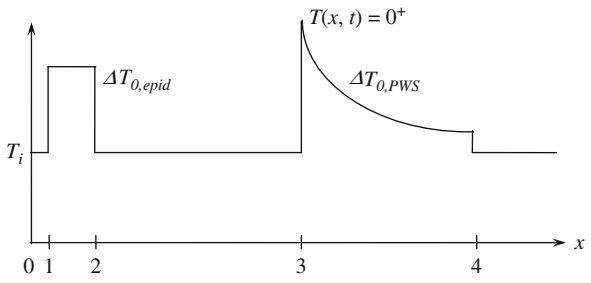


Fig. 10.20 Increase in temperature at end of laser pulse. The curve provides the initial temperatures for the heating solution

Note that the entire effect of the spray cooling is lumped into the convective heat transfer coefficient, h. According to Newton's law of cooling, Eq. (10.9), it should be possible to derive a value for h from a correlation relation such as Eq. (10.79).

$$\dot{q} = \frac{\dot{Q}}{A} = h\left(T_\infty - T_s\right) = h\Delta T \tag{10.85}$$

For pre-cooling, the boundary heat transfer effect creates a temperature profile such as depicted in Fig. 10.21 where T_i is the non-cooled initial temperature at the surface of the skin.

Examples of cooling temperature profiles as a function of cooling times which are typically in tens of milliseconds are presented in Chapter 23.2.5 and Ref. [22].

By incorporating the laser heating cycle into the initial conditions and assuming that blood perfusion has no significant thermal effect, the transient tissue temperature is described by the basic diffusion equation (Eq. (10.8)) which is written in a single spatial dimension with constant property values as

$$\frac{\partial T\left(x,t\right)}{\partial t} = \alpha \frac{\partial^2 T}{\partial x^2} \tag{10.86}$$

The general form of the heat conduction equation due to laser heating is given by Carslow and Jaeger [13].

$$\Delta T\left(x,\ t\right) = \frac{1}{2\sqrt{\pi\,\alpha\,t}} \left\{ \int_0^\infty \Delta t_0\left(x'\right) \left[e^{\frac{-\left(x-x'\right)^2}{4\alpha t}} - e^{\frac{-\left(x+x'\right)^2}{4\alpha t}} \right] dx' \right. \tag{10.86a}$$

where α is the thermal diffusivity. Equation (10.86a) provides a general solution for any shape of $\Delta t_0(x)$.

Several application cases can be considered: thermal response to laser heating with no cooling, with precooling, and with both precooling and postcooling. The general form of the solution to the diffusion problem in a semi-infinite system is

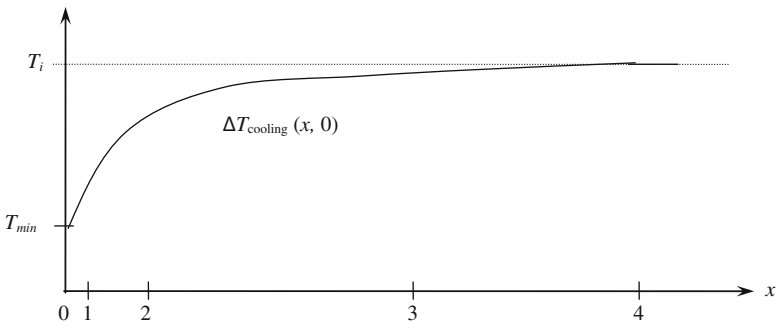


Fig. 10.21 Temperature profile produced by precooling at the onset of the laser pulse

the error function. Accordingly, the analytical solutions follow for each of the three cases, based on Carslaw and Jaeger [13]. The zero point of the adjustable time scale is set at the completion of the laser irradiation of duration t_{laser} as defined by $t^* = t–t_{\text{laser}}$. Since there are two spatial temperature sources in the system, the temperature rise at any location in the skin is the superimposed sum of the effects diffusing from both of the initial areas of temperature rise (in the epidermis and port wine stain layers). Thus,

$$\Delta T\left(x,t^*\right) = \Delta T_{\text{epid}}\left(x,t^*\right) + \Delta T_{PWS}\left(x,t^*\right) \tag{10.87}$$

Case 1. No Cooling: For the first case in which there is no cooling, the boundary condition is insulated and the temperatures at $t^* = 0$ is given by Eq. (10.84) and depicted in Fig. 10.20. The solutions for the two components of heating in the preceding equation are

$$\Delta T_{\text{epid}}\left(x,t^*\right) = \frac{\Delta T_{0,\text{epid}}}{2}\left\{\text{erf}\left(\frac{x_i}{2\sqrt{\alpha t^*}} - \frac{x}{2\sqrt{\alpha t^*}}\right) + \text{erf}\left(\frac{x_i}{2\sqrt{\alpha t^*}} + \frac{x}{2\sqrt{\alpha t^*}}\right)\right\}_{x_1/2\sqrt{\alpha t^*}}^{x_2/2\sqrt{\alpha t^*}} \tag{10.88}$$

$$\Delta T_{PWS}\left(x,t^*\right) = \frac{\Delta T_{0,PWS}}{2}\exp\left[2\mu_a^{\text{blood}}\sqrt{\alpha\, t^*}\left(\frac{x_3}{2\sqrt{\alpha\, t^*}} + \mu_a^{\text{blood}}\sqrt{\alpha\, t^*}\right)\right]$$
$$\left\{\exp\left[-2\frac{x}{2\sqrt{\alpha t^*}}\mu_a^{\text{blood}}\sqrt{\alpha t^*}\right] erf\left(\frac{x_i}{2\sqrt{\alpha t^*}} - \frac{x}{2\sqrt{\alpha t^*}} + \mu_a^{\text{blood}}\sqrt{\alpha t^*}\right)\right.$$
$$\left. + \exp\left[2\frac{x}{2\sqrt{\alpha t^*}}\mu_a^{\text{blood}}\sqrt{\alpha t^*}\right] erf\left(\frac{x_i}{2\sqrt{\alpha t^*}} + \frac{x}{2\sqrt{\alpha t^*}} + \mu_a^{\text{blood}}\sqrt{\alpha t^*}\right)\right\}_{x_3/2\sqrt{at^*}}^{x_4/2\sqrt{at^*}} \tag{10.89}$$

where μ_a^{blood} is the effective optical absorption coefficient of blood at the laser wavelength (see Chapter 3). $\mu_a^{\text{blood}} = c\,(BL)\,\mu_{BL}\,(BL)$ where $c(BL)$ is the concentration of blood in the PWS layer and $\mu_a(BL)$ is the wavelength dependent absorption coefficient of blood based on Hb and HbO (see Chapter 3).

Case 2. Precooling: When the skin is precooled prior to laser irradiation, the initial conditions are altered as the cooling wave penetrates the skin. For the period prior to irradiation, time is measured from the beginning of cooling, and the temperature history is described in terms of the classic problem of diffusion in a semi-infinite medium with a convective boundary condition.

$$\Delta T_{\text{cooling}}\left(x,t\right) = \left(T_\infty - T_0\right)\left\{\text{erfc}\left(\frac{x}{2\sqrt{\alpha t}}\right) - \left[\exp\left(\frac{hx}{k} + \frac{h^2\alpha t}{k^2}\right)\right]\times \times\left[\text{erfc}\left(\frac{x}{2\sqrt{\alpha t}} + \frac{h\sqrt{\alpha t}}{k}\right)\right]\right\} \tag{10.90}$$

The complementary error function, erfc, is defined by erfc = 1–erf. If the cooling spray is discontinued when the short laser irradiation starts, the cooling imposes an initial temperature distribution as shown in Fig. 10.21 along with the initial temperature distribution due to absorption of laser light. An exact analytical function does not exist for this problem, so an approximation has been developed [23], although there may be many options for an approximation beyond that chosen by those authors. The result is that the post heating temperature is given by a superposition of the combined effects of cooling and the two layers of laser heating where there still exist residual temperature alterations owing to the cooling after that process is terminated.

$$\Delta T\left(x,t^*\right)=\Delta T_{\text{epid}}\left(x,t^*\right)+\Delta T_{PWS}\left(x,t^*\right)+\Delta T_{\text{cooling}}\left(x,t^*\right) \tag{10.91}$$

Case 3: Simultaneous Cooling and Later Heating: For the conditions of continuous cooling a superposition solution is given in Equation (10.91) with the individual terms determined for the appropriate initial and boundary conditions. The expression of cooling effect applies throughout the entire process, Equation (10.90). The combined heating and cooling terms are given by

$$\Delta T_{\text{epid}}\left(x,t^*\right)=\frac{\Delta T_{0,\text{epid}}}{2}\left\{\text{erf}\left(\frac{x_i}{2\sqrt{\alpha t^*}}-\frac{x}{2\sqrt{\alpha t^*}}\right)+\text{erf}\left(\frac{x_i}{2\sqrt{\alpha t^*}}+\frac{x}{2\sqrt{\alpha t^*}}\right)\right.$$
$$\left.-2\exp\left[\frac{x}{2\sqrt{\alpha t^*}}+\frac{x_i}{2\sqrt{\alpha t^*}}\right]^2\exp\left[\frac{x}{2\sqrt{\alpha t^*}}\right]^2\text{erfc}\left(\frac{h}{k}\sqrt{\alpha t^*}+\frac{x_i}{2\sqrt{\alpha t^*}}+\frac{x}{2\sqrt{\alpha t^*}}\right)\right\}_{x_1/2\sqrt{\alpha t^*}}^{x_2/2\sqrt{\alpha t^*}} \tag{10.92}$$

$$\Delta T_{PWS}\left(x,t^*\right)=\Delta T_{0,PWS}\exp\left[2\mu_a^{\text{blood}}\sqrt{\alpha\, t^*}\,\frac{x_3}{2\sqrt{\alpha\, t^*}}\right]\bullet$$
$$\left\{\exp\frac{\left[2\mu_a^{\text{blood}}\sqrt{\alpha t^*}\right]^2}{2}\left(\exp\left[-2\mu_a^{\text{blood}}\sqrt{\alpha t^*}\,\frac{x}{2\sqrt{\alpha t^*}}\right]\text{erf}\left[\frac{x_i}{2\sqrt{\alpha t^*}}-\frac{x}{2\sqrt{\alpha t^*}}+\mu_a^{\text{blood}}\sqrt{\alpha t^*}\right]\right.\right.$$
$$\left.+\left[\frac{\mu_a^{\text{blood}}\sqrt{\alpha t^*}+\frac{h}{k}\sqrt{\alpha t^*}}{\mu_a^{\text{blood}}\sqrt{\alpha t^*}-\frac{h}{k}\sqrt{\alpha t^*}}\right]\exp\left[2\mu_a^{\text{blood}}\sqrt{\alpha t^*}\,\frac{x}{2\sqrt{\alpha t^*}}\right]\text{erf}\left[\frac{x_i}{2\sqrt{\alpha t^*}}-\frac{x}{2\sqrt{\alpha t^*}}+\mu_a^{\text{blood}}\sqrt{\alpha t^*}\right]\right)$$
$$+\left[\frac{\frac{h}{k}\sqrt{\alpha t^*}}{\mu_a^{\text{blood}}\sqrt{\alpha t^*}-\frac{h}{k}\sqrt{\alpha t^*}}\right]\exp\left[-\left(\frac{x}{2\sqrt{\alpha t^*}}+\frac{x_i}{2\sqrt{\alpha t^*}}\right)^2-2\frac{x_i}{2\sqrt{\alpha t^*}}\mu_a^{\text{blood}}\sqrt{\alpha t^*}\right]\bullet$$
$$\left.\exp\left[\frac{x_i}{2\sqrt{\alpha t^*}}+\frac{x}{2\sqrt{\alpha t^*}}+\mu_a^{\text{blood}}\sqrt{\alpha t^*}\right]^2\text{erfc}\left(\frac{x_i}{2\sqrt{\alpha t^*}}+\frac{x}{2\sqrt{\alpha t^*}}+\mu_a^{\text{blood}}\sqrt{\alpha t^*}\right)\right\}_{x_3/2\sqrt{\alpha t^*}}^{x_4/2\sqrt{\alpha t^*}} \tag{10.93}$$

These transient temperature solutions can be applied with a damage rate model to calculate the level of injury to tissue as a function of position and treatment protocol.

10.5 Comments and Conclusion

Heat transfer is an important component of the laser irradiation process as there is a direct coupling with the temperature field. The temperature governs critical rate processes relating to homeostasis, the manifestation of thermal injury, and thermal therapy processes. Laser irradiation is used for many cases to elicit injury or therapeutic procedures. Thermal analysis provides an effective tool to predict the outcome of specific irradiation protocols and to design procedures for accomplishing specific objectives.

References

1. Incroprera FP, DeWitt DP, Bergman TL, and Lavine AS. *Fundamentals of heat and mass transfer*. Wiley, Hoboken, NJ, 6th edition (2007).
2. Kreith F and Bohn MS. *Principles of heat transfer*. West publishing Co., St. Paul, 5th edition (1993).
3. Minkowycz WJ, Sparrow EM, Schneider GE, and Pletcher RH. (eds). *Handbook of numerical heat transfer*. Wiley, New York (1988).
4. Kreith F and Goswami Y. *The CRC handbook of mechanical engineering*. CRC, Boca Raton, 2nd edition (2005).
5. Pennes HH. Analysis of tissue and arterial blood temperatures in the resting forearm. *J. Appl. Physiol.*, 1:88–122 (1948).
6. Charney CK. Mathematical models of bioheat transfer. *Adv. Heat Trans*., 22:19–156 (1992).
7. Wissler EH. Pennes' 1948 paper revisited. *J. Appl. Physiol*., 85:35–41 (1998).
8. Diller KR, Valvano J, and Pearce JA. Bioheat transfer. In: F Kreith and Y Goswami (eds) *The CRC handbook of mechanical engineering*. CRC, Boca Raton, 2nd edition, pp. 4-282–4-361 (2005).
9. Alexiades V and Solomon AD. *Mathematical modeling of melting and freezing processes.* Hemisphere Publishing, Washington (1993).
10. Tong LS. *Boiling heat transfer and two phase flow*. Taylor and Francis, Washington, DC (1997).
11. Crank J. *Free and moving boundary problems.* Clarendon, Oxford (1984).
12. Valvano JW. Tissue thermal properties and perfusion. In: AJ Welch and MJC van Gemert (eds) *Optical-thermal response of laser irradiated tissue,* 2nd edition. Springer (2010).
13. Carslaw HS and Jaeger JC. *Conduction of heat in solids.* Oxford University, London, 2nd edition (1959).
14. Morse PM and Feshback H. *Methods of theoretical physics, parts I and II.* McGraw-Hill, New York (1953).
15. Kays WM, Crawford MF, and Weigand B. *Convective heat and mass transfer.* McGraw Hill, New York, 4th edition (2004).
16. Bejan A. *Convection heat transfer.* Wiley, Hoboken, NJ (2004).
17. Siegel R and Howell JR. *Thermal radiation heat transfer.* Taylor and Francis, New York, 4th edition (2002).
18. Modest MF. *Radiative heat transfer.* McGraw Hill, New York (1993).
19. Howell JR. *Catalogue of radiation configuration factors.* McGraw Hill, New York (1982).
20. Chato JC. Heat transfer to blood vessels. *J. Biomech. Engr.*, 102:110–118 (1980).

21. Chen MM and Holmes KR. Microvascular contributions in tissue heat transfer. *Ann. N.Y. Acad. Sci.*, 335:246–251 (1980).
22. Anvari B, Milner TE, Tanenbaum BS, Kimel S, Svaasand LO, and Nelson JS. Selective cooling of biological tissues: Application for the thermally mediated therapeutic procedures. *Phy. Med. Biol.*, 40:241–252 (1995).
23. Anvari B, Tanenbaum BS, Milner TE, Kimel S, Svaasand LO, and Nelson JS. A theoretical study of the thermal response of skin to cryogen spray cooling and pulsed laser irradiation: implications for treatment of port wine stain birthmarks. *Phy. Med. Biol.*, 40:1451–1465 (1995).
24. Pikkula BM, Torres JH, Tunnell JW, and Anvari B. Cryogen spray cooling: Effects of droplet size and spray density on heat removal. *Lasers Surg. Med*., 28:103–112 (2001).
25. Hodgson JW and Sutherland JE. Heat transfer from a spray cooled isothermal cylinder. *Indy. Engr. Chem. Fundam.*, 7:567–571 (1968).
26. Martin H. Heat and mass transfer between impinging gas jets and solid surfaces, *Advan. Heat Trans.*, 13:1–60 (1977).
27. Ghodbane B and Holman JP. Experimental study of spray cooling with Freon-113. *Int. J. Heat Mass Transfer.*, 34:1163–1174 (1991).
28. Holman JP and Kendall CM. Extended studies of spray cooling with Freon-113. *Int. J. Heat Mass Transfer.*, 36:2239–2241 (1993).
29. Torres JH, Nelson JS, Tanenbaum BS, Milner TE, Goodman DM, and Anvari B. Estimation of internal skin temperatures in response to cryogen spray cooling: Implications of laser therapy of port wine stains. *IEEE J. Sel. Top. Quantum Electron.*, 59:23–32 (1999).
30. Tunnell JW, Torres JH, and Anvari B. Methodology for estimating time-dependent surface heat flux due to cryogen spray cooling. *Annals Biomed. Engr.*, 30:19–33 (2002).
31. Tunnell JW, Wang LV, and Anvari B. Optimum pulse duration and radiant exposure for vascular laser therapy of dark port-wine skin: a theoretical study. *Appl. Opt*., 42:1367–1378 (2003).

Chapter 11
Temperature Measurements

John A. Pearce, Jonathan W. Valvano, and Stanislav Emelianov

11.1 Introduction

The objective of this chapter is to present the fundamental mechanisms, instrumentation techniques, and error analyses for temperature measurements in laser irradiated biologic media. Because temperature is a significant biological parameter, it is important to understand and minimize potential measurement errors [1–8]. Temperature measurements in a radiative field are particularly difficult because:

- There is direct optical absorption of laser energy into a temperature sensor;
- The temperature can be up to 300°C;
- The spatial temperature gradients can exceed 50°C/mm; and
- The laser pulse durations can be as short as 100 fs.

Laser applications involving temperature measurements extend over a wide range of temperatures and time scales. Hyperthermia typically involves temperatures less than 50°C for durations on the order of an hour. In this case perfusion is a very important factor. Photodynamic therapy also has long irradiation times at low temperatures. On the other hand, the use of lasers for coagulation involves pulse times ranging from 250 μs to 1 s and achieves temperatures up to 100°C. The most challenging application is ablation where pulse times can be as short as 10 ns, and temperatures routinely exceed 300°C.

J.A. Pearce (✉)
Department of Electrical and Computer Engineering, The University of Texas, Austin, TX, USA
e-mail: jpearce@mail.utexas.edu

A.J. Welch, M.J.C. van Gemert (eds.), *Optical-Thermal Response of Laser-Irradiated Tissue*, 2nd ed., DOI 10.1007/978-90-481-8831-4_11,

11.2 General Concepts

11.2.1 Classification of Temperature Transducers

11.2.1.1 Thermal Expansion

An increase in thermal energy causes an increase in the average spacing between molecules. For an ideal gas, the kinetic energy ($1/2\ mv^2$) equals the thermal energy ($3/2\ kT$). For liquids and most gases, an increase in kinetic energy is also accompanied by an increase in potential energy. The mercury thermometer and bimetallic strips are common transducers that rely on thermal expansion. A resistance temperature device (RTD) is constructed from long narrow gauge metal wires. Over a moderate range of temperature (e.g., 50°C), the wire resistance, R, is approximately linearly related to its temperature.

$$R = R_0\left[1 + \gamma\,(T - T_0)\right] \tag{11.1}$$

where R_0 is the resistance at a fixed temperature T_0. R_0 is a function of the resistivity and geometry of the wire. The fractional sensitivity, γ, for various metals is given in Table 11.1. Although many metals such as gold, nichrome (nickel-chromium), nickel and silver can be used, platinum is typically selected because of its excellent stability.

11.2.1.2 Boltzmann Factor

The rates of first order processes are proportional to the Boltzmann factor, $e^{-E/kT}$. This strong temperature dependence can be exploited to make measurements. Rates that are governed by the Boltzmann factor include the evaporation of liquids and the population of charge carriers in a conduction band. The conductance, G (1/resistance), of a thermistor is proportional to the occupation of charge carriers in the conduction band:

$$G = G_0 e^{-E/kT} \tag{11.2}$$

Table 11.1 Electrical properties of various materials [9]

Material	γ, fractional sensitivity (1/°C)	ρ, electrical resistivity (Ω-cm)
Gold	+0.004	2.35×10^{-6}
Nickel	+0.0069	6.84×10^{-6}
Platinum	+0.003927	10^{-5}
Copper	+0.0068	1.59×10^{-6}
Silver	+0.0041	1.673×10^{-6}
NTC Thermistor	−0.04	10^{3}
PTC Thermistor	+0.1	

where E is the activation energy [J], k is Boltzmann's constant [1.3805×10^{-23}J/K], and T is the temperature in K. The electrical resistance, R, is the reciprocal of the conductance:

$$R = R_0 e^{+E/kT} = R_0 e^{+\beta/T} \quad \text{or} \quad 1/T = H_0 + H_1 \ln(R) \tag{11.3}$$

where β[K], $R_0[\Omega]$, H_0 and H_1 are calibration coefficients. H_0 is $-\ln(R_0)/\beta$, and H_1 is $1/\beta$.

11.2.1.3 Seebeck Effect

A thermocouple is constructed using two wires of different metals welded to form two junctions. One junction is placed at a known reference temperature (e.g., 0°C ice in thermal equilibrium with liquid and gaseous water at 1 atmosphere) and the other junction is used to measure the unknown temperature. The Seebeck effect involves thermal to electrical energy conversion. In a closed loop configuration, the current around the loop is proportional to the temperature difference between the two junctions. In the open loop configuration, a voltage is generated proportional to the temperature difference:

$$V = a(T_2 - T_1) + b(T_2 - T_1)^2 \tag{11.4}$$

where a and b are calibration coefficients.

11.2.1.4 Radiation

Electromagnetic radiation is emitted by all materials above absolute zero, of which black or gray bodies are a special case. The spectral properties of a surface are a function of the material and its temperature. The net wide-band radiation power, W [W/cm^2] of an object in an enclosure, is a function of the fourth power of surface temperature.

$$W = \sigma \left(T^4 - T_0^4\right) \tag{11.5}$$

where σ is the Stefan-Boltzmann constant [5.6697×10^{-12} W/cm$^2 - K^4$], T is the surface temperature of the object (emitted power), and T_0 [K] is the temperature of the surroundings (absorbed power).

11.2.1.5 Phase Transitions

The sudden transition of state that occurs for a particular substance can be used to measure temperature. In particular, this property is frequently exploited to produce temperature references. An ice bath at 0°C (water freezing point), is often used as the reference junction of a thermocouple. The melting point of gallium is conveniently 29.7714°C [10]. A triple point occurs under pressure when a substance exists simultaneously in all three states: gas, liquid, and solid. The triple point of water is

0.01±0.0005°C at 1 atmosphere [11], the triple point of rubidium is 39.265°C, and the triple point of succinonitrile is 58.0805°C.

11.2.2 Transducer Specifications

11.2.2.1 Transducer Linearity

Let T_i be the temperature input and X_i be the output (voltage or resistance) signal of the transducer, e.g., Fig. 11.1. The *linearity* is a measure of the straightness of the static calibration curve. Let $X_i = sT_i + b$ be the best fit line through the transducer data. Linearity (or deviation from it) as a figure-of-merit can be expressed as percentage of reading or percentage of full scale. Let $X_{\max}$ be the largest transducer output.

$$\text{Average linearity of reading (percent)} = \frac{100}{n} = \sum_{i=l}^{n} \frac{|X_i - sT_i - b|}{X_i} \tag{11.6}$$

$$\text{Average linearity of full scale (percent)} = \frac{100}{n} = \sum_{i=l}^{n} \frac{|X_i - sT_i - b|}{X_{\max}} \tag{11.7}$$

11.2.2.2 Transducer Sensitivity

Two definitions for sensitivity are used for temperature transducers. The *static sensitivity* is:

$$s = \frac{\partial X}{\partial T}\,[X/^{\circ}\mathrm{C}] \tag{11.8}$$

If the transducer is linear, then the static sensitivity is the slope, s, of the straight line through the static calibration curve that gives the minimum mean squared error.

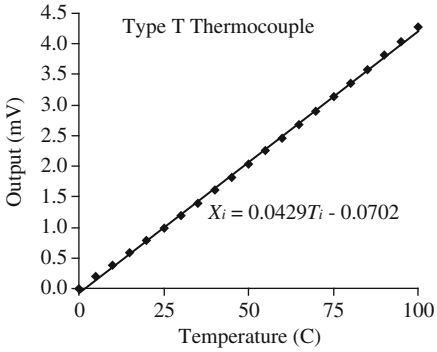


Fig. 11.1 Typical thermocouple response

Thermistors can be manufactured to have a resistance value at 25°C ranging from 30 Ω to 20 MΩ [12–16]. Because the interface electronics can just as easily convert any resistance into a voltage, a 20 MΩ thermistor is not more sensitive than a 30 Ω thermistor. In this situation, it makes more sense to define *fractional sensitivity* as:

$$\gamma = \frac{1}{X} \cdot \frac{\partial X}{\partial T} \tag{11.9}$$

11.2.2.3 Time Constant

The transient response of a temperature transducer to a sudden change in tissue temperature can be approximated by an exponential equation (assuming first-order response):

$$T(t) = T_f + (T_0 - T_f)_e^{-t/\tau} \tag{11.10}$$

where T_0 and T_f are the initial and final temperature respectively. The *time constant*, τ, of a transducer is the time to reach 63.2% of the final output after the temperature is instantaneously increased. This time is dependent on both the transducer and the experimental setup. Manufacturers often specify the time constant of thermistors and thermocouples in well-stirred oil (fastest) or still air (slowest). In biomedical applications, one must consider the situation. If the transducer is placed in a high flow artery like the pulmonary artery, it is reasonable to use the stirred oil time constant. If the transducer is embedded in a resting muscle, then thermal conduction in the tissue will determine the time constant almost independently of the type of transducer.

11.2.2.4 Specificity

Unfortunately, temperature transducers are often sensitive to factors other than temperature (e.g., humidity, pressure, motion, acceleration, vibration, shock, radiation fields, electric fields and magnetic fields.) *Specificity* is a measure of relative sensitivity of the transducer to the desired signal (in our case temperature) compared to the sensitivity of the transducers to these other unwanted influences. A transducer with a good specificity will respond only to temperature and be independent of these disturbing factors. On the other hand, a transducer with a poor specificity will respond to temperature as well as to some of these disturbing factors. If all these disturbing factors are grouped together as noise, then the *signal-to-noise ratio (S/N)* is a quantitative measure of the specificity of the transducer.

11.2.2.5 Transducer Impedance

The input range is the allowed range of the input temperature, *T*. The *input impedance* is the steady state or sinusoidal effort (voltage, force, pressure) divided

by the steady state or sinusoidal flow (current, velocity, flow). The input impedance of a thermal sensor is a measure of the thermal perturbation that occurs due to the presence of the probe itself in the tissue. For example, a thermocouple needle inserted into a laser irradiated tissue will affect the tissue temperature because heat will conduct down the stainless steel shaft. A thermocouple has a low input impedance (which is bad) because the transducer itself loads (reduces) the tissue temperature. On the other hand, an infrared detector measures surface tissue temperature without physical contact. Infrared detectors therefore have a very high input impedance (which is good) because the presence of the transducer has no effect on the temperature to be measured. In the case of temperature sensors, the driving force for heat transfer is the temperature difference, ΔT. The resulting heat flow, q [W], can be expressed using Fourier's law of thermal conduction:

$$q = kA\frac{\partial T}{\partial x} \approx k4\pi\, a^2\frac{\Delta T}{a} \tag{11.11}$$

where k [W/cm-°C] is the probe thermal conductivity, and A [cm^2] is the probe surface area and a [cm] the radius of a spherical transducer. The steady state input impedance of a spherical temperature probe can thus be approximated by:

$$Z \equiv \frac{\Delta T}{q} \approx \frac{1}{4\pi ak} \quad [°\text{C/W}] \tag{11.12}$$

Again, the approximation in the above equation assumes a spherical transducer. Similar discussions can be constructed for the sinusoidal input impedance. Since most thermal events can be classified as step events rather than sinusoidally varying events, most researchers prefer the use of time constant to describe the transient behavior of temperature transducers.

The output signal strength of the transducer can be specified by its electrical output resistance, R_{out}, and output capacitance, C_{out}. See Fig. 11.2. Some transducers are completely passive (e.g., thermocouples, thermopiles, and infrared detectors), while others are active requiring external power (e.g., piezoelectric crystals, resistance temperature devices, and thermistors.)

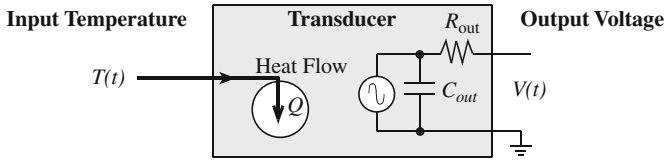


Fig. 11.2 Block diagram of a temperature transducer

11.2.3 Instrument Specifications

11.2.3.1 Accuracy

The *instrument accuracy* is the absolute error referenced to the National Institute of Standards and Technology (NIST, in the US) of the entire system including transducer, electronics, and software. Let T_{mi} be the values as measured by the instrument, and let T_{ti} be the true values from NIST references. Accuracy can be defined many ways:

$$\text{Average accuracy } [^\circ\text{C}] = \frac{1}{n}\sum_{i=l}^{n} |T_{ti} - T_{mi}| \tag{11.13}$$

$$\text{Maximum error } [^\circ\text{C}] = \max |T_{ti} - T_{mi}| \tag{11.14}$$

$$\text{Standard error } [^\circ\text{C}] = \sqrt{\frac{1}{n-1}\sum_{i=l}^{n} [T_{ti} - T_{mi}]^2} \tag{11.15}$$

In biomedical applications, the temperature of interest is often a relative quantity and the Celsius and Fahrenheit scales have arbitrary zeroes (e.g., 0°C is the freezing point of water, 37°C is a typical body temperature). Unless you are measuring differential temperature (between two positions or from one time to another) or absolute temperature (e.g., 0–300 K), it is inappropriate to specify temperature error as a percentage of reading. For example, if your device measures temperature at 20°C, with an error of 1°C, one might be tempted to specify the error is 5% of reading. However, if you simply convert the temperature to °F, then the device measures this same temperature, 68°F, with an error of 1.8°F. Inappropriately, we can now claim this same device has an error of 2.6% of reading. On the other hand, if you wish to specify accuracy as a percent, it is appropriate to scale the measurement errors by the full scale range of the system.

$$\text{Average accuracy } [\%] = \frac{100\sum_{i=l}^{n} |T_{ti} - T_{mi}|}{n\,(T_{\max} - T_{\min})} \tag{11.16}$$

$$\text{Maximum error } [\%] = \max \frac{100\,|T_{ti} - T_{mi}|}{(T_{\max} - T_{\min})} \tag{11.17}$$

where the instrument range is $T_{\min}$ to $T_{\max}$.

11.2.3.2 Resolution

The *instrument resolution* is the smallest temperature difference, ΔT, that can be detected by the entire system including transducer, electronics, and software. The resolution of the system is sometimes limited by noise processes in the transducer itself (e.g., thermal imaging) and sometimes limited by noise processes in

the electronics (e.g., thermistors, RTDs, and thermocouples). One way to determine resolution is to take repeated measurements at one fixed temperature, increase the temperature by ΔT, then take a second set of measurements. An unpaired Students t-test can be used to test the null hypothesis that the means of the measured data between groups are equal. If the means of the two sets are statistically different, one can conclude the resolution is less than or equal to ΔT.

The *spatial resolution* (or *spatial frequency response*) of the transducer is the smallest distance, Δx, between two independent measurements. The spatial resolution is determined by the size and thermal properties of the probe. A metal probe will disturb the existing tissue temperature field more than a glass probe. Hence, a glass probe has a smaller spatial resolution than a metal probe of the same size. Thermal imaging systems exhibit excellent spatial resolution because the instrument does not disturb the thermal field. The spatial resolution of a thermal imaging system is the tissue surface area from which the thermal radiation originates that is eventually focused onto the detector during the imaging of a single pixel, the so-called instantaneous field of view, IFOV.

Another way to define spatial sensitivity is to consider a probe that is placed in a tissue having spatial temperature gradients. The goal of an instrument is to have the probe output (volts or ohms) be dependent on the surrounding tissue temperature. But in actuality, the probe output is only a function of its own temperature. In other words, the system measurement for a probe transducer (thermistor, resistance temperature device, or thermocouple) is determined by the spatial average temperature within the active region of the probe itself. The spatial resolution of an invasive probe depends on its size and thermal properties. In order for a system to accurately measure tissue temperature, heat transfer between the tissue and the probe must allow the probe tip (the active volume) to reach the temperature of the surrounding tissue.

11.2.3.3 Precision

Precision is the number of distinguishable alternatives from which the given result is selected – precision represents the lower limit of repeatability. Precision can be expressed in bits or decimal digits. Consider an instrument with a temperature range of 0–100°C. The system displays the output using 3 digits (e.g., 12.3°C). In addition, the system can resolve each temperature T from the temperature T+0.1°C. This system has 1001 distinguishable outputs, and hence has a precision of 1001 or about 10 bits. There is a simple relationship among range, resolution and precision: range (e.g., 100°C) equals resolution (e.g., 0.1°C) times precision (e.g., 1001 alternatives), where "range" is the maximum minus minimum temperature, and precision is specified in terms of number of alternatives.

11.2.3.4 Reproducibility or Repeatability

Reproducibility (or repeatability) is a parameter that specifies whether the instrument has equal outputs given identical inputs over some period of time. This

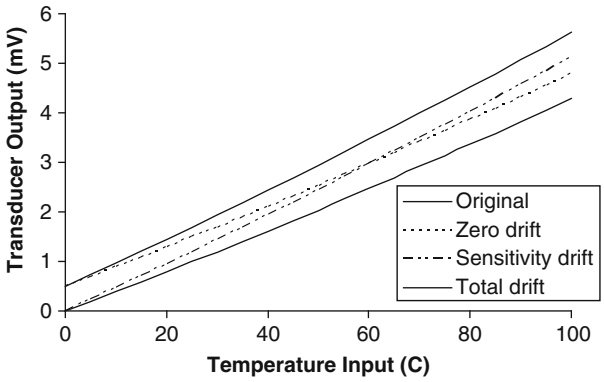


Fig. 11.3 Transducer output vs. temperature showing sensitivity drift and zero drift

parameter can be expressed as the full range or standard deviation of output results given a fixed input, where the number of samples and time interval between samples are specified. *Statistical control* is the probabilitistic model which defines the errors due to nonremovable noise. The parameter includes the noise model (e.g., normal, chi-squared, uniform, salt and pepper) and the parameters of the model (e.g., average, standard deviation).

The *zero drift* is the change in the static sensitivity curve intercept, b, as a function of time or other factor. See Fig. 11.3. The *sensitivity drift* is the change in the static sensitivity curve slope, s, as a function of time or some other factor. These drift factors determine how often the transducer must be calibrated. For example, thermistors have a drift much larger than that of Resistance Temperature Devices (RTDs) or thermocouples. Thermistors are typically aged at 300°C for long periods of time to improve their reproducibility [12–16].

11.3 Thermistors

11.3.1 Basic Principles

Thermistors are a popular temperature transducer made from a ceramic-like semiconductor [4, 11–16]. An NTC (negative temperature coefficient) thermistor is made from combinations of metal oxides of manganese, nickel, cobalt, copper, iron, and titanium [14]. A mixture of milled semiconductor oxide powders and a binder is shaped into the desired geometry. The mixture is dried and sintered (under pressure) at an elevated temperature. The wire leads are attached and the combination is coated with glass or epoxy. By varying the mixture of oxides, a range of resistance values from 30 Ω to 20 MΩ (at 25°C) is possible. Table 11.2 lists the tradeoffs between thermistors and thermocouples.

Table 11.2 Tradeoffs between thermistors and thermocouples

Thermistors	Thermocouples
More sensitive	More sturdy
Better temperature resolution	Faster response
Less susceptible to noise	Inert, interchangeable V vs. T curves
Less thermal mass	Requires less frequent calibration
Does not require a reference	More linear
$k = 0.001$W/cm-°C, $\alpha = 0.001$ cm^2/s	$k = 4.01$W/cm-°C, $\alpha = 1.17$ cm^2/s
Temperature can vary within the probe	Uniform temperature within probe
Probe itself causes less disturbance	Probe can alter temperature fields in tissue

11.3.2 Thermistor Instrumentation

Electronic instrumentation converts the thermistor resistance into a voltage that can be measured by the data acquisition system. A common method uses a bridge to convert thermistor resistance, R_T, into a voltage difference, $V_1 - V_2$, as shown in Fig. 11.4. The bridge should be powered by a low-noise reference voltage. We set $R1$ equal to $R2$, and we choose the resistance $R1$ large enough to prevent self-heating error in the thermistor. Assuming the ADC voltage range is 0 to $V_{\max}$ (e.g., 0 to +5 V), the resistance $R3$ is selected to equal the thermistor resistance at the maximum desired temperature. The gain of the instrumentation amp is selected so that the minimum desired temperature maps into the maximum ADC voltage, $V_{\max}$. The thermistor interfaced in Fig. 11.4 has a resistance of 125 kΩ at 0°C and a resistance of 10 kΩ at 50°C. The ADC sampling rate is chosen to be at least twice the maximum frequency component in the signal. The two-pole Butterworth low pass filter removes aliasing, and the cutoff frequency is selected at the ADC sampling rate. The precision of the ADC, the 60 Hz pickup noise on the thermistor, power supply noise, and electronic circuit noise all affect the overall performance of the instrument. An

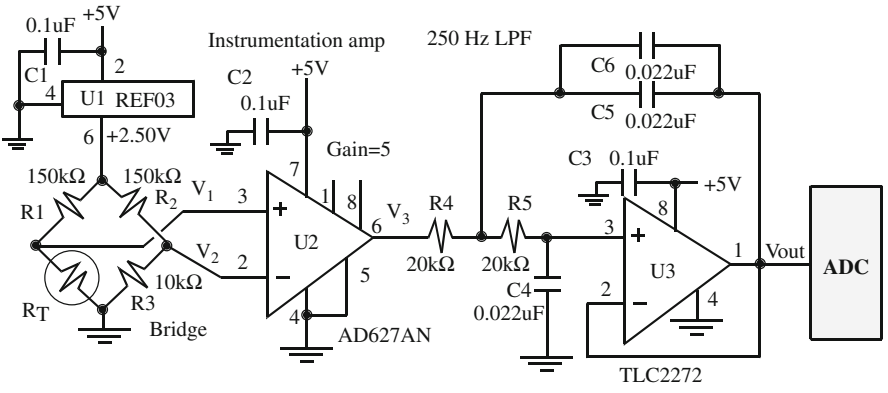


Fig. 11.4 Thermistor instrumentation using a bridge and differential amplifier

effective method is to use a calibration table with 10–20 points (ADC sample vs. temperature) and linear interpolation to convert ADC sample into temperature.

A second method to convert thermistor resistance into voltage utilizes a constant current source. In this method, it is convenient to first measure thermistor resistance, then use Eq. (11.3) to calculate temperature. This method is appropriate when one wishes to build systems with interchangeable thermistor probes and instruments, because the ohmmeter functionality can be separately calibrated from the resistance to temperature calibration [15].

A third method uses an astable multivibrator (or a monostable multivibrator) to convert resistance into a period (or pulse width). This method is inherently digital and can be used to build very low cost thermometers utilizing the timer capabilities of a microcontroller [15].

11.4 Thermocouples

11.4.1 Thermocouple Construction

A thermocouple is constructed by spot welding two different metal wires together [4, 11, 12, 15, 16, 20]. Probe transducers include a protective casing which surrounds the thermocouple junction. Probes come in many shapes including round tips, conical needles and hypodermic needles. Bare thermocouple junctions provide faster response but are more susceptible to damage and noise pickup. Ungrounded probes allow electrical isolation but are not as responsive as grounded probes. Commercial thermocouples have been constructed in 16–30 gauge hypodermic needles [12] – a 30 gauge needle has an outside diameter of above 0.03 cm. Bare thermocouples can be made from 30-μm wire producing a tip with an 80 μm diameter. A spot weld is produced by passing a large current through the metal junction that fuses the two metals together.

Cain and Welch [20] constructed Cu–Ni thermocouples with tip diameters as small as 10 μm. They began by using a heating flame to draw a 1-mm quartz rod to a tip diameter of 5 μm. Then they vacuum-deposited the two thermocouple metals on the sides of the probe such that the metals only overlapped at the tip. Lead wires were attached to the other end with silver paint. The device was covered with a thin coat of electrical insulation. Time constants for these probes were as short as 0.2 ms. Very fast probes were necessary to measure tissue temperature response to pulsed laser irradiation.

11.4.2 Thermocouple Physics

Placing the wires in a loop as depicted in Fig. 11.5 causes a current to flow depending on the temperature difference at the two junctions. This thermal to electrical

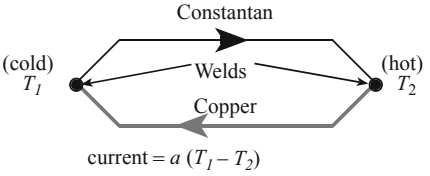


Fig. 11.5 The thermocouple thermal to electrical Seebeck effect

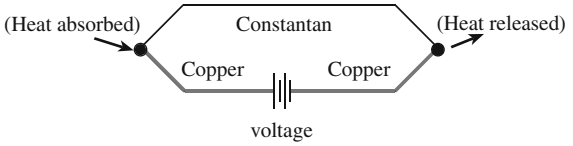


Fig. 11.6 The thermocouple electrical to thermal Peltier effect

energy conversion is called the Seebeck effect. If the temperature difference is small, then the current, *I*, is linearly proportional to the temperature difference *T*1–*T*2.

Breaking the loop and applying an electrical voltage cause heat to be absorbed at one junction and released at the other. This electrical to thermal energy conversion is called the Peltier effect (see Fig. 11.6). If the voltage is small, then the heat transferred is linearly proportional to the voltage, *V*.

To measure temperature, we use the thermocouple with an open configuration, as shown in Fig. 11.7. A voltage will develop depending on the temperature difference between the two junctions because of the Seebeck effect. If the temperature difference is small, then the voltage, *V*, is nearly linearly proportional to the temperature difference $T_1 - T_2$. Thermocouples are characterized by: (1) low impedance (resistance of the wires), (2) low temperature sensitivity (45 μV/°C for copper/constantan), (3) low power dissipation, (4) fast response (because of the metal), (5) high stability (because of the purity of the metals), and (6) interchangeability (again because of the physics and the purity of the metals).

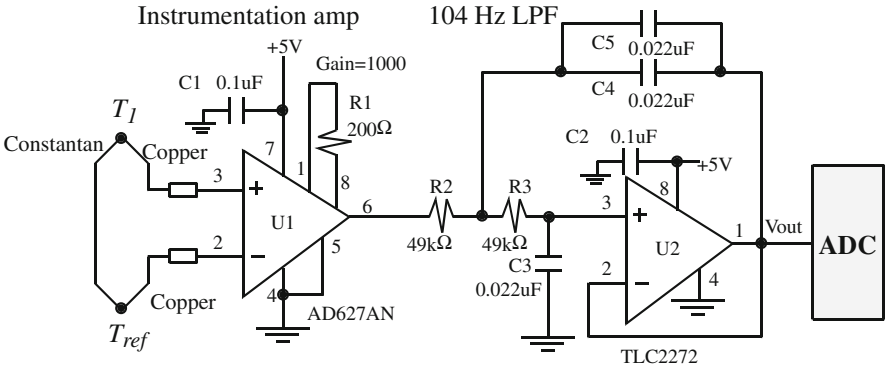


Fig. 11.7 Thermocouple instrumentation using an ice-bath reference

11.4.3 Thermocouple Characteristics

In this section we will design a thermocouple-based thermometer. The range is 0–100°C and the resolution will be 0.1°C. If we wish to measure temperature with a frequency range from 0 to 100 Hz, then the amplifier must pass 0–100 Hz and the ADC must be sampled at a rate faster than 200 Hz. Table 11.3 gives the sensitivities and temperature ranges of typical thermocouple devices. Constantan is a copper/nickel alloy. The response of a Type T copper/constantan transducer is shown in Fig. 11.1. An instrumentation amplifier with a gain of 1000 converts the 0–4.2 mV thermocouple voltage into 0–4.2 V. Figure 11.7 shows the interface electronics connecting the thermocouple to data acquisition system. Typically, an ice bath ($T_{ref} = 0°C$) is used for the reference junction. The purpose of the amplifier is to convert the transducer output voltage (e.g., 0–4.2 mV) into the range of the ADC converter (e.g., 0–5 V). A 10-bit ADC will provide the desired 0.1°C measurement resolution.

If the temperature range is less than 25°C, then linear approximation can be used to measure temperature. Let N be the digital sample from the ADC converter for the unknown tissue temperature, T_1. A calibration is performed under the conditions of a constant reference temperature: typically, one uses the extremes of the temperature range (T_{min} and T_{max}). A precision thermometer system is used to measure the "truth" temperatures. Let N_{min} and N_{max} be the digital samples at T_{min} and T_{max} respectively. Then the following equation can be used to calculate the unknown tissue temperature from the measured digital sample.

$$T_1 = T_{min} + (N - N_{min}) \cdot \frac{T_{max} - T_{min}}{N_{max} - N_{min}} \tag{11.18}$$

Because the thermocouple response is not exactly linear, errors using the above linear equation will increase as the temperature range increases. For instruments with a larger temperature range, a quadratic equation can be used,

$$T_1 = H_0 + H_1 \cdot N + H_2 \cdot N^2 \tag{11.19}$$

where H_0, H_1 and H_2 are determined by calibration of the instrument over the range of interest.

Table 11.3 Temperature sensitivity and range of various thermocouples [12]

Type	Thermocouple	Sensitivity μV/°C at 37°C	Useful range °C
Type T	Copper/Constantan	42	−150 to +350
Type J	Iron/Constantan	53	−150 to +1000
Type K	Chromel/Alumel	41	−200 to +1200
Type E	Chromel/Constantan	62	−200 to +900
Type S	Platinum/Platinum-rhodium	6	0 to +1500

11.4.4 Analysis of Thermocouple-Specific Errors

11.4.4.1 Nonhomogeneous Wires

A homogeneous thermocouple wire is one which has constant thermoelectric properties. Temperature gradients along a homogeneous thermocouple wire will not affect the voltage output, as long as the junction temperatures are not affected. See Fig. 11.8. If Metal A is homogeneous, the voltage output, V_{out}, will be independent of $T3$.

On the other hand, impurities in Metal A together with a temperature gradient along the wire will cause an error in the temperature measurement, as shown in Fig. 11.9. The magnitude of the error will be a function of the impurity and the temperature difference, $T_3 - T_4$. Let a' be the effective temperature sensitivity between the pure Metal A and the impure Metal A′. Thus:

$$\text{voltage error} = a' \cdot (T_3 - T_4) \tag{11.20}$$

This error can be minimized by constructing thermocouples with special high purity metals [12].

11.4.4.2 Cable Connections

If Metal C (the cable) is introduced into the circuit, then two more junctions are created. Fig. 11.10 shows that the temperatures at these two new junctions are T_3

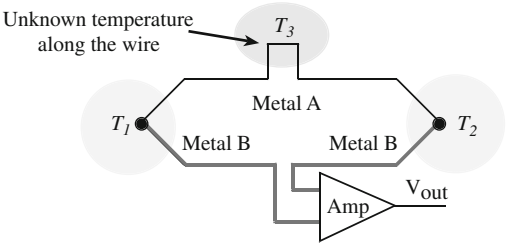


Fig. 11.8 If the thermocouple wire is homogeneous, then the temperature along the wire has no effect on the voltage output

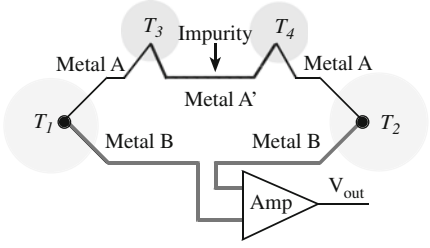


Fig. 11.9 The effect of a temperature gradient along a nonhomogeneous thermocouple wire

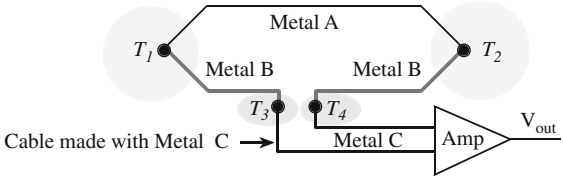


Fig. 11.10 The effect of unknown temperatures at the connection of the thermocouple wire to the instrument cable

and T_4. Let a_{BC} be the effective temperature sensitivity between the Metal B and the Metal C. The voltage error is:

$$\text{voltage error} = a_{\text{BC}} \cdot (T_3 - T_4) \tag{11.21}$$

If the BC and CB junctions are at the same temperature ($T_3 = T_4$), then there is no net effect on the output, and the voltage V_{out} is independent of the properties of Metal C and temperature T_3. This allows copper wires to be used in thermocouple circuits even if neither Metal A nor Metal B is copper. The temperature difference ($T_3 - T_4$) can be reduced by minimizing the physical distance between the BC CB junctions and insulating the two junctions from the environment.

11.4.4.3 Thermocouple Reference

The standard thermocouple tables are generated for a reference temperature, T_{ref}, of 0°C. For small temperature ranges, the output of the system varies linearly with the temperature of interest, T_1

$$V_{\text{out}} = a\,(T_1 - T_{\text{ref}}) \tag{11.22}$$

where the sensitivity a includes both the thermocouple response and the gain of the instrumentation amplifier. The sensitivity in Fig. 11.11 varies from 39 μV/°C to 46 μV/°C as the temperature increases from 0 to 100°C, even though the data points quite closely match the straight line approximation. Consequently, if one wishes to use a reference temperature other than 0°C, it will be appropriate to explicitly calibrate the device, rather than relying on the standard calibration curves.

On the other hand, if the reference temperature unknowingly changes, then this drift translates "one to one" into a temperature measurement error. The temperature of ice and water in a bucket can vary up to 3°C if careful procedures are not followed. First, one should fill the container with as much ice as will fit, adding more ice as melting occurs. Second, cracked ice and water should be used because they produce fewer spatial temperature gradients than cubed ice. Third, one should insulate all walls of the container so as to reduce melting and decrease spatial temperature gradients within the container. Some reference junctions are mounted in copper blocks. Fourth, it is important to position the thermocouple reference junction in the center of the ice bucket and not let it rest up against the side of the container.

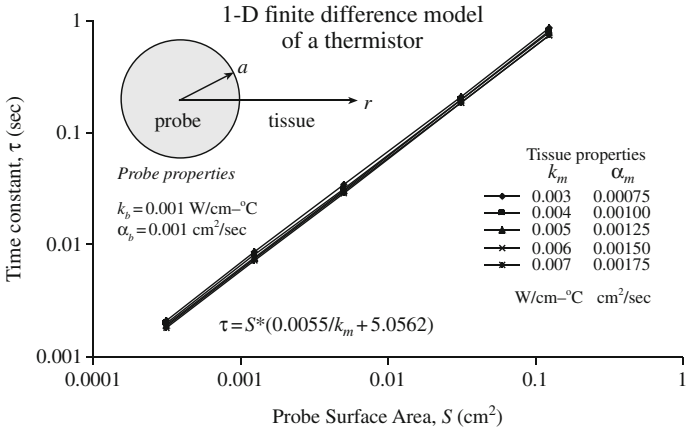


Fig. 11.11 Time constant vs. surface area of a thermistor assuming conduction-dominated heat transfer at $t = 0$, $T_{probe} = 1°C$, $T_{tissue} = 0°C$

We can also employ an electronic reference. In this system a thermistor or RTD probe is placed in close proximity to the thermocouple T_{ref} junction. No ice is required, but we will insulate the thermocouple T_{ref} junction and the thermistor/RTF probe so they will be at the same temperature. The thermistor/RTD probe is used to measure T_{ref} (Fig. 11.4), and the thermocouple system (Fig. 11.7) measures V_{out}. We rearrange Eq. (11.22) to measure T_1

$$T_1 = V_{out}/a + T_{ref} \tag{11.23}$$

11.5 Error Analysis of Probe-Based Transducers

11.5.1 Time Constant

11.5.1.1 Conduction-Dominated Time Constant

If the transducer is placed in a perfect heat sink, its time response will be determined only by the size and thermal properties of the transducer probe. Since this is the situation in which the probe responses are fastest, it is often the parameter quoted by the manufacturer. The only biologic situations in which this assumption can be assumed are the cases where the probe is located in a high flow artery or where the probe is directly heated by external energy sources like lasers and microwaves. Table 11.4 shows that the materials used to construct temperature transducers have a wide range of thermal properties.

If the probe is inserted in tissue then the thermal properties of the tissue will also affect the measurement time constant. In this situation, heat transfer will be dominated by thermal conduction between the temperature probe and the tissue.

Table 11.4 Thermal properties of various materials at 25°C

Material	c, specific heat (J/g-°C)	ρ, density g/cm^3	k, thermal conductivity (W/cm-°C)	α, thermal diffusivity (cm^2/s)
Copper[4]	0.385	8.933	4.01	1.17
Stainless steel 304[4]	0.477	7.900	0.149	0.0395
Stainless steel 316[4]	0.468	8.238	0.134	0.0348
glass[4]	0.75	2.5	0.014	0.0075
J.B. Weld epoxy[1]			0.00590	
tissue[1]	1	1.07	0.005	0.0013
Teflon[4]	1.25	2.2	0.0035	0.0013
5 min clear epoxy[1]			0.00218	
Lexan[3]	1.25	1.2	0.0019	0.0013
NTC Thermistor[2]	1	1	0.001	0.001

[1]Measured with self-heated thermistors [17–19]
[2]Estimated from self-heated thermistor calibration measurements [18]
[3]Manufacturer's specifications
[4]Incropera, *Fundamentals of Heat Transfer*, Wiley (1990)

A simple spherical 1-D finite difference model has been used to investigate the various factors which affect the transient response of a thermal probe. The transducer and tissue were assumed to have spherical shapes and homogeneous thermal properties. The 700 nodes in the probe and tissue were equally spaced. The tissue radius was 7 times larger than the probe radius. No thermal contact resistance, convection, radiation, or blood perfusion was included. The initial temperatures of the probe and tissue were 1 and 0°C respectively. The temperature measured by the device was assumed to be the volume average temperature within the probe. The time constant, τ, was calculated as the time required for the probe to reach $1/e$ of the final temperature. The analysis was performed as a function of thermistor radius ($a = 0.005$, 0.01, 0.02, 0.05, 0.1 cm), and for five tissues ($k_m = 0.003$ $\alpha_m = 0.00075$, $k_m = 0.004$ $\alpha_m 0.001$, $k_m = 0.005$ $\alpha_m = 0.00125$, $k_m = 0.006$ $\alpha_m = 0.0015$, and $k_m = 0.007$ W/cm-°C $\alpha_{\mathrm{m}} = 0.00175\ \mathrm{cm}^2/s$). The results as a function of probe surface area in Fig. 11.11 show the time constant of a thermistor is strongly affected by its surface area, and weakly dependent on tissue thermal properties.

The analysis was repeated for three different probes (copper $k_b = 4.01$, $\alpha_b = 1.17$, stainless steel $k_b = 0.14$, $\alpha_b = 0.0387$, and thermistor $k_b = 0.001$W/cm-°C, $\alpha_\beta = 0.001\mathrm{cm}^2/\mathrm{s}$) and a single tissue. Once again, the time constant is linearly related to probe surface area, as illustrated in Fig. 11.12. Even though the probe thermal diffusivities varied considerably (0.001–1.17 cm^2/s), the time constant is most dependent on probe size and only slightly dependent on tissue thermal properties. Furthermore, for a conduction-dominated environment, the time constant is almost independent of the thermal properties of the probe. In other words, in situations where heat is transferred from the tissue to the probe by thermal conduction, thermistors are almost as fast as thermocouples.

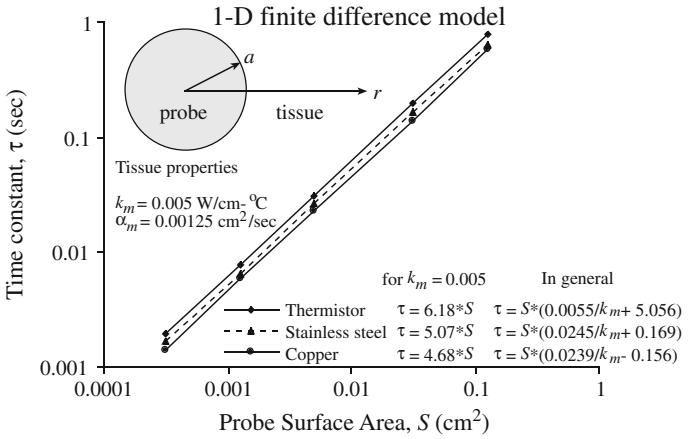


Fig. 11.12 Time constant vs. surface area assuming conduction-dominated heat transfer at $t = 0$, $T_{\text{probe}} = 1°\text{C}$, $T_{\text{tissue}} = 0°\text{C}$

11.5.1.2 Convective Boundary Condition

If the probe is located in an artery, then a convective boundary condition occurs on the probe surface. Additional 1-D finite difference models were used to compare conductive vs. convective heat transfer in the tissue medium. The results plotted in Fig. 11.13 show that the heat transfer in the medium strongly affects the time

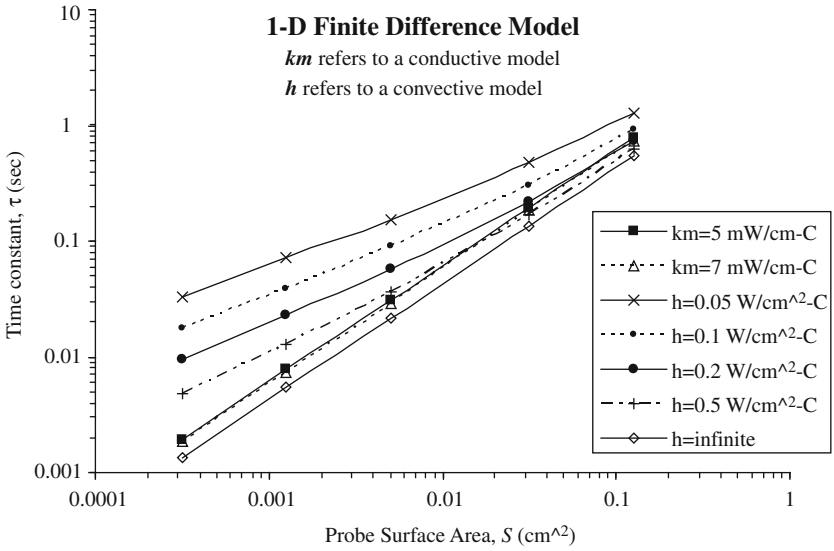


Fig. 11.13 Time constant vs. probe surface area for a thermistor probe at $t = 0$, $T_{\text{probe}} = 1°\text{C}$, $T_{\text{tissue}} = 0°\text{C}$. Convective boundary conditions at radius a of 0°C

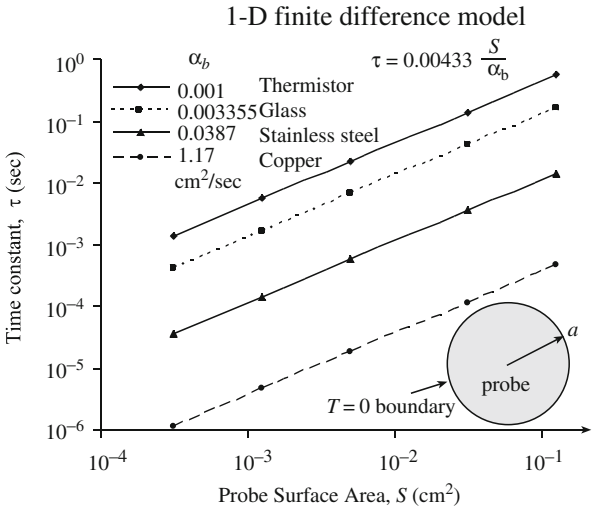


Fig. 11.14 Time constant vs. probe surface area assuming a high flow boundary condition at $t = 0$, $T_{probe} = 1°C$, $T_{surface} = 0°C$

constant of the temperature probe. Again, the transducer was assumed to have a spherical shape and homogeneous thermal properties. 100 nodes were equally spaced in the probe. The initial temperature of the probe was 1°C and a $T = 0°C$ convective boundary condition was created at the probe surface. The time constant, τ, was again calculated as the time required for the volume average probe temperature to reach $1/e$ of the final temperature.

The analysis was repeated for five probe radii and four probe thermal properties assuming an infinite convection coefficient. Figure 11.14 shows for a highly convective medium, the time constant is linearly related to probe surface area, and inversely proportional to the probe thermal diffusivity. The 20 data points in Fig. 11.14 fit the following equation with an average error of 1.3%

$$\tau = 0.00433 \frac{S}{\alpha_b} \tag{11.24}$$

where S is the probe surface area [cm^2] and α_b is the probe thermal diffusivity [cm^2/sec]. If the probe is positioned in a medium in which the heat transfer is dominated by thermal convection, then the probe surface area and probe thermal diffusivity are significant factors in determining the time constant of the transducer.

11.5.1.3 Manufacturer Specifications

Figure 11.15 plots the time constant vs. probe surface area for four unsheathed fine gauge thermocouples manufactured by Omega [12] and 23 thermistors of various

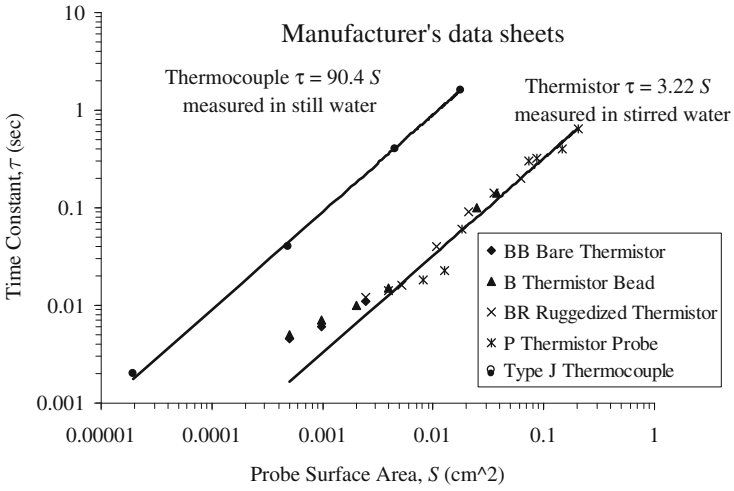


Fig. 11.15 Time constant as specified by Omega and Thermometrics vs. probe surface area

types manufactured by Thermometrics [13, 14]. This data supports the fact that time constant varies linearly with probe surface area. As shown by the numerical analyses, time constant strongly depends on the experimental conditions in the medium. The data in Fig. 11.15, however, do not prove thermistors are faster than thermocouples. Rather, the data illustrate that Omega and Thermometrics do not measure time constants in the same manner. In fact, Omega measured time constant in still water, while Thermometrics measured it in stirred water. It is good practice to measure (or model) time constants in the same environment in which the devices will be ultimately used. The purpose of manufacturer's specifications is only to compare the various sizes and types from that company.

11.5.2 Measurement Errors in the Presence of Spatial Tissue Temperature Gradients

A 2-D axisymmetric steady state finite element model was used to study how the presence of the probe itself disturbs the tissue temperature field. Figure 11.16 shows the geometry and boundary conditions of the first numerical model. The 2-D model was rotated about the z-axis to simulate a 3-D cylindrical probe (with radius a and depth 0.5 cm) in a cylindrical tissue (with a radius ten times the probe radius). In this first model, the probe sensor, probe shaft, and wire leads had uniform thermal properties. A tissue temperature gradient of 6.25°C/cm was created by the $T = 0$ and $T = 10$ boundary conditions at the top ($z = 1.6$ cm) and bottom ($z = 0$). An insulating boundary condition was applied to the sides ($r = 0$ and $r = 1$ cm), creating a one dimensional spatial temperature gradient. If the probe thermal conductivity equaled

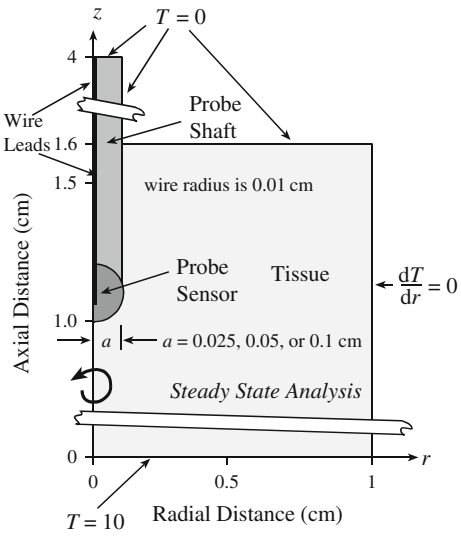


Fig. 11.16 The geometry and boundary conditions for the 2-D axi-symmetric finite element model that is used to analyze 1-D temperature gradients

the tissue conductivity, then the presence of the probe would not affect the tissue temperature field. It was assumed that the temperature measured by the device will be the volume average temperature at the tip of the probe (labeled "probe sensor" in Fig. 11.16). The temperature error is the difference between the average temperature at the probe tip and the average calculated when k_b equals k_m. Clearly, the size of the probe significantly affects its transient response [11–13, 20]. On the other hand, calculations for three probe sizes ($a = 0.025, 0.05, 0.1$ cm) showed that the steady state temperature errors in this situation were independent of probe radius. This fact can be explained by considering the heat flux boundary condition at the probe/tissue interface.

$$k_b \frac{\partial T}{\partial r} = k_m \frac{\partial T}{\partial r} \quad \text{at } r = a \tag{11.25}$$

There is no explicit dependence on probe radius, a, in this situation. If the local tissue temperature gradient is linear in the direction along the probe shaft, then the error is independent of probe diameter.

The steady state temperature in a slab 1.6 cm thick with boundary temperatures of 0°C ($z = 1.6$) and 10°C at $z = 0$ is a straight line as illustrated in Fig. 11.17. When a probe is placed in the geometry, heat transfer between the probe and tissue alters the temperature field as illustrated in Fig. 11.17 where centerline temperature, $T(0, z)$, for five probe thermal conductivities is plotted. For the case of the thermistor, the probe's lower thermal conductivity causes the tip of the probe to slightly overestimate the tissue temperature. On the other hand, the high thermal conductivity of the

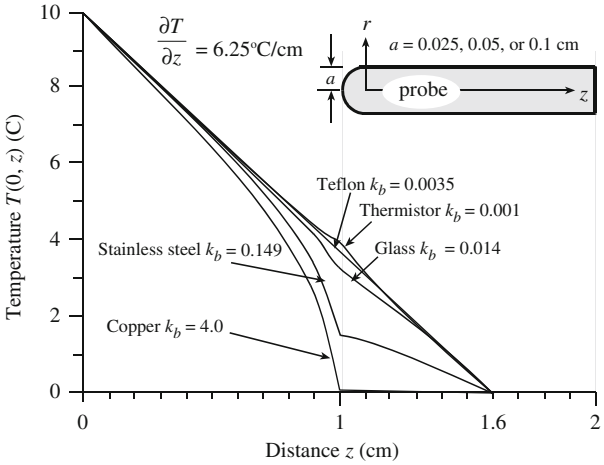


Fig. 11.17 The steady state centerline temperature for five different probes. The probe radius has no effect on this data

copper drove the entire thermocouple to the T=0 boundary temperature. A significant temperature error occurs with the metal probes. The temperature measurement errors of the five numerical studies can be expressed as:

$$T_{\text{error}} = \left(-0.047 - 0.065 \cdot \ln \frac{k_b}{k_m}\right) \cdot \frac{\partial T}{\partial z} \tag{11.26}$$

The analyses were repeated for other boundary conditions at the ($z = 1.6$ cm) top surface of the model (convective, constant $\partial T/\partial z$, and constant heat flux boundary conditions). In each case, large errors occur with the metal probes. These results strongly discourage the use of metal probe shafts because the presence of the probe itself will significantly alter the temperature field. Unfortunately, the largest temperature error occurs at the measurement sensitive tip of the probe. To minimize this error one should construct the probe with a shaft that has a thermal conductivity the most similar to tissue.

Two very striking conclusions arise from this work. The first observation is that stainless steel shafts cause significant temperature measurement errors. The magnitude of this error increases with the tissue temperature gradient ($\partial T/\partial z$) and mismatch in probe/tissue thermal conductivity ($\ln k_b/k_m$). If the tissue temperature gradient exists only along the probe shaft, then the measurement error is independent of shaft diameter. The measurement error can be minimized by using a shaft with thermal properties similar to tissue. Teflon is a good choice. This work shows that thermistors have a smaller perturbing effect than thermocouples. The second way that errors can be minimized is to align the probe shaft along an isotherm. In this way heat will not conduct along the shaft, and errors will not develop at the probe tip.

11.5.3 Surface Measurement Errors

The measurement of temperature at the surface of a tissue is quite difficult when using contact transducers like thermistors and thermocouples. The temperature of the tissue surface is strongly dependent on the delicate balance of the various heat transfer modalities occurring at the surface. The convective heat transfer is dependent on the air temperature and the air flow pattern. The presence of the probe may alter the air flow pattern at the tissue surface. The evaporative component of heat transfer (sweating) can be significant for human skin and the presence of the probe will significantly affect evaporation at the probe site. The heat transfer due to thermal radiation depends on the emissivity of the surface. A perturbation will occur if the emissivity of the transducer surface does not match the value of tissue. Surface measurements are difficult to obtain because in many biological situations significant spatial temperature gradients occur at the tissue surface. In addition, surface measurement probes are subject to the same types of errors that were studied in the previous section.

11.5.4 Direct Absorption of External Energy into the Probe

When external energy is applied to the tissue for diagnostic or therapeutic purposes, it is possible for this energy to be directly deposited into the transducer. Both ultrasonic and laser energy will directly heat thermocouples and thermistors. Figure 11.18 plots the temperature measured by a microthermocouple placed in a rabbit eye vitreous during 12-ms argon laser pulse [20]. There is an instantaneous

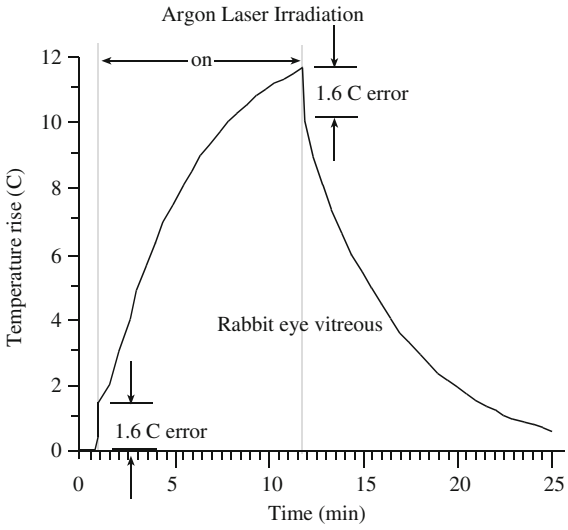


Fig. 11.18 The temperature vs. time measured with a microthermocouple during argon laser heating [20]

1.6°C temperature rise (and fall) at the beginning (and end) of the laser pulse. This is due to direct absorption of the laser energy into the thermocouple. This error always produces an overestimation of the tissue temperature. The magnitude of this error can be experimentally determined using an experimental protocol similar to that of Fig. 11.18.

11.6 Infrared Temperature Measurement

Thermographic imaging offers an alternative non-contact method of estimating surface temperatures with extremely high spatial resolution. Thermal cameras measure surface energy in the infrared band, which is due primarily to molecular and atomic agitation in the material. Surface temperature is then inferred from the radiation measurements. There are several confounding effects which can result in incorrect temperature estimation. Some attention to these will ensure that reliable estimates of surface temperature are obtained from thermal infrared imagery.

All electromagnetic (EM) radiation, including thermal infrared emission, comes from the acceleration of charge. In radio antennas, the charges are electrons pumped back and forth in the antenna elements by the transmitter. In visible light optical emission, the light photons are emitted as charge carriers relax (decelerate) from an excited, higher energy state to a lower level energy state. The photon wavelength, and thus its energy, is determined by the energy difference between the states. The measured signal in thermal imaging comes from photons emitted by the material under observation due to the relaxation of thermally excited atoms and molecules. That is, the internal thermal energy contained in the various vibrational, rotational and translational degrees of freedom results in thermal generation and recombination processes upon collision; photon emission is associated with these momentum exchange events. Electromagnetic radiation has both a wave nature (wavelength, frequency, interference, and harmonic expression) and a particle nature (photon emission, momentum, scattering and absorption); these two descriptions will be used interchangeably in this discussion.

11.6.1 Electromagnetic Radiation Spectrum

The electromagnetic spectrum describes radiated electromagnetic energy in terms of its relative wavelength or frequency. Assuming propagation in free space (vacuum or, approximately, air at standard temperature and pressure) with electric permittivity $\varepsilon_0 = 8.85 \times 10^{-12}[\mathrm{F \cdot m^{-1}}]$, and magnetic permeability $\mu_0 = 1.26 \times 10^{-6}[\mathrm{H \cdot m^{-1}}]$, the speed of propagation is $c = 3.0 \times 10^{8}[\mathrm{m \cdot s^{-1}}]$. For a wave in a non-dispersive medium, the wavelength, λ (m), and photon energy, E [J] are:

$$\lambda = \frac{c}{f} \quad \text{and} \quad E = hf \tag{11.27}$$

Table 11.5 Electromagnetic energy spectrum

Signal	Frequency	Wavelength	Photon Energy (J)
A-M Radio	500–1600 kHz	188–600 m	$3.3–10.6 \times 10^{-28}$
F-M Radio	88–108 MHz	2.8–3.4 m	$5.8–7.2 \times 10^{-24}$
Microwave Oven	2.45 GHz	12.2 cm	1.62×10^{-24}
HF Radar	40 GHz	7.5 mm	26.5×10^{-24}
Thermal Imaging	$25–37.5 \times 10^{12}$	8–12 μm	$0.17–0.25 \times 10^{-19}$
	$60–100 \times 10^{12}$	3–5 μm	$0.4–0.66 \times 10^{-19}$
Nd:YAG Laser	283×10^{12}	1.06 μm	1.9×10^{-19}
Visible Light	$0.43–1.0 \times 10^{15}$	300–700 nm	$2.8–6.6 \times 10^{-19}$
Excimer Laser	$1.18–1.55 \times 10^{15}$	193–255 nm	$7.8–10.3 \times 10^{-19}$

where f is the frequency [Hz] and h is Planck's constant (6.63×10^{-34}[J-s]). The lower energy photons have the longer wavelengths. Table 11.5 illustrates the relationships among these parameters for familiar types of EM radiation.

11.6.2 Planck Radiation Law

There was a considerable controversy surrounding the origins of the radiated flux of an ideal, or "black body," radiator at the close of the 19th century. Classical mechanics formulations suggested that the emissive power from internal thermal energy modes should increase without bound as the wavelength decreased; the so-called "ultraviolet catastrophe." Plainly, the total radiated power from a surface must be finite, and measurements indicated that a peak in radiation was encountered at a specific wavelength, λ_{max}. The wavelength of peak emission is found from Wien's displacement Law:

$$\lambda_{max} T = 2989 \mu m^{\circ}K \quad \text{or} \quad \lambda_{max} T = 5216 \mu m^{\circ}R \tag{11.28}$$

where the temperatures, T, are in absolute degrees [K or R]. Below λ_{max} the monochromatic emissive power decreases. In Fig. 11.19 the monochromatic emissive power is plotted as a function of wavelength.

The figure is a plot of the Planck monochromatic radiation law (Eq. (11.29)), which very closely matches measured values. The peak emission wavelengths, from Wien's displacement law, are also shown in the figure. Note that the wavelength at peak emission decreases as the surface temperature increases. Our sun, for example, is a black body at approximately 6500 [K] (λ_{max} about 445 nm).

$$W_b(\lambda, T) = \frac{2\pi h c^2 \lambda^{-5}}{\left(e^{\frac{hc}{\lambda kT}} - 1\right)} \tag{11.29}$$

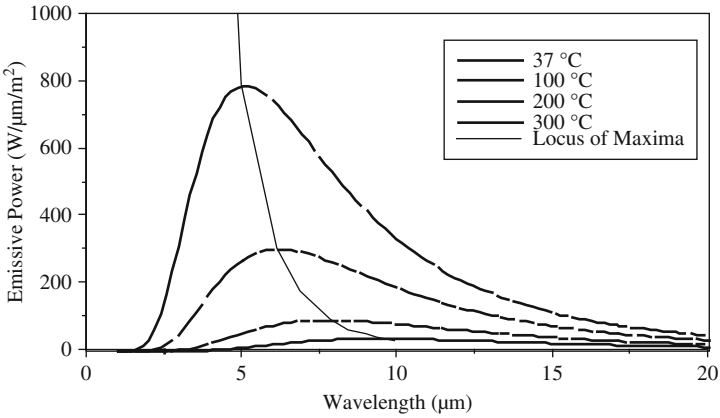


Fig. 11.19 Monochromatic emissive power $[W \cdot \mathrm{m}^{-2} \cdot \mu\mathrm{m}^{-1}]$ and wavelengths of maximum emissive power

where W_b is the black body (ideal) monochromatic emissive power $[W \cdot \mathrm{m}^{-2} \cdot \mu\mathrm{m}^{-1}]$, and $k =$ Boltzmann's constant $(1.38 \times 10^{-23}[\mathrm{J} \cdot \mathrm{K}^{-1}])$.

11.6.3 Wide Band Radiation Heat Transfer

The total emissive power for a black surface may be calculated by integrating the Planck radiation law over all wavelengths:

$$E_b(T) = \int_0^\infty W_b(\lambda, T) d\lambda = \frac{2\pi^5 k^4}{15c^2 h^3} T^4 = \sigma T^4 \tag{11.30}$$

where $\sigma = 5.67 \times 10^{-8} \mathrm{W} \cdot \mathrm{m}^{-2} \mathrm{K}^{-4}$ is the Stefan-Boltzmann constant. Equation (11.30) is the Stefan-Boltzmann law of thermal radiation from a black surface.

11.6.4 Band-Limited Radiation

Thermal imaging devices are necessarily band-limited in response by the spectral sensitivity of the photo detectors used. The two most common imaging bands for measuring emitted infrared energy are 3–5 μm and 8–12 μm. These limits are imposed by band gaps in useful photo detectors and also to avoid atmospheric absorption between 5 and 8 μm. The radiated emissive power in those two bands may be quickly calculated from Eq. (11.29) over the specific bands. Figure 11.20 illustrates the relative black body power distribution between the two fixed imaging bands as a function of temperature.

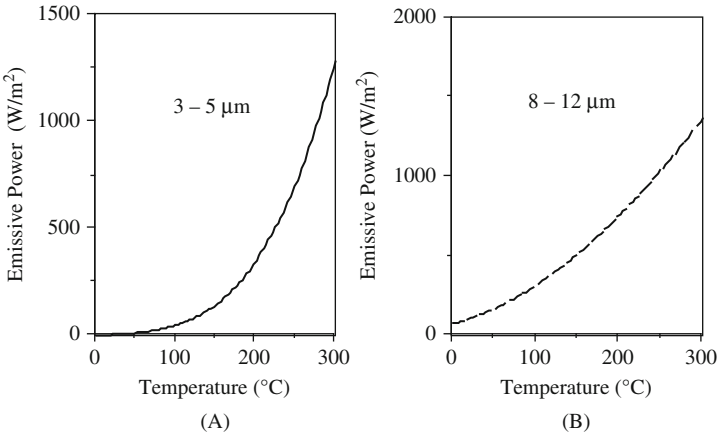


Fig. 11.20 Band-limited emissive power for the two common imaging bands in $W \cdot m^{-2}$: (**a**) 3–5 μm and (**b**) 8–12 μm

The functional relationship between temperature and emissive power can be easily expressed as fifth order polynomial fits to the curves in the figure ($r^2 = 1.00$ in each case). For the 3–5 μm band the data may be modeled by:

$$E_b(T) = 1.987 + 8.257 \cdot 10^{-2} T + 1.1817 \cdot 10^{-3} T^2 + 9.232 \cdot 10^{-6}\, T^3 + 1.3709 \cdot 10^{-7} T^4 + 1.071 \cdot 10^{-10} T^6 \tag{11.31}$$

where T is in K. For the 8−12 μm band:

$$E_b(T) = 74.296 + 1.4603\, T + 8.933 \cdot 10^{-3} T^2 + 1.4858 \cdot 10^{-5}\, T^3 - 7.0592 \cdot 10^{-8} + 9.286 \cdot 10^{-11}\, T^5 \tag{11.32}$$

where T is also in K. It is interesting to note the approximate linearity of the 8–12 μm band compared to the 3–5 μm band. Also, the 8–12 μm band has more signal at lower temperatures: approximately 26% of the wide-band emissive power (Eq. (11.49)) compared to about 1% for the 3–5 μm band at 37°C. However, the noise figure in typical 8–12 μm band thermal detectors is much higher than for 3–5 μm band detectors; so the signal to noise ratio is about the same in comparable devices operating over these bands.

11.6.5 Making a Black Body

Though an ideal black body is impossible to make, one can come quite close by careful design. Figure 11.21 shows a close approximation: an arbitrarily shaped cavity with an infinitesimally small aperture through which the wide band emissive power $E_b(T_w)$ may be observed. The operating assumption is that the walls are

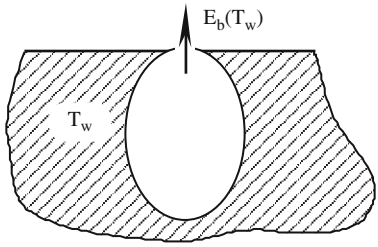


Fig. 11.21 Ideal black body radiator is a nearly totally enclosed cavity with an isothermal wall at temperature T_w

isothermal at temperature T_w. Then, in terms of radiation heat transfer, the shape factors and view factors simplify to 1. That is, a small segment of wall surface is in thermal steady state at T_w and is exchanging heat by radiation with other cavity wall segments (with the exception of the vanishingly small aperture) at T_w, and so is in thermal equilibrium with them. Under these conditions, the radiative flux escaping the aperture will approach the black body curve as the aperture area approaches a negligible fraction of the cavity surface area.

Creating a sufficiently accurate ideal aperture requires a very large volume device. Useful close approximations can be obtained by using approximate designs with interior walls coated with high emissivity paint to make them approach black radiators. In Fig. 11.22 two useful approximations are shown. A large area black body with a milled cavity surface is shown in part A and a deep drilled cylindrical cavity in B. The large surface in part A consists of a crossed pattern of 30° (included angle) slots milled to leave pyramidal shapes in the surface of a large copper cylinder. The 30° included angle is, in some sense, an optimal compromise between black body volume and surface patch view factors. That is, a differential patch on the side of one of the pyramids "sees" mostly other surfaces at the same temperature, rather than the environment. Thus, its resultant emissive power approaches that of a black surface. The block is heated by heater wires wrapped around the cylinder or embedded within it. Control and surface temperature indication are obtained from an embedded thermistor or thermocouple just below the surface. The large surface

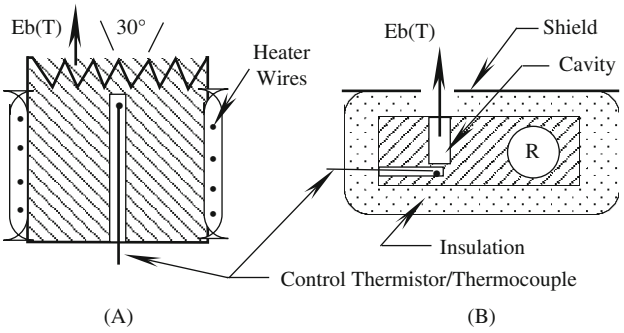


Fig. 11.22 Two practical "black body" designs

design of part A is not suitable for small laser spot experiments owing, primarily, to its height. The low profile black body, shown in cross-section in part B of the figure, is more easily used in laser spot experiments. The small block is heated by a quarter watt or eighth watt resistor, R in the figure, which gives an idea of the scale of the copper block. The black body is a drilled cavity coated with high emissivity paint – we use Floquil "Engine Black," a carbon black pigment paint available at model railroad stores. In fact, any paint which has carbon black as the pigment will do as well. The diameter of the cavity must be larger than the instantaneous field of view of the detector (see discussion below) to obtain a useful black body source. Insulation and a radiation shield (uncoated polished metal) are essential to protect tissues from damage and to provide a clear indication of the location of the black body in the thermal image.

11.6.6 Gray Bodies

In general, real surfaces emit only a portion of the ideal black surface radiation. The relative efficiency of a radiator is called the emissivity, ε range: $0 \leq \varepsilon \leq 1$, a function of wavelength. Consequently, the monochromatic emissive power is less than that of a black surface:

$$W(\lambda, T) = \varepsilon(\lambda)\, W_b(\lambda, T) \tag{11.33}$$

Many interesting surfaces can be characterized as approximately "gray"; that is, $\varepsilon(\lambda) = \varepsilon$, a constant. For those surfaces the emissive power is simply related to the equivalent black body:

$$E(T) = \varepsilon E_b(T) = \varepsilon\, \sigma\, T^4 \tag{11.34}$$

where Eq. (11.34) has been written for the wide band case. For a band-limited view of a gray surface, one needs simply to multiply the polynomials of Eqs. (11.31) and (11.32) by ε. While the gray body assumption may be questionable in the wide band case, it is much more realistic in the band-limited case for thermal imaging. The surface need only have constant emissivity over the imaging bandwidth. Consequently, nearly all calibration schemes tacitly assume a gray surface is viewed by the thermal camera.

11.6.7 Image Formation Effects

Thermal imaging devices measure the surface radiosity, $J(x, y)[\mathrm{W} \cdot \mathrm{m}^{-2}]$, which consists of transmitted, reflected and emitted flux. Incident flux may, likewise, be transmitted, reflected or absorbed. For purposes of analysis one may define a reflectivity, r, an absorptivity, A, and transmissivity, t , which are the respective fractions of the incident flux. Kirchoff's radiation law states that for an incident flux the sum of the absorbed, reflected and transmitted fractions must be 1:

$$A + r + t = 1 \tag{11.35}$$

When viewing a gray surface it must be borne in mind that the radiosity signal consists of emitted, reflected and (perhaps) transmitted components. The surface temperature field, $T(x, y)$, is subsequently inferred from measurements of the radiosity, $J_m(x, y)$. If the surface emissivity is known and the surface emittance of an equivalent black body at temperature T is estimated, then Eqs. (11.31) or (11.32) can be used to estimate T from $E_b(T)$. The reflected and transmitted portions must be removed during the calibration step. Assuming an opaque surface ($t = 0$), a complete and isothermal enclosure at temperature T_∞, and simple shape factor and view factor relations, the emitted flux for an equivalent black body, $E_b(T)$ may be estimated from:

$$E_b(T) = \frac{J_m(T, T_\infty, \varepsilon) \; - \; (1 - \varepsilon)E_b(T_\infty)}{\varepsilon} \tag{11.36}$$

where for an opaque surface $r = 1 - \varepsilon$ and J_m is the measured radiosity (thermal camera result). Equation (11.36) is only practical when the surface emissivity is above about 0.5, since division of J_m by a small ε value increases the uncertainty in E_b due to photo detector noise.

For the usual laser irradiation situation (where the enclosure is cooler than tissue temperatures) all of the errors in thermal imaging contribute to an underestimation of the true surface emitted flux component, $\varepsilon E_b(T)$, and thus an underestimate of surface temperature is obtained. In addition to the nonlinear response (Eqs. (11.31) and (11.32) and to the reflectance errors described above, an underestimate of J_m can also be obtained due to: (1) optical pathway attenuation (lenses and mirrors), (2) thermal gradients along the optical axis, and (3) an imaging spot size too large to resolve the laser spot profile. These errors occur when the detector is scanned (as in a flying spot 2-D imaging device) as well as when it is used as a staring sensor with an optical fiber.

11.6.7.1 Optical Pathway Attenuation

Errors due to optical pathway attenuation may be simply managed by including black body reference sources at known temperatures, T_1 and T_2, in the thermal scene at the level of the experimental surface. Then the camera response to the reference sources may be used to calibrate the video scale since $E_b(T_1)$ and $E_b(T_2)$ are easily obtainable from Eqs. (11.31) or (11.32) – note that photo detectors have an approximately linear relationship between received power ($\int J_m dA$) and output signal, a change in either the photo detector current or resistance. If the emissivity of the surface in view is known, then Eq. (11.36) can be used to estimate $E_b(T(x, y))$ from $J_m(x, y)$, and Eqs. (11.31) or (11.32) inverted to estimate $T(x, y)$.

11.6.7.2 Optical Axis Thermal Gradients

The observed radiosity, J_m, is largely a surface phenomenon; however, the definition of surface is often not as straight forward as it might at first seem. When imaging ordinary surfaces under conduction and convection heating conditions, the thermal gradient normal to the surface is not usually significant compared to the viewing depth of the photo detector. That is, the radiative flux emitted from the surface actually comes from volume emission, $P(T, z)[W \cdot m^{-3}]$, at some depth to the surface. Imagine an isotropic volume emitter whose emission is attenuated primarily by absorption as it propagates toward the surface as shown in Fig. 11.23.

For the 3–5 μm and 8–12 μm imaging wavelength bands, λ_i, scattering processes in tissue may be neglected in favor of absorption. Then, a thermal field which is isothermal over a volume deeper than the optical depth of the imaging band and with a surface footprint much larger in diameter than the photo detector field of view has $P(T, z) = P(T)$. The measured surface emitted flux, $E_m(T)$ is the sum of all embedded slabs at depth "z" and a one-dimensional analysis based on Beer's law absorption may be used:

$$E_m(T) \int_0^\infty P(T, z) \ \exp[-\mu_a(\lambda_i)z]dz = \frac{P(T)}{\mu_a(l_i)} \tag{11.37}$$

where $\mu_a(\lambda_i)$ is the average absorption coefficient of the thermal radiation $[m^{-1}]$ in the imaging wavelength band.

Under the condition that the temperature is constant over the viewing depth of the camera, the standard calibration methods previously outlined will provide acceptable estimates of the surface temperature. If, however, there is a significant thermal gradient along the optical axis (Z-direction), then the true surface temperature may be significantly underestimated. For example, assume that a high absorption coefficient laser, such as 193 μm excimer or 2.93 μm Er:YAG laser, creates a temperature profile which is invariant over the imaging spot field of view (X- and Y-directions) but varies significantly in the Z-direction:

$$P(T, x, y, z) = P(T, z) = P(T, 0)e^{-\beta z} \tag{11.38}$$

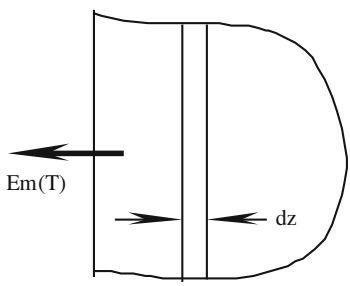


Fig. 11.23 Volume radiation from differential layers of thickness dz propagates toward the surface to form the total measured emitted flux, $E_m(T)$

where β is the decay constant of the temperature profile [m^{-1}]. Then, if the goal is to estimate the surface temperature (at $z = 0$) from the measured surface emittance, E_m must be corrected:

$$E_m(T) = \int_0^\infty P(T,z)\ \exp[-\mu_a(\lambda_i)z]dz = \int_0^\infty P(T,0)\ \exp\{-[\mu_a(\lambda_i)+\beta]z\}dz$$
$$= \frac{P(T,0)}{\mu_a\,(\lambda_i)\ +\ \beta} \tag{11.39}$$

If no temperature gradient existed in the Z-direction, an equivalent constant temperature emittance, $E_c(T)$, would be measured from the surface temperature, T. This equivalent constant temperature emittance is simply related to $P(T, 0)$ and can be used in Eqs. (11.30) or (11.31) to estimate the true surface temperature:

$$E_c(T) = \frac{P(T,0)}{\mu_a\,(\lambda_i)} = \frac{\mu_a(\lambda_i)+\beta}{\mu_a\,(\lambda_i)}E_m(T) \tag{11.40}$$

Note that for many laser types the correction coefficient, $(\mu_a(\lambda_i) + \beta)/\mu_a(\lambda_i)$, may be quite large and is applied to the emittance (and refers to absolute zero temperature); see Fig. 11.20.

At the onset of laser irradiation, before heat transfer processes smear the temperature profile, β will equal the tissue absorption coefficient $\mu_a(\lambda_1)$ at the laser wavelength, λ_1. For example, in water the absorption coefficient for CO_2 laser light (at 10.6 μm wavelength) is approximately 7.92×10^4 [m^{-1}] (a depth of penetration of about 12.6 μm). This means that an 8–12 μm thermal camera gets approximately 86% of its signal from the 25 μm just below the surface of the water (and a very similar depth in tissue). Excimer lasers have depths of penetration between approximately 1 and 30 μm depending on wavelength. If an 8–12 μm detector was used to image a temperature profile with an absorption coefficient in that range (i.e., $33.3 \times 10^3 < \mu_a < 1.0 \times 10^6\ m^{-1}$), the correction factor applied to E_m would lie somewhere between 1.4 and 13.6. An uncorrected surface emittance severely underestimates the actual surface temperature. By comparison, deeply penetrating lasers, such as diode or Nd:YAG, have negligible correction factors – near 1.04 and 1.003, respectively, depending on tissue optics and absorption.

Axial thermal conduction modifies the temperature profile expression. The time at which heat transfer significantly changes these correction factors depends critically on the geometry of the laser spot. A 1 μm penetration depth laser spot differs significantly from the simple exponential decay of Eq. (11.38) after about 1 μs, while a 10 μm depth of penetration wavelength decays with a time constant of about 0.1 ms. As time progresses, the axial temperature gradient becomes small and $\beta << \mu_a(\lambda_i)$.

11.6.7.3 Image Spot Size Effects

The photo detector has a finite acceptance angle, called the Instantaneous Field Of View, IFOV. Large IFOV angles are equivalent to large numerical apertures in microscopy. The IFOV is defined in Eq. (11.41) and illustrated in Fig. 11.24. Typical photo detectors in thermal cameras are used as "flying spot" scanners to generate a video raster scan from a single detector; usually either Indium Antimonide (In–Sb) or Mercury–Cadmium–Telluride (Hg–Cd–Te). The IFOV is scanned along a line by either a rotating prism (Agema device) or oscillating mirror (Inframetrics). Two alternate technologies are the CCD (Charge Coupled Device) imager (by Hughes) and the pyro-electric converter (by Thomsen-CSF). The CCD imager uses multiple linear arrays of staring photo detectors. The received signal is sequentially interrogated from each line of detectors by sliding the signal from detector to detector with a clock signal – it works much like an old fashioned fire fighting bucket brigade and sometimes these devices are called BBDs for bucket brigade devices. Because of the deleterious effects of thermal recombination and generation noise processes, all of the detectors, except the pyro-electric device, must be cooled to at least liquid nitrogen temperatures (77 K) and superior thermal images are obtained at liquid helium temperatures (4 K). The single room temperature device, the pyro-electric converter, exposes a thermally sensitive target to incoming infrared flux and interrogates the target with an electron beam, as in a television vidicon tube. The pyro-electric vidicon is not able to respond to constant signals, however, and a DC infrared signal must be chopped in order to form an image.

$$\mathrm{IFOV} = \arctan\{\mathrm{d/D}\} \tag{11.41}$$

High spatial resolution images are obtained from devices with a small IFOV and a fast response (large band width). The band width required of image processing electronics is determined by the line rate and number of samples (pixels) in a line:

$$\mathrm{BW} = \frac{\text{lines}}{\text{second}} \cdot \frac{\text{pixels}}{\text{line}} \tag{11.42}$$

For a standard video display (15,750 lines $\cdot\, \mathrm{s}^{-1}$) and 256 pixels/line, the signal bandwidth would be 4.03 MHz and the required minimum digital sampling rate would be 8.06 MHz. Thermal detectors are slower than this – the fastest among them have a band width of about 2 MHz – so thermal imagers collect fewer than the 67,200 samples a 60 Hz video field would require. Competing commercial systems

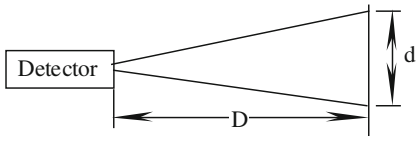


Fig. 11.24 Instantaneous field of view, IFOV, calculation

solve this dilemma in different ways. Some store a line scan of thermal data and read it out more than once in a video field, others put out data at video rates (30 frames/s, 60 fields/s) but only collect 1 or 2 thermal images in a second.

Photo detectors have a weighted response. That is, over an IFOV, more weight is given to information in the center than at the edges. Typical spatial response models use a Gaussian window, w, to describe the IFOV:

$$w(x, y) = \frac{1}{2\pi\sigma^2} \mathrm{e}^{\frac{(x^2+y^2)}{2\sigma^2}} \tag{11.43}$$

where the IFOV diameter $2w_L$is 4σ [m]. A Gaussian weighted IFOV matches experimentally determined point spread functions with good accuracy. For a Gaussian IFOV, the video output signal, $f(x, y)$ [V], may be expressed as:

$$f(x, y) = \text{gain} \cdot \left[\int_{\text{IFOV}} \int w(x - \omega, y - \xi) \left\{ \int_{\lambda_1}^{\lambda_2} S(\lambda) J_m(\lambda, \omega, \xi)\, d\lambda \right\} d\omega d\xi \right] + \text{offset} \tag{11.44}$$

where gain [V/W] and offset [V] describe the video system, w is the window function, S the detector spectral response over the imaging band λ_1 to λ_2, and ω and ξ are the dummy variables of convolution. Thermal gradients within the IFOV are averaged out with weights determined by the window function, w.

A large IFOV can contribute significant error to the estimate of maximum center temperature when the laser spot is small and significant thermal gradients exist within the IFOV. For example, if a surface temperature, T_0, uniform over the IFOV is imaged, the total power received by the detector is: $P_{\text{tot}} = 2\pi\sigma^2 E_0$, where E_0 is the emissive power at T_0.

Next, consider the case of an axisymmetric Gaussian-weighted surface temperature distribution $T(x, \mathrm{y})$, as from a laser spot, with parameter Σ_T such that the surface emissive flux, $E(x, y)[\mathrm{W/m^2}]$ is:

$$E(x, y) = E_0 e^{-\frac{(x^2+y^2)}{2\Sigma^2}} \tag{11.45}$$

where E_0 is the center emissive power $[\mathrm{W} \cdot \mathrm{m}^{-2}]$ due to temperature T_0, the desired center temperature. When this distribution is imaged with a Gaussian distributed flying spot scanner with variance parameter Σ_I, as above, such that the center of the laser spot and the IFOV coincide, it may quickly be shown that the total received power is:

$$P_{\text{tot}} = 2\pi E_0 \frac{\left(\Sigma_I^2\, \Sigma_T^2\right)}{\left(\Sigma_I^2 + \Sigma_T^2\right)} \tag{11.46}$$

A corrected value for the received power may be determined by multiplying the measured power, P_m, by the correction factor $(\Sigma_I^2 + \Sigma_T^2)/\Sigma_T^2$, to obtain the power which would have been measured if the entire IFOV had been at T_0. The value of T_0 can then be estimated as previously described. Significant corrections (> 4%)

Table 11.6 Small spot correction factors

Temperature	Correction
Distribution (Σ_T/Σ_I)	Factor
∞	1
100	1.0001
50	1.0004
20	1.0025
10	1.01
5	1.04
2	1.25
1	2
0.5	5
0.2	52
0.1	101

are necessary for spots with Σ_T less than $5\Sigma_I$ (see Table 11.6). For other window functions, some restoration of blurred small spots can be accomplished by Wiener filtering [21] or by constrained least squares convolution [21]. The effect of a large IFOV is, once again, to underestimate maximum center temperature.

A scanned detector in flying spot imagers has an additional source of error. The slow nature of cooled thermal detectors leads to lag when the detector encounters a sharp temperature rise. The detectors are slow because while cooling beneficially decreases thermal noise by squelching thermal generation-recombination processes, it also increases charge carrier transit times by decreasing mobility. Detector slew rate limitations can lead to significant underestimates of peak temperatures in small laser spots.

These problems can be eliminated with the use of array detectors. However, the cost of these devices is significant.

11.6.8 Summary

Thermal imaging offers unmatched spatial sampling of surface temperature distributions which can be used to directly verify models of laser tissue interactions. There are several significant sources of error common to any application which must be compensated; namely, optical pathway attenuation, surface emissivity and reflected irradiance effects. Laser-tissue applications have several special error sources; optical axis temperature gradients, and small spot size effects. Nevertheless, thermal imaging is a powerful experimental tool which, when carefully applied, provides information otherwise unobtainable.

11.7 Acoustic Measurements of Temperature

Temperature mapping and, therefore, thermal imaging can also be performed using several imaging techniques including magnetic resonance imaging (MRI) [22, 23], ultrasound imaging [24–26], and photoacoustic imaging [27, 28]. All of these

imaging systems are capable of depth-resolved thermal imaging with high spatial resolution. In all approaches, the imaging system is used to measure the temperature induced change in some physical parameter (e.g., speed of sound) and the differential temperature is then reconstructed using a known relationship.

Compared to various methods including MRI [23], microwave radiometry [29], and impedance tomography [30, 31], acoustic methods (e.g., ultrasound imaging and photoacoustic imaging) have several advantages: they are relatively inexpensive, portable, non-invasive and real-time imaging techniques capable of providing instantaneous feedback of laser-tissue interaction. Therefore, the remainder of this chapter will describe thermal imaging based on ultrasound and photoacoustic imaging systems.

11.7.1 Ultrasound-Based Thermal Imaging

11.7.1.1 Ultrasound Imaging

Ultrasound imaging is a widely accepted imaging technique used in the fields of medicine and biology with applications ranging from fundamental science to clinical studies. In ultrasound imaging, a short pulse (typically 2–3 cycles) is transmitted into the body using an ultrasound probe. Today, the ultrasound probe consists of a 1-D or even 2-D array of 128–256 or more ultrasound transducers, although high frequency imaging is often performed using a single element transducer that is mechanically swept over the region of interest. As pressure waves and tissue interact, some of the ultrasound energy is reflected back to the probe which now acts as a receiver and converts the backscattered energy into electric signals. Knowing the time delay between the transmitted ultrasound pulse and received ultrasound echo signals (i.e., using the time-of-flight principle), the 2-D and even 3-D images of the tissue are formed based on the received ultrasound signals.

11.7.1.2 Principles of Ultrasound-Based Thermal Imaging

For reliable monitoring of laser-tissue interaction, it is critical to monitor the temperature distribution in tissues non-invasively and in real time. Several ultrasound parameters have been investigated for temperature estimation including frequency dependent attenuation [32], change in backscattered power [33] and apparent shifts in ultrasound radiofrequency (RF) signals [24–26, 34, 35]. Although all approaches are capable of remote measurements of temperature, here we limit our discussion to a widely used approach utilizing apparent shifts in ultrasound RF signal.

Indeed, using a real time ultrasound imaging system, the temperature change during laser-tissue interaction can be estimated by measuring the thermally induced differential motion of the ultrasound speckle. The diagram in Fig. 11.25 demonstrates the changes in the ultrasound signal with temperature and outlines the procedural steps to estimate the local temperature from the temperature-induced changes in ultrasound signal.

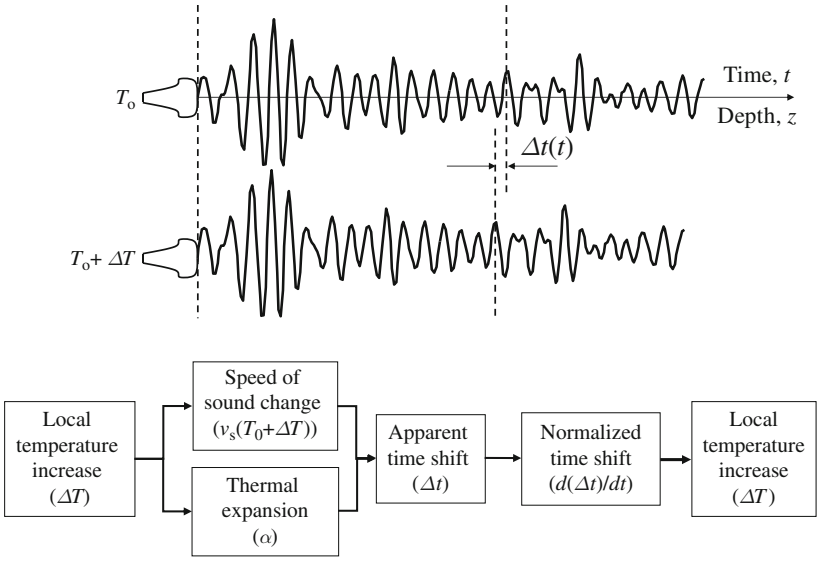


Fig. 11.25 Schematic view of the ultrasound signal change due to temperature, and the block diagram illustrating the principles of ultrasound-based thermal imaging

The time of echo for ultrasound in a homogenous medium is given by

$$t(T_0) = \frac{2 \cdot z}{v_s(T_0)} \tag{11.47}$$

where $t(T_0)$ is the time delay of echo from scatterers at position z at initial temperature T_0 and $v_s(T_0)$ is the speed of sound in the medium at temperature T_0. Note that time t and depth z can be used interchangeably given the relationship in Eq. (11.47). However, since the ultrasound signals are captured as a function of time, the temperature mapping is formulated here using time variable.

When the temperature changes by ΔT, there are apparent time shifts Δt in the ultrasound signal (Fig. 11.25) due to thermal expansion and speed of sound change. The time of echo is now given by

$$t(T_0 + \Delta T) = \frac{2 \cdot z(1 + \alpha \cdot \Delta T)}{v_s(T_0 + \Delta T)} \tag{11.48}$$

where α is the linear coefficient of thermal expansion and $v_s(T_0 + \Delta T)$ is the speed of sound at temperature $T_0 + \Delta T$.

The temperature-induced time shift Δt in ultrasound RF signals is primarily dependent on the speed of sound change while the effects of thermal expansion can be considered negligible for temperatures below 60°C [34, 35]. Note that the change in thermal expansion is not negligible, just that its effect on the time shift is small compared to the effect of speed of sound. Therefore, neglecting $\alpha \cdot \Delta T$, the time shift can be expressed as

$$\Delta t = t(T_0 + \Delta T) - t(T_0) = 2 \cdot z \cdot \left[\frac{1}{v_s(T_0 + \Delta T)} - \frac{1}{v_s(T_0)} \right] \tag{11.49}$$

The apparent time shift depends on the temperature induced change of the speed of sound. Furthermore, the apparent time shift varies with depth – for the same temperature rise, ultrasound RF signals received from deeper structures (larger z) will have greater time shifts. The undesired depth dependency of the apparent time shift can be removed by normalizing the time shifts with respect to depth or time since time and depth are related via Eq. (11.47). Therefore, the temperature elevation is proportional to normalized apparent time shift, i.e., by measuring the apparent time shift Δt in the ultrasound signal, the temperature change ΔT can be estimated using

$$\Delta T \approx k \cdot \frac{d(\Delta t)}{dt} \tag{11.50}$$

where k is a material constant that can be determined experimentally [25], and $d(\Delta t)/dt$, referred to as the normalized time shift, is the gradient of the apparent time shift [25, 26].

Equation (11.50) can be used to describe both the uniform temperature increase in a sample as well as a local and/or spatially varying temperature change ($\Delta T(z)$). To estimate the spatial temperature map during laser heating, multiple ultrasound images are first acquired during the procedure. Generally, ultrasound frames can be acquired at tens and hundreds of frames per second and single ultrasound RF signals can be acquired at tens of thousands of lines per second. Then, the apparent time shift profiles $\Delta t(t)$ or, equivalently, $\Delta t(z)$, are calculated between successive ultrasound images or ultrasound RF signals. Finally, the normalized time shift profiles ($d(\Delta t(z))/dt$) are computed by spatially differentiating the apparent time shifts along the direction of ultrasound beam or ultrasound wave propagation (z). The normalized time shift can now be directly related to the local temperature elevation (Fig. 11.25 and Eq. 11.50).

The normalized apparent time shift ($d(\Delta t(z))/dt$) in Eq. (11.50) can also be viewed as an apparent strain (not to be confused with a true thermal strain due to thermal expansion of tissue; for small temperature elevations such a strain is very small and could be ignored). To estimate apparent strain and, therefore, temperature, strain estimation techniques routinely used in ultrasound-based elasticity imaging can be used [36]. Indeed, the apparent time shift (or, generally, the displacement) between successive frames can be calculated using time-delay estimation techniques or motion tracking block matching algorithms similar to those adapted in elasticity imaging [37]. Overall, the normalized time shifts and axial strains are equivalent since both are obtained from the gradient changes in either time shifts or displacements.

The schematic view of the experimental setup for ultrasound thermal imaging is presented in Fig. 11.26. Here a continuous wave or pulsed laser is used to heat the tissue, and ultrasound imaging system, capable of capturing ultrasound frames in

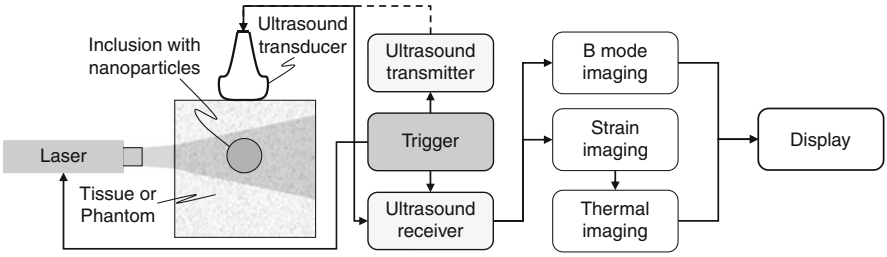


Fig. 11.26 Setup for ultrasound thermal imaging of laser-tissue interaction

real time and computing the normalized time shift (or, equivalently, strain), is used to measure the temperature maps. Generally, the apertures of the ultrasound transducer and laser beam can be separate (as shown in Fig. 11.26) or combined. Consequently, the light may be directed to tissue using an air beam or through the optical fibers. Furthermore, both light delivery and ultrasound imaging can be applied from the outside (Fig. 11.26) or endoscopically (e.g., intravascular application). It also must be noted that ultrasound imaging requires direct contact or liquid coupling with tissue.

To compute a thermal image during laser-tissue interaction, the multiple ultrasound frames are captured during the procedure and stored offline. First, a motion tracking algorithm is applied to successive ultrasound frames to estimate the apparent time shifts (or displacements) [37]. Then the apparent time shifts are differentiated along the axial direction to obtain the normalized time shifts (or strain). Finally, the normalized time shifts are converted to temperature maps by utilizing the a priori known and independently measured relationship between speed of sound and temperature for the specific tissue under investigation.

11.7.1.3 Relationship Between Speed of Sound and Temperature

Temperature mapping using ultrasound imaging requires a look-up table or independent measurements to convert the normalized time shift to temperature changes [24, 25]. Here, we illustrate the independent measurements needed to obtain a relationship between the normalized time shift and temperature.

The temperature response of the sample was determined using a temperature controlled water bath, i.e., using the experimental setup presented in Fig. 11.26 where the laser was replaced with a water bath. Prior to the experiment, the sample was placed in a water bath and a thermistor was inserted in the center of the sample to measure temperature independently. Initially, a baseline ultrasound frame was captured. Then, the temperature of the water bath was gradually increased from 24 to 35°C and ultrasound frames were captured for every 1°C temperature increment.

Temperature distribution in the sample was assumed to be spatially homogenous at steady state. Normalized time shifts in the ultrasound signal due to temperature increase were computed in a 10 by 10 mm homogenous region near the thermistor.

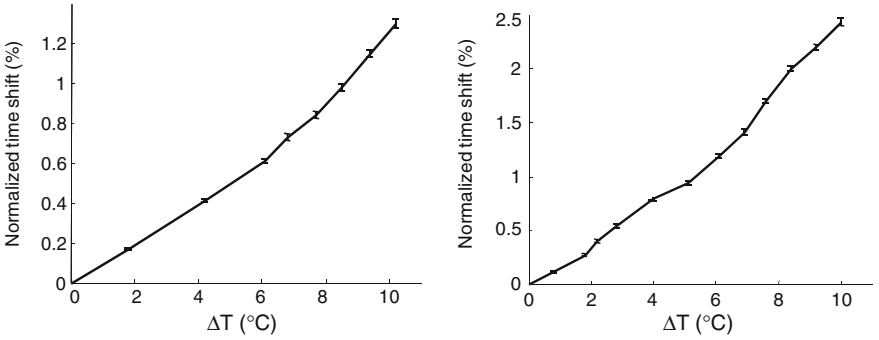


Fig. 11.27 Temperature dependence of normalized time shift for (**a**) polyvinyl alcohol (PVA) phantom and (**b**) sample of porcine muscle tissue. The error bars represent standard deviation obtained from 10 measurements. (Adapted from [26])

Thus, a normalized time shift vs. temperature dependence was measured for the phantom made out of polyvinyl alcohol (PVA) and for a sample of porcine muscle tissue (Fig. 11.27). The measured dependence was approximated using a second-order polynomial fit, and the coefficients of the polynomial fit were utilized later in remote measurements of temperature by relating the measured normalized time shifts to corresponding temperature during laser heating experiments.

In the temperature range up to 55°C, the normalized time shift in ultrasound signals is primarily caused by the speed of sound changing with temperature since thermal expansion of tissue is small and can be ignored [34, 35]. The speed of sound linearly increases for water and water based tissues between 10 and 55°C, and linearly decreases for lipid-based tissues [38]. Therefore, the calibration curves obtained at 24–35°C are valid at a physiological temperature of 37°C. Furthermore, the temperature elevations up to 20°C can be measured with the curves presented in Fig. 11.27.

11.7.1.4 Illustration of Ultrasound-Based Thermal Imaging

The ultrasound-based thermal imaging is first demonstrated using a tissue mimicking PVA phantom measuring 50 by 50 by 50-mm. PVA has been utilized to make both optical and ultrasound phantoms since it scatters light and has a speed of sound similar to tissues, and the modest optical absorption can be altered by additives such as dyes or nanoparticles [39, 40]. The phantom was fabricated by pouring 8% PVA solution into a mold and set to a desired shape by applying two freeze-thaw cycles. Gold nanoparticles of 70 nm diameter, having their absorption resonance close to 532 nm, were added to a circular inclusion (Fig. 11.28a) as photoabsorbers. Acoustic contrast was obtained by adding a higher (1.5%) concentration of silica particles in the inclusion as compared to the background (0.75%). To heat the phantom, a radiant energy of a continuous wave laser (532 nm wavelength, 1 W/cm^2 intensity) was used for 180 s while ultrasound frames were captured in real-time.

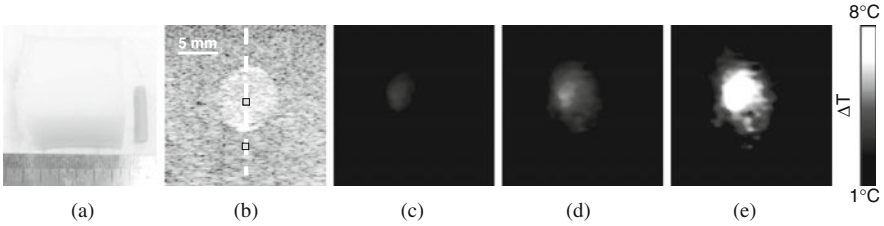


Fig. 11.28 (**a**) Photograph of the PVA phantom with the cylindrical inclusion. (**b**) Ultrasound image of the hyperechoic, optically absorbing inclusion. (**c–e**) Thermal images recorded at 30, 60 and 180 s of laser irradiation, respectively. The ultrasound and thermal images cover an 18 by 20 mm region. (Adapted from [26])

The ultrasound and thermal images covering an 18 by 20 mm field of view are presented in Fig. 11.28b–e. In these images, the ultrasound transducer is positioned at the top, and the laser beam is delivered from the bottom. The inclusion can be easily identified in the gray-scale B-mode ultrasound image (Fig. 11.28b) obtained using a 5–7 MHz linear transducer array. The contrast in this ultrasound image is related to elevated concentration of silica particles – ultrasound imaging cannot directly visualize nanoparticles. However, the ultrasound thermal imaging does not rely on the contrast.

The thermal maps computed at 30, 60 and 180 s of laser irradiation show the progressive increase in temperature. At 180 s, the inclusion has a peak temperature of over 7°C while the surrounding material has a temperature of less than 2°C. The temperature rise was confined to the inclusion due to the presence of nanoparticles. Since the thermal images are obtained from ultrasound images, the images in Fig. 11.28b–e are spatially co-registered.

The spatial and temporal rise of temperature is further examined in Fig. 11.29. The profile of temperature distribution (Fig. 11.29a) along a vertical line passing through the center of the inclusion shows that the temperature rise is primarily localized to the area with nanoparticles (i.e., the cylindrical inclusion). After 180 s, the temperature rises from the baseline of room temperature by more than 7°C. The temperature rise in a 1 mm by 1 mm region inside and outside of the inclusion was also examined (Fig. 11.29b). Mean temperature in the inclusion rises monotonically over time. However, the rate of temperature rise in the inclusion is non-linear due to heat diffusion into surrounding material. The region outside the inclusion has a significantly lower temperature elevation over time.

The laser irradiation and temperature monitoring was also carried out on a fresh sample of porcine muscle tissue. The ultrasound and thermal images (20 by 15 mm field of view) are presented in Fig. 11.30. Gold nanoparticles were injected in the porcine tissue at the site indicated in the ultrasound image (Fig. 11.30a). After 20 s of laser irradiation at 2, 3 and 4 W/cm^2, progressive temperature elevations were observed. Furthermore, higher laser energy resulted in higher temperature rise. In comparison, negligible temperature rise was observed in the control tissue sample injected with water (Fig. 11.30b).

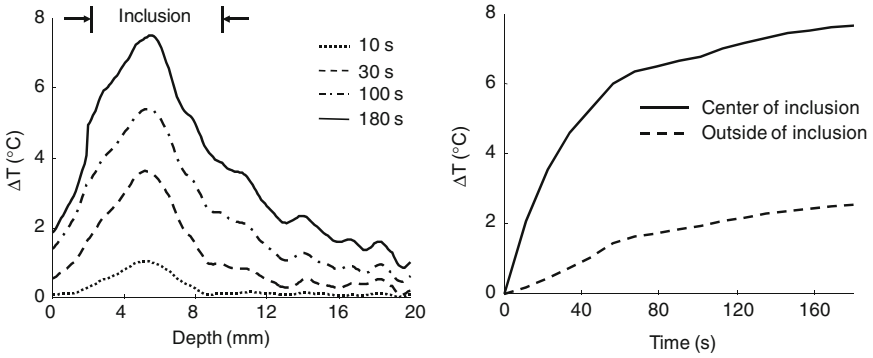


Fig. 11.29 (**a**) Spatial temperature profiles along the vertical line passing through the center of the inclusion (see Fig. 11.28b). (**b**) Temporal temperature profiles measured inside and outside of the inclusion in the locations indicated by the squares in Fig. 11.28b. (Adapted from [26])

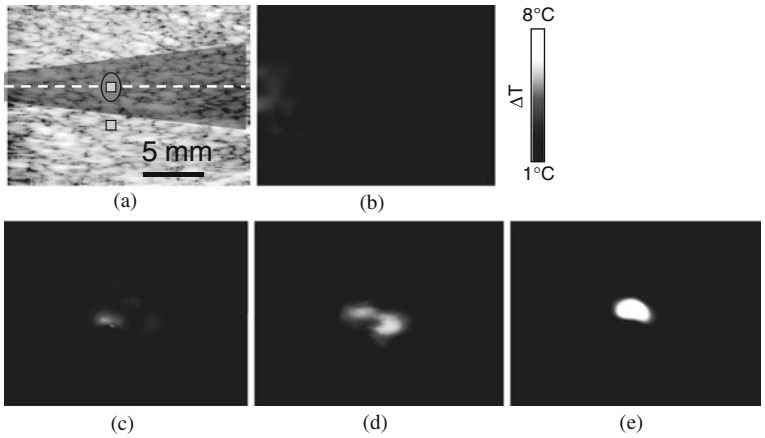


Fig. 11.30 (**a**) Ultrasound image of the porcine tissue sample. The photoabsorber injection site with respect to the laser beam is indicated by the *circle*. (**b**) The thermal image of the tissue sample injected with water and irradiated for 120 s at 2 W/cm^2 indicates negligible laser heating of the specimen. In contrast, thermal images (**c–e**) of the tissue injected with photoabsorbers and irradiated for 20 s at 2, 3 and 4 W/cm^2 respectively, clearly indicate progressive and localized temperature increase. All images cover a 20 mm by 15 mm region. (Adapted from [26])

Spatial temperature profile (Fig. 11.31a) was computed along the direction of the laser beam illustrated by the dashed line in the ultrasound image (Fig. 11.30a). After 120 s of laser heating at 2 W/cm^2, both control and nanoparticle-injected tissue had a temperature rise of close to 4°C near the surface which rapidly decreased to below 1°C at a depth of 3 mm. However, in the nanoparticle injected tissue, the temperature rise exceeded 12°C at the injection site. In the control tissue sample, no measurable temperature rise was seen deeper within the specimen.

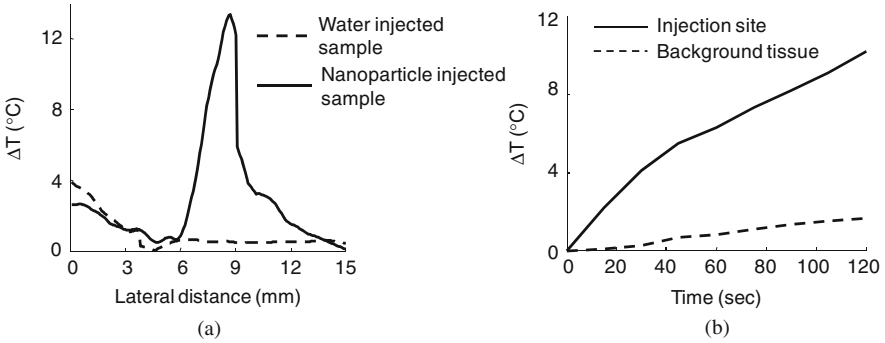


Fig. 11.31 (**a**) Spatial temperature profile measured along the *horizontal dotted line* shown in Fig. 11.30a for tissue specimens injected with either aqueous solution of photoabsorbers or water without photoabsorbers. (**b**) Temporal temperature profiles measured inside and outside of the targeted area (site of injection). The 1.5 by 1.5 mm regions for temporal assessment of temperature are indicated by *squares* in Fig. 11.30a. (Adapted from [26])

The temporal temperature profiles (Fig. 11.31b), measured in a 1.5 by 1.5 mm region within the injection site and in the surrounding tissue, were similar to those seen in the phantom experiment (Fig. 11.29b). In the area injected with nanoparticles, there is a steady rise in temperature with temperature reaching over 10°C elevation, while the background tissue shows a more gradual temperature increase due to thermal diffusion and lower optical absorption.

Ultrasound images, acquired before and after laser irradiation (Fig. 11.32a–b), illustrate the spatial location and extent of thermal lesion created during the laser heating procedure. In the ultrasound image after heating, there is an increase in echogenicity at the injection site as well as shadowing below the region (Fig. 11.32b). The thermal map (Fig. 11.32c) acquired after 180 s laser heating at 3 W/cm^2 showed that the temperature in the nanoparticles-injected region increased by over 25°C. Visual inspection of the samples (Fig. 11.32d) showed that the location of the nanoparticle injection was consistent with the site of temperature elevation seen in ultrasound and thermal images.

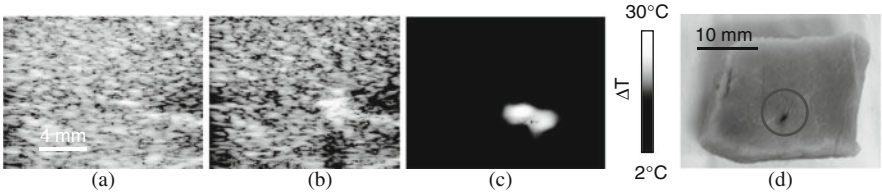


Fig. 11.32 Photothermal laser heating of porcine muscle tissue performed using 3 W/cm^2 laser irradiation for 180 s. Ultrasound images (**a**) before and (**b**) after laser irradiation indicate the location of heated region associated with changes in echogenicity. The location and the extent of the heated region are further confirmed in (**c**) thermal image captured after 180 s of laser heating. All images are 16 by 12 mm. The region of elevated temperature is consistent with injection site visible in the photograph (**d**) of the tissue sample. (Adapted from [26])

Finally, monitoring of the temperature produced by a pulsed laser irradiation was performed using a composite PVA phantom with an inclusion containing 30-μm graphite flakes acting as optical absorbers. The inclusion was embedded within the homogenous body of the phantom. Heat depositions were produced by 532 nm wavelength, 5 ns pulsed laser irradiation with energy fluencies of 30, 60, and 85 mJ/cm^2. A single laser pulse was fired and multiple ultrasound beams were acquired immediately following the laser pulse. Using 40 MHz single element transducer, approximately 1000 beams were captured covering a 100 ms observation time. The experimental setup was undisturbed and sufficient time was allowed for the phantom to return to room temperature before repeating the measurement at a different energy level.

The M-mode representation of the temporal variation in temperature along a single line passing through the inclusion is shown in Fig. 11.33a, c and e for laser radiant energies of 30, 60, and 85 mJ/cm^2, respectively. The horizontal axis in these plots represents an observation time of 100 ms and the vertical axis is the depth (4 mm). In the images, the ultrasound transducer is located at the top, while the laser beam irradiated the sample from the bottom. The circular inclusion, containing graphite particles, is approximately centered at a distance of 3.75 mm – towards the outer edge of the phantom where the laser irradiation was applied. Clearly, the temperature rise is predominantly in the inclusion, and there is a negligible temperature rise outside the inclusion.

The maximum temperature rise and temperature dynamics following a single laser pulse are shown quantitatively in Fig. 11.33b, d and f. The temperature was measured in a small region within the inclusion. The maximum temperatures measured after the laser pulse for radiant energy of 30, 60, and 85 mJ/cm^2 are 0.7±0.3, 2.9±0.2, and 5.0±0.2°C. Subsequently, an exponential decrease in temperature due to heat diffusion was observed. The thermal relaxation time measured by the time required for the temperature to decrease exponentially to 1/*e* (or 37%) of the maximum value was approximately 5 ms for the above radiant energy conditions.

11.7.1.5 Advantages and Limitations of Ultrasound-Based Thermal Imaging

Clearly, ultrasound imaging methods can be used to map the temperature change during both continuous and pulsed laser-tissue interactions. The temperature distribution during laser heating is affected by two processes – heat generation due to absorption of radiant energy and spatial redistribution of heat due to thermal diffusion. Therefore, it is important to monitor the rise of temperature both spatially and temporally. Ultrasound-based thermal images (Figs. 11.28 and 11.33) illustrate the feasibility of temporal tracking of temperature throughout the region of interest.

The temperature resolution in thermal ultrasound imaging is limited by the variance obtained from the motion estimation algorithm. For example, in the phantom experiments, variance in the normalized time shifts was ±0.02% corresponding to temperature resolution of about ±0.1°C. The spatial resolution for temperature measurements is limited by the kernel size employed in motion tracking algorithms. The kernel size is selected given the trade-off between spatial resolution

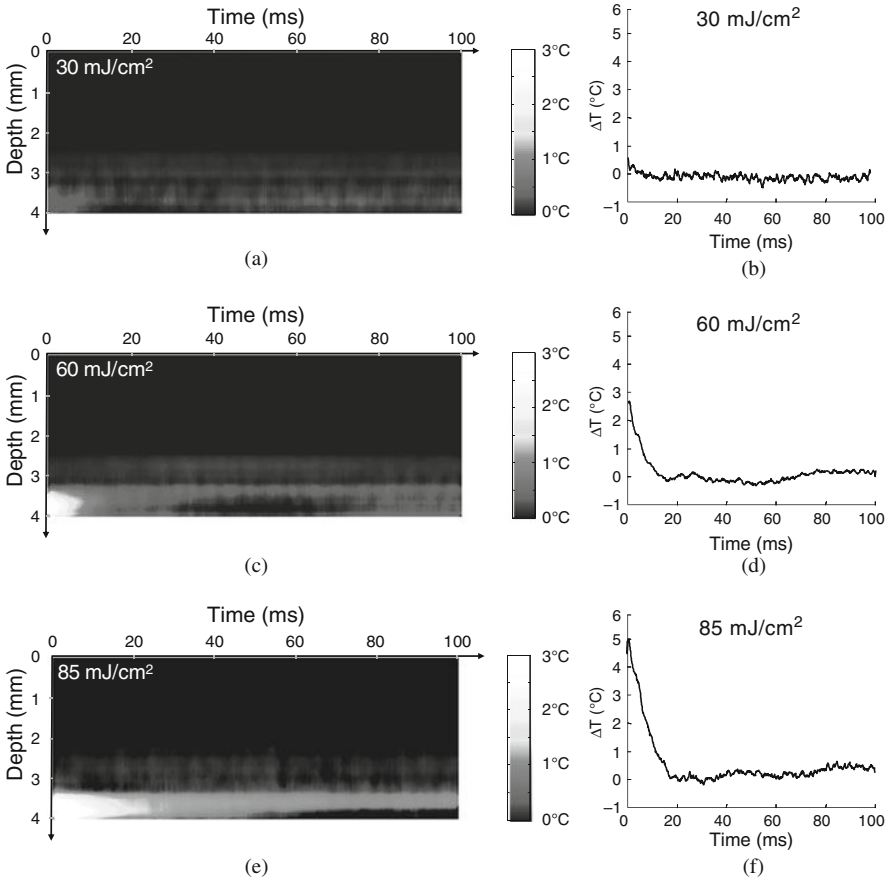


Fig. 11.33 Spatio-temporal ultrasound estimation of temperature in phantom following laser excitation. The M-mode image representation and the 1-D maximum temperature profile through the inclusion at laser energies (**a, b**) 30 mJ/cm^2, (**c, d**) 60 mJ/cm^2, (**e, f**) 85 mJ/cm^2 monitored for 100 ms following laser excitation shows an exponential decay. The thermal relaxation time was estimated to be approximately 5 ms. (Adapted from [41])

and displacement signal-to-noise ratio (SNR). A larger kernel leads to higher displacement and, therefore, strain SNR, while a smaller kernel is needed for better spatial resolution [37, 42]. In the phantom experiment (Fig. 11.28), a 0.6 mm axial by 2.1 mm lateral kernel was used, thus limiting the spatial temperature resolution to sub-millimeter axially and several millimeters laterally. Generally, however, spatial resolution scales with ultrasound frequency: an ultrasound pulse with higher bandwidth results in images with better spatial resolution although the penetration depth is sacrificed.

Physiological motion (e.g., cardiac, respiratory) could lead to artifacts in the ultrasound-based thermal imaging. For example, periodic heart beats cause tissue motion which appears as time shifts in the ultrasound signal and could lead to an

error in the temperature measurement. However, an electrocardiogram (ECG) could be utilized to capture ultrasound data at the same point in the cardiac cycle and thus potentially minimizing motion artifacts [43]. Finally, any abrupt changes in acoustic or structural properties of tissue in response to laser irradiation (e.g., tissue boiling and/or formation of gas bubbles) may cause errors in temperature estimation due to unexpected alteration of the ultrasound pulse-echo signals.

11.7.2 Thermal Photoacoustic Imaging

11.7.2.1 Photoacoustic Imaging

Photoacoustic [44, 45] or optoacoustic [46] and, generally, thermoacoustic [47] imaging relies on the absorption of electromagnetic energy, such as light, and the subsequent emission of an acoustic wave. To perform photoacoustic imaging, the tissue is irradiated with short laser pulses. The corresponding thermal expansion due to the absorption of light energy results in acoustic transients. Using an ultrasound sensor, the pressure pulses can be detected and spatially resolved to form an image of the tissue. If the laser pulse satisfies the condition of temporal stress confinement, the pressure rise p of the generated acoustic wave can be described by

$$p = \left(\frac{\beta v_s^2}{C_p}\right)\mu_a\psi = \Gamma\mu_a\psi \tag{11.51}$$

where β is the volume expansion coefficient, v_s is the speed of sound, C_p is the heat capacity at constant pressure, Γ is the Grueneisen coefficient, μ_a is the optical absorption coefficient of the medium and ψ is the laser fluence [46, 48]. Selection of the correct optical wavelength is extremely important in photoacoustic imaging – by utilizing the wavelength in the near infra-red (NIR) spectrum, photoacoustic transients can be generated at depths of a few centimeters [49].

11.7.2.2 Principles of Photoacoustic Thermal Imaging

Photoacoustic imaging can be utilized to monitor the temperature distribution in tissues non-invasively and in real time [27, 28, 50]. As evident from Eq. (11.51), the photoacoustic pressure amplitude is directly dependent on the dimensionless Grueneisen parameter.

$$\Gamma = \frac{\beta \cdot v_s^2}{C_p} \tag{11.52}$$

The volume expansion coefficient and the speed of sound are both temperature-dependent and linearly proportional to the temperature for water-based and fatty tissues in 10–55°C range [38, 51, 52]. Therefore, the Grueneisen parameter, and thus the photoacoustic signal, is directly related to temperature. For example, when

the temperature increases from 20 to 30°C, the volume expansion coefficient of water changes from $0.2{\cdot}10^{-3}(\mathrm{K})^{-1}$ to $0.3{\cdot}10^{-3}(\mathrm{K})^{-1}$ while the speed of sound varies from 1481 to 1507 m/s [51, 52]. Thus, for water, about 50% increase in photoacoustic signal amplitude occurs for a 10°C increase in temperature; almost a 5% per degree centigrade change in signal amplitude. A similar change in the photoacoustic signal is expected for tissue. For example, the change in the speed of sound is 1.83 m/(°C·s) for bovine liver and –7.4 m/(°C·s) for bovine fat, compared to 2.6 m/(°C·s) for water [38]. Furthermore, water bearing tissues have a volume expansion coefficient in the range of $0.26 \cdot 10^{-3}(\mathrm{K})^{-1} - 0.37.10^{-3}(\mathrm{K})^{-1}$ while lipid bearing tissues have a volume expansion coefficient in the range of $-0.8{\cdot}10^{-3}(\mathrm{K})^{-1}$ to $-1.76 \cdot 10^{-3}(\mathrm{K})^{-1}$ [51]. The positive changes in the speed of sound and volume expansion coefficients with temperature for tissue containing water and the corresponding negative response for lipid (fat) permit photoacoustic classification of unknown tissue or determination of the boundary between water-based tissue and lipid. Either the temperature dependent amplitude or phase shift (associated with the speed of sound) provides a clear indication of tissue type [53].

Therefore, under constant optical fluence, photoacoustic thermal imaging can be performed using the following relationship

$$\Delta T = a \cdot \frac{\Delta p}{p} = a \cdot \frac{\Delta \Gamma}{\Gamma} \tag{11.53}$$

where a is a tissue dependent constant that can be experimentally determined, Δp is the pressure rise when the temperature rises by ΔT, and $\Delta\Gamma$ is the corresponding change in the Grueneisen parameter [28].

The schematic view of a photoacoustic thermal imaging setup is presented in Fig. 11.34. In this particular setup, two lasers are used: a continuous wave laser is employed to produce laser-tissue interaction and a nanosecond pulsed laser is used to generate photoacoustic transients, thus allowing photoacoustic thermal imaging. It is possible to use a single pulsed laser with a higher energy mode employed for heating and a lower energy mode used for photoacoustic imaging. An ultrasound transducer is used to capture photoacoustic transients. The same transducer can be used for pulse-echo ultrasound imaging. Thus, both ultrasound and photoacoustic images can be acquired by this system. Furthermore, since the same transducer is used in both ultrasound and photoacoustics, the images will be spatially co-registered.

To measure the temperature, multiple photoacoustic frames are acquired during the laser-tissue interaction. The change in the normalized photoacoustic signal amplitude ($\Delta p/p$) is then computed and converted to temperature change.

11.7.2.3 Temperature Dependence of the Grueneisen Parameter

Photoacoustic-based thermal imaging, similar to ultrasound-based thermal imaging, requires a priori knowledge of the temperature dependence of the Grueneisen parameter to convert changes in the photoacoustic signal amplitude to changes

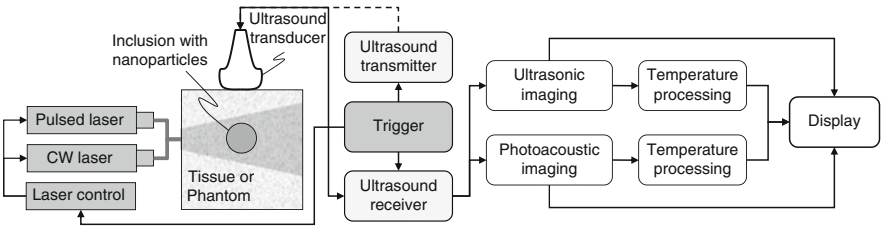
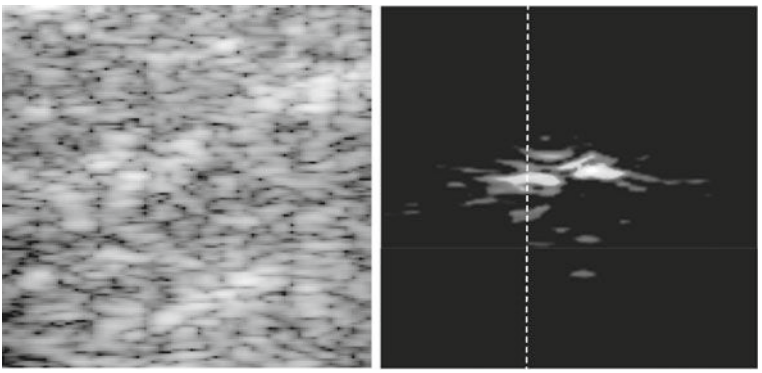


Fig. 11.34 Block diagram of photoacoustic (and ultrasound) thermal imaging system



(a) (b)

Fig. 11.35 (**a**) Ultrasound and (**b**) photoacoustic images of a porcine muscle tissue. Images cover a 15 by 15 mm field of view. (Adapted from [28])

in temperature. Generally, a tissue-specific database, containing the temperature-induced variations in the speed of sound and the thermal expansion coefficient for various tissues, can be created.

The temperature dependence of a porcine tissue was determined using a temperature controlled water bath experiment. In other words, the setup presented in Fig. 11.34 was modified to replace the CW laser with a water bath to provide known and well controlled temperature change of the sample. The initial ultrasound and photoacoustic images of the tissue sample placed in the water tank are presented in Fig. 11.35. Here the laser beam irradiated only a small region located in the center of the tissue sample. During the experiment, the temperature of the water bath was gradually increased and multiple ultrasound and photoacoustic frames were acquired.

In water-based tissue, the amplitude of the photoacoustic signal increases with temperature as indicated in Fig. 11.36a where multiple photoacoustic RF signals, collected from the region between the arrows in Fig. 11.35b at different temperatures, are plotted. There is a steady increase in the signal intensity from left (yellow) to right (red) as the temperature increases. In addition, the photoacoustic signal is shifting and appears to move closer to the transducer as the temperature increases. This effect can also be observed by comparing three photoacoustic RF signals (Fig. 11.36b) captured at the baseline temperature and temperature elevated by 5

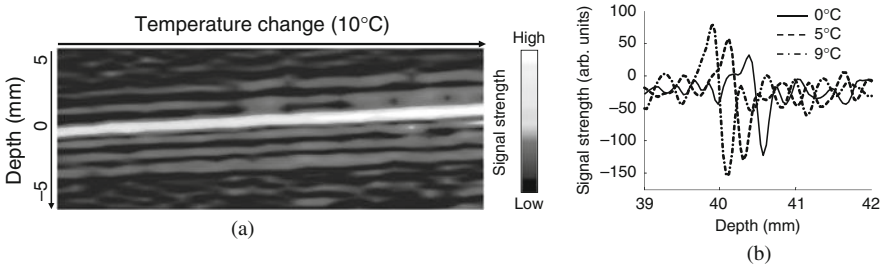


Fig. 11.36 Photoacoustic signal change with an increase in temperature. (**a**) Successive photoacoustic RF signals plotted for increasing temperatures show the increase of the amplitude of the photoacoustic signal. (Adapted from [28]) (**b**) RF signals at three different temperatures demonstrate the temporal shift and amplitude increase in photoacoustic signal

and 9°C, respectively. Overall, the photoacoustic signal has shifted by about 1 mm axially towards the transducer. This effect is due to thermally induced change of the speed of sound and is used to monitor temperature in ultrasound-based thermal imaging. Therefore, for accurate photoacoustic thermal imaging, the time shift in the photoacoustic signal must be accounted for before comparing photoacoustic signals.

Overall, the photoacoustic signal intensity changes significantly with a modest rise in temperature. Note that the changes of speed of sound and thermal expansion coefficient with temperature for water-based and lipid-based tissues have opposite signs. Thermal changes in photoacoustic signals for porcine muscle and porcine fat are compared in Fig. 11.37. As expected, the photoacoustic signal increased with temperature for porcine muscle (Fig. 11.37a), while the photoacoustic signal decreased for porcine fat (Fig. 11.37b) for the same temperature rise. Thus, temperature dependent changes in the photoacoustic signal potentially can be utilized to differentiate water-based and lipid-bearing tissues.

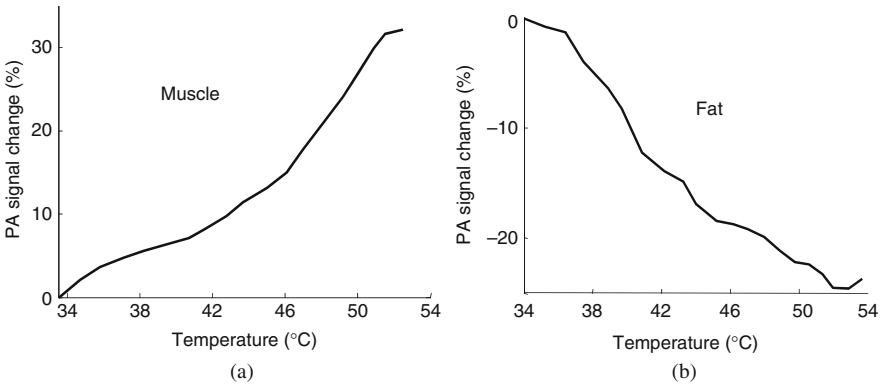


Fig. 11.37 Temperature induced change in photoacoustic signal measured in (**a**) porcine muscle and (**b**) porcine fat. (Adapted from [54])

Recently, metal nanoparticles were introduced as optical imaging and therapeutic agents due to their ability to interact with cells on a molecular level. For example, gold nanoshells and nanorods are commonly used as photoabsorbers in experimental investigations of photothermal therapy [28, 55, 56]. Thermal expansion of the tissue, due to this additional temperature increase, generates a photoacoustic response that is greater than the response from the tissue without any additives [57] since the addition of photoabsorbers effectively enhances the local optical absorption coefficient. However, the speed of sound and the volume expansion coefficient of the surrounding bulk tissue (and, therefore, the Grueneisen parameter) are responsible for the thermally induced changes in the photoacoustic signal (Eqs. (11.52)–(11.53)). Therefore, the temperature dependence of the photoacoustic signal will not change in the presence of photoabsorbers.

To demonstrate this effect, the photoacoustic signals from two PVA samples are compared – one phantom with gold nanoparticles added and another phantom without any additives. While the magnitude of photoacoustic signal was larger from the sample containing gold nanoparticles, the relative change in the photoacoustic signal for a 10°C increase in temperature was the same for both phantoms (Fig. 11.38), confirming that gold nanoparticles affect the magnitude of the photoacoustic signal (i.e., signal-to-noise ratio or SNR) but do not affect the rate of the signal change with temperature.

11.7.2.4 Illustration of Thermal Photoacoustic Imaging

To demonstrate photoacoustic thermal imaging, experiments were performed using a sample of porcine muscle tissue. Composite gold nanoparticles broadly absorbing in the near infra-red spectrum were directly injected into the sample. The injection site is shown by the dashed circle in the ultrasound image (Fig. 11.39a). There is strong photoacoustic response from the tissue due to the presence of nanoparticles

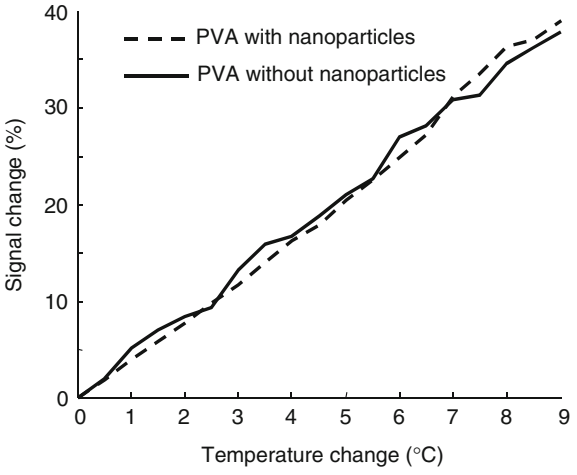


Fig. 11.38 Photoacoustic signal change for a 10°C temperature increase in plain PVA phantom and PVA phantom containing gold nanoparticles. (Adapted from [28])

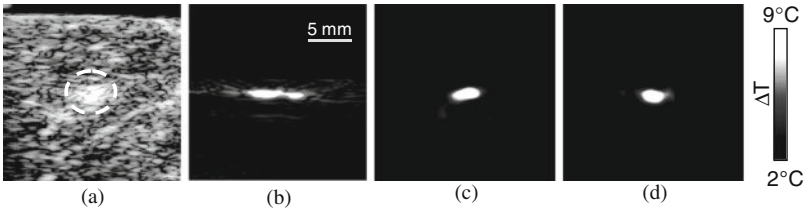


Fig. 11.39 Thermal imaging performed on porcine muscle tissue during photothermal continuous wave laser therapy. (**a**) Ultrasound and (**b**) photoacoustic images before therapy. (**c, d**) Temperature map after 3 min of photothermal therapy obtained from thermal ultrasound and photoacoustic imaging respectively. All images are 20 by 20 mm regions. (Adapted from [28])

(Fig. 11.39b). In addition, the photoacoustic signal from the native tissue is fairly weak as the imaging was performed at 800 nm – at this wavelength the tissue does not significantly absorb the radiant energy.

Following the ultrasound and photoacoustic imaging, laser heating was performed using a continuous wave laser (1 W/cm^2, 800 nm wavelength) for 4 min. During the laser therapy, thermal imaging was performed by both ultrasound and photoacoustic imaging techniques (Fig. 11.39c and d, respectively). After 4 min of therapy, the region with injected photoabsorbers reached a temperature elevation of greater than 9°C, while the surrounding region had a temperature rise of less than 2°C. The therapeutic region was spatially and temporally co-registered in both ultrasound and photoacoustic thermal images.

The temperature elevation, measured by the ultrasound and photoacoustic thermal imaging, was compared in a 1.5 by 1.5 mm region approximately centered in the therapeutic zone. The temperature rise computed by both methods was highly correlated (Fig. 11.40) throughout the procedure. The maximum temperature difference between the two imaging techniques was less than 0.5°C while the mean absolute temperature difference over the 4 min procedure was 0.26°C.

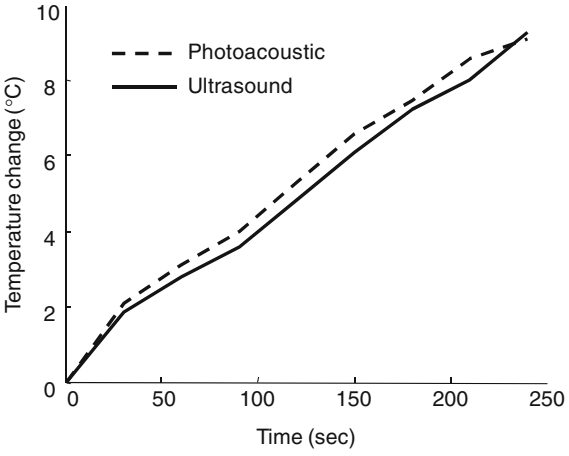


Fig. 11.40 Comparison of ultrasound and photoacoustic measurements of temperature. The measurements were obtained from a 1.5 by 1.5 mm region. (Adapted from [28])

11.7.2.5 Advantages and Limitations of Thermal Photoacoustic Imaging

Thermal photoacoustic imaging can be used to monitor temperature elevations (Fig. 11.39) during laser-tissue interaction in general and photothermal therapy of tissue in particular. The temperature elevation measured photoacoustically was comparable to thermal ultrasound imaging (Fig. 11.40). In addition, during nanoparticle-enhanced photothermal therapy, the contrast in photoacoustic images is significantly enhanced due to the presence of exogenous contrast agents. Therefore, photoacoustic thermal imaging might have higher contrast-to-noise ratio (CNR) as compared to ultrasound thermal imaging [28]. However, for deeper penetration, photoacoustic imaging with nanoparticles is often performed at near infrared wavelengths. Consequently, the photoacoustic response from surrounding tissue without nanoparticles may not have sufficient SNR for reliable measurements of temperature. In this case, a combination of the two imaging techniques may be needed to monitor temperature throughout the tissue sample where photoacoustic imaging is used to monitor temperature in the target region with enhanced optical contrast while ultrasound imaging is utilized for thermal imaging of the surrounding tissue.

Since thermal photoacoustic imaging depends on the photoacoustic signal strength, the temperature resolution is inherently limited by the pulse-to-pulse laser stability. In our experiments, the average standard deviation for the photoacoustic signal change was 5% which lead to a temperature resolution of 0.3°C. However, the temperature resolution of thermal photoacoustic imaging can be significantly improved by using highly stable lasers and/or performing averaging of several laser pulses.

Photoacoustic thermal imaging relies on the photoacoustic signal change with temperature, and, therefore, any changes in the signal related to physiological motion could potentially introduce artifacts in the temperature monitoring. However, photoacoustic thermal imaging is dependent on signal strength to estimate temperature and is, therefore, less sensitive to tissue motion as compared to ultrasound thermal imaging which is highly sensitive to tissue motion. Nevertheless, for both imaging techniques, it is possible to reduce motion artifacts by utilizing the gating techniques (e.g., ECG) to collect motion-synchronized images.

11.7.3 Summary

Real-time ultrasound and photoacoustic thermal imaging is possible. Each imaging system alone or combined imaging system is capable of accurate monitoring of thermal laser-tissue interactions. Both imaging techniques can be integrated with optical systems and require minimal changes in the experimental setups. Overall, ultrasound or photoacoustic imaging can be used to visualize volumetric distribution of temperature at sufficient depth and with high temporal and spatial resolution, high sensitivity, accuracy and signal-to-noise ratio.

Acknowledgments Dr. A. J. Welch provided us with copious amounts of information on thermistors and thermocouples. David Yuan performed many of the finite element analyzes. Jignesh Shah and Shriram Sethuraman performed many ultrasound and photoacoustic measurements of temperature. Experiments shown in this chapter were funded in part by grants from the Whitaker Foundation, the National Institutes of Health (GM31931-01A3, EB004963, EB008101 and HL084076), and the Texas Advanced Technology Program (#003658-570).

References

1. Bligh J. Regulation of body temperature in man and other mammals. In: A Shitzer and RC Eberhart (eds) *Heat transfer in medicine and biology*, Vol. 1. Plenum, New York, pp. 15–52 (1985).
2. Rowell LB and Wyss CR. Temperature regulation in exercising and heat-stressed man. In: A. Shitzer and RC Eberhart (eds) *Heat transfer in medicine and biology*, Vol. 1. Plenum, New York, pp. 53–78 (1985).
3. Lipton JM. Thermoregulation in pathological states. In: A Shitzer and RC Eberhart (eds) *Heat transfer in medicine and biology*, Vol. 1. Plenum, New York, pp. 79–106 (1985).
4. Webster JG. Basic sensors and principles. In: JG Webster (ed) *Medical instrumentation application and design*. Houghton Mifflin, Boston, pp. 385–391, 414–417 (1978).
5. Jain RK. Analysis of heat transfer and temperature distributions in tissues during local and whole-body hyperthermia. In: A Shitzer and RC Eberhart (eds) *Heat transfer in medicine and biology*, Vol. 2. Plenum, New York, pp. 3–54 (1985).
6. McGrath JJ. Preservation of biological material by freezing and thawing. In: A Shitzer and RC Eberhart (eds) *Heat transfer in medicine and biology*, Vol. 2. Plenum, New York, pp. 185–238 (1985).
7. Trezek GJ. Thermal analysis for cryosurgery. In: A Shitzer and RC Eberhart (eds) *Heat transfer in medicine and biology*, Vol. 2. Plenum, New York, pp. 239–260 (1985).
8. Davis TP. In vivo temperature measurements. *Acta. Ophthalmol*., (Suppl) 76:41–50 (1963).
9. RC Weast (ed). *CRC handbook of chemistry and physics*, B220, D165, E12, F170. CRC, Cleveland, 58th edition (1977).
10. Sostman HE. Melting point of Gallium as a temperature calibration standard. *Rev. Sci. Instrum*., 48:127–130 (1977).
11. Cobbold RSC. *Transducers for biomedical measurements: Principles and applications*. Wiley, New York, pp. 57–113 (1974).
12. *The temperature handbook*. Omega Technologies, Stamford CT, pp. A36–A38 (1989).
13. *Thermistors*, Catalog No. 182. Thermometrics, Edison, NJ (1992).
14. *Thermistor sensor handbook*. Thermometrics, Edison, NJ, pp. 2–27 (1987).
15. Valvano JW. Temperature measurements. *Adv. Heat Transfer*, Academic Process, 22:359–436 (1992).
16. Welch AJ and Pearce JA. Thermocouples. In: JG Webster (ed) *Encyclopedia of medical devices and instrumentation*, Vol. 4. Wiley, New York, pp. 2739–2746 (1988).
17. Balasubramaniam TA and Bowman HF. Thermal conductivity and thermal diffusivity of biomaterials: A simultaneous measurement technique. *J. Biomech. Eng.*, 99:148–154 (1977).
18. Valvano JW, Allen JT, and Bowman HF. The simultaneous measurement of thermal conductivity, thermal diffusivity and perfusion in small volumes of tissue. *J. Biomech. Eng.*, 106:192-197 (1984).
19. Valvano JW, Cochran JR, and Diller KR. Thermal conductivity and diffusivity of biomaterials measured with self-heated thermistors. *Int. J. Thermophys.*, 6:301–311 (1985).
20. Cain CP and Welch AJ. Thin-film temperature sensors for biological measurements. *IEEE Trans. Biomed. Eng.*, BME-21:421–423 (1974).
21. Rosenfeld A and Kak AC. *Digital picture processing*, Academic, New York, 2nd Edition (1982).

22. Ishihara Y, Calderon A, Watanabe H, Okamoto K, Suzuki Y, Kuroda K, and Suzuki Y. A precise and fast temperature mapping using water proton chemical shift. *Magn. Reson. Med*., 34:814–823 (1995).
23. Quesson B, de Zwart JA, and Moonen CT. Magnetic resonance temperature imaging for guidance of thermotherapy. *J. Magn. Res. Imaging*, 12:525–533 (2000).
24. Seip R and Ebbini ES. Non-invasive monitoring of ultrasound phased array hyperthermia and surgery treatments. *IEEE 17th Annu. Conf. Eng. Med. Biol. Soc.*, 1:663–664, 661 (1995).
25. Varghese T, Zagzebski JA, Chen Q, Techavipoo U, Frank G, Johnson C, Wright A, and Lee FT Jr. Ultrasound monitoring of temperature change during radiofrequency ablation: Preliminary in-vivo results. *Ultrasound Med. Biol*., 28:321–329 (2002).
26. Shah J, Aglyamov SR, Sokolov K, Milner TE, and Emelianov SY. Ultrasound imaging to monitor photothermal therapy – feasibility study. *Opt. Express*, 16:3776–3785 (2008).
27. Larina IV, Larin KV, and Esenaliev RO. Real-time optoacoustic monitoring of temperature in tissues. *J. Phys. D: Appl. Phys*., 38:2633–2639 (2005).
28. Shah J, Park S, Aglyamov S, Larson T, Ma L, Sokolov K, Johnston K, Milner T, and Emelianov S. Photoacoustic imaging and temperature measurement for photothermal therapy. *J. Biomed. Opt*., 13:034024 (2008).
29. Meaney PM, Paulsen KD, and Ryan TP. Microwave thermal imaging using a hybrid element method with a dual mesh scheme for reduced computation time. *Proc. 15th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 96–97 (1993).
30. Paulsen KD, Moskowitz MJ, and Ryan TP. Temperature field estimation using electrical impedance profiling methods. I. Reconstruction algorithm and simulated results. *Int. J. Hyperthermia*, 10:209–228 (1994).
31. Moskowitz MJ, Paulsen KD, Ryan TP, and Pang D. Temperature field estimation using electrical impedance profiling methods. Ii. Experimental system description and phantom results. *Int. J. Hyperthermia*, 10:229–245 (1994).
32. Seip R and Ebbini ES. Noninvasive estimation of tissue temperature response to heating fields using diagnostic ultrasound. *IEEE Trans. Biomed. Eng*., 42:828–839 (1995).
33. Arthur RM, Straube WL, Starman JD, and Moros EG. Noninvasive temperature estimation based on the energy of backscattered ultrasound. *Med. Phys*., 30:1021–1029 (2003).
34. Maass-Moreno R and Damianou CA. Noninvasive temperature estimation in tissue via ultrasound echo-shifts. Part i. Analytical model. *J. Acoustical Soc. Am*., 100:2514–2521 (1996).
35. Maass-Moreno R, Damianou CA, and Sanghvi NT. Noninvasive temperature estimation in tissue via ultrasound echo-shifts. Part ii. In vitro study. *J. Acoustical Soc. Am*., 100:2522–2530 (1996).
36. Shah J, Aglyamov SR, Sokolov K, Milner TE, and Emelianov SY. Ultrasound-based thermal and elasticity imaging to assist photothermal cancer therapy – preliminary study. *Proc. 2006 IEEE Ultrasonics Sympos*., 1029–1032 (2006).
37. Lubinski MA, Emelianov SY, and O'Donnell M. Speckle tracking methods for ultrasonic elasticity imaging using short-time correlation. *IEEE Trans. Ultrasonics, Ferroelectrics Frequency Control*, 46, 82–96 (1999).
38. Bamber JC and Hill CR. Ultrasonic attenuation and propagation speed in mammalian tissues as a function of temperature. *Ultrasound Med. Biol*., 5:149–157 (1979).
39. Hassan CM and Peppas NA. Structure and applications of poly(vinyl alcohol) hydrogels produced by conventional crosslinking or by freezing/thawing methods. *Adv. Polymer Sci*., 153:37–65 (2000).
40. Kharine A, Manohar S, Seeton R, Kolkman RG, Bolt RA, Steenbergen W, and Mul FFd. Poly(vinyl alcohol) gels for use as tissue phantoms in photoacoustic mammography. *Phys. Med. Biol*., 48, 357–370 (2003).
41. Sethuraman S, Aglyamov SR, Smalling R, and Emelianov SY. Ultrasound based remote temperature measurements in intravascular photoacoustic imaging, *Ultrasound Med. Biol*., 34:299–308 (2008).

42. Srinivasan S, Righetti R, and Ophir J. Trade-offs between the axial resolution and the signal-to-noise ratio in elastography. *Ultrasound Med. Biol*., 29:847–866 (2003).
43. Simon C, VanBaren PD, and Ebbini ES. Motion compensation algorithm for noninvasive two-dimensional temperature estimation using diagnostic pulse-echo ultrasound. *SPIE-Surg. Appl. Energy*, 3249:182–192 (1998).
44. Wang X, Xu Y, Xu M, Yokoo S, Fry ES, and Wang LV. Photoacoustic tomography of biological tissues with high cross-section resolution: Reconstruction and experiment. *Med. Phys*., 29:2799–2805 (2002).
45. Kruger RA, Liu P, Fang YR, and Appledorn CR. Photoacoustic ultrasound (paus)-reconstruction tomography. *Med. Phys*., 22:1605–1609 (1995).
46. Oraevsky AA and Karabutov AA. *Optoacoustic tomography*. In: T Vo-Dinh (ed) *Biomedical photonics handbook*, CRC Press, Boca Raton, Florida (2003).
47. Kruger RA, Reinecke DR, and Kruger GA. Thermoacoustic computed tomography-technical considerations. *Med. Phys*., 26:1832–1837 (1999).
48. Oraevsky AA, Jacques SL, and Tittle FK. Determination of tissue optical properties by time-resolved detection of laser-induced stress waves. *Soc. Photo-Optical Instrum. Engineers (SPIE) Conf*., 86:1882 (1993).
49. Esenaliev RO, Karabutov AA, and Oraevsky AA. Sensitivity of laser opto-acoustic imaging in detection of small deeply embedded tumors. *IEEE J. Sel. Top. Quant. Electron*., 5:981–988 (1999).
50. Shah J, Park S, Aglyamov S, Larson T, Ma L, Sokolov K, Johnston K, Milner T, and Emelianov S. Photoacoustic and ultrasound imaging to guide photothermal therapy: Ex-vivo study. *Proc. SPIE*, 6856:68560U (2008).
51. Duck FA. *Physical properties of tissue*. Academic, New York (1990).
52. Kikoin IK. *Handbook of chemistry and physics*. CRC, Cleveland, OH (1974).
53. Huang S, Kim K, Witte R, Olafsson R, and O'Donnell M. Inducing and imaging thermal strain using a single ultrasound linear array. *IEEE Trans. Ultrasonics, Ferroelectrics and Frequency Control*, 54:1718–1720 (2007).
54. Shah J, Thompson S, Milner TE, and Emelianov SY, Ultrasound guidance and monitoring of laser-based fat removal. *Lasers Surg. Med*., 40:680–687 (2008).
55. Hirsch LR, Stafford RJ, Bankson JA, Sershen SR, Rivera B, Price RE, Hazle JD, Halas NJ, and West JL. Nanoshell-mediated near-infrared thermal therapy of tumors under magnetic resonance guidance. *Proc. Natl. Acad. Sci. U.S.A*., 100:13549–13554 (2003).
56. Huang X, El-Sayed IH, Qian W, and El-Sayed MA. Cancer cell imaging and photothermal therapy in the near-infrared region by using gold nanorods. *J. Am. Chem. Soc*., 128:2115–2120 (2006).
57. Wang Y, Xie X, Wang X, Ku G, Gill KL, O'Neal DP, Stoica G, and Wang LV. Photoacoustic tomography of a nanoshell contrast agent in the in vivo rat brain. *Nano Lett*., 4:1689–1692 (2004).

Chapter 12
Tissue Thermal Properties and Perfusion

Jonathan W. Valvano

12.1 Basic Definitions

The transport of thermal energy in living tissue is a complex process involving multiple phenomenological mechanisms including conduction, convection, radiation, metabolism, evaporation, and phase change. There are three objectives of this chapter. First, the basic definitions of thermal properties and perfusion are presented. Second, experimental techniques to measure thermal properties and perfusion are developed in enough detail that the reader can perform their own measurements. Lastly, equilibrium thermal property and perfusion data are presented in graphical and tabular forms.

12.1.1 Significance

The understanding of heat-transfer mechanisms occurring in biomaterials requires the accurate knowledge of both tissue thermal properties and perfusion [1]. Applications of bio-heat transfer include the burn-injury process, cyropreservation, cryoprotection, laser arterial grafting, laser angioplasty, hyperthermia cancer therapy, thermography, electrocautery, and environmental stress. Because many biologic functions are temperature dependent, thermal conductivity and thermal diffusivity play a role in understanding these physiologic and pathophysiologic phenomena. Even though perfusion strongly affects heat transfer in vivo, a data base of in vitro thermal properties, well categorized by temperature and tissue type, will translate into an improved understanding of the many heat transfer applications in medicine.

Because most medical applications of heat and cold are transient, the tissue thermal diffusivity is also important. Accurate measurements of this transient thermal property are required in order to develop realistic time-dependent thermal models.

J.W. Valvano (✉)
Department of Electrical and Computer Engineering, The University of Texas, Austin, TX, USA
e-mail: valvano@mail.utexas.edu

A.J. Welch, M.J.C. van Gemert (eds.), *Optical-Thermal Response of Laser-Irradiated Tissue*, 2nd ed., DOI 10.1007/978-90-481-8831-4_12,

12.1.2 Thermal Conductivity

Thermal conductivity, k, is defined as the steady-state ability of an object to conduct heat when subject to a temperature gradient on its surfaces (see Fig. 12.1). In flat wall (e.g., across the skin), the steady state heat flow across a solid object is given by

$$q = \frac{dQ}{dt} = -kA\frac{\Delta T}{\Delta L} \tag{12.1}$$

where Q is the thermal energy, q is the heat flow, k is the thermal conductivity of the object, A is the surface area of the object, $\Delta T/\Delta L$ is the temperature gradient across the object. The minus sign is due to the second law of thermodynamics indicating heat flows from the hot side to the cold side.

In a cylindrical wall (e.g., across an arterial wall) the steady state heat flow across a cylindrical shell is given by

$$q = \frac{dQ}{dt} = -2\pi\, kL\frac{\Delta T}{\ln\left(r_2/r_1\right)} \tag{12.2}$$

where L is the cylinder length, r_2 is the outer radius and r_1 is the inner radius.

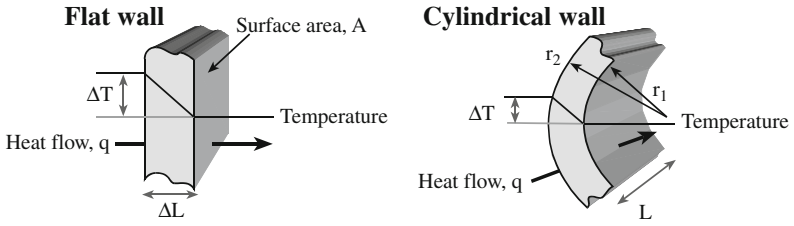


Fig. 12.1 Conduction through a stationary medium

12.1.3 Thermal Diffusivity

Thermal diffusivity, α, is defined as the transient ability of an object to conduct heat when subject to a spatially varying temperature. The general form of the conduction equation is

$$\frac{\partial}{\partial x}\left(k\frac{\partial T}{\partial x}\right) + \frac{\partial}{\partial y}\left(k\frac{\partial T}{\partial y}\right) + \frac{\partial}{\partial z}\left(k\frac{\partial T}{\partial z}\right) + S = \rho c\frac{\partial T}{\partial t} \tag{12.3}$$

where S is the internal heat generation [W/cm^3], ρ is the density and c is the specific heat at constant pressure. Assuming the thermal conductivity is homogeneous (doesn't vary with position), the conduction equation simplifies to

$$\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2} + \frac{S}{k} = \frac{1}{\alpha}\frac{\partial T}{\partial t} \tag{12.4}$$

where

$$\alpha = \frac{k}{\rho c} \tag{12.5}$$

In cylindrical coordinates, the conduction equation is

$$\frac{\partial^2 T}{\partial r^2} + \frac{1}{r}\frac{\partial T}{\partial r} + \frac{1}{r^2}\frac{\partial^2 T}{\partial \phi^2} + \frac{\partial^2 T}{\partial z^2} + \frac{S}{k} = \frac{1}{\alpha}\frac{\partial T}{\partial t} \tag{12.6}$$

In spherical coordinates, the conduction equation is

$$\frac{1}{r}\frac{\partial^2}{\partial r^2}(rT) + \frac{1}{r^2 \sin\psi}\frac{\partial}{\partial \psi}\left(\sin\psi \frac{\partial T}{\partial \psi}\right) + \frac{1}{r^2 \sin\psi}\frac{\partial^2 T}{\partial \phi^2} + \frac{S}{k} = \frac{1}{\alpha}\frac{\partial T}{\partial t} \tag{12.7}$$

To better understand the meaning of thermal diffusivity, first consider a semi-infinite 1-D example, as shown in Fig. 12.2. Initially the temperature in the solid is T_i. At time zero, the temperature of the flat surface is fixed at T_s. The solution of Eq. (12.4) under these simple conditions is

$$\frac{T(x,t) - T_s}{T_i - T_s} = erf\left(\frac{x}{\sqrt{4\alpha t}}\right) \tag{12.8}$$

where x is the distance from the surface and erf() is defined as

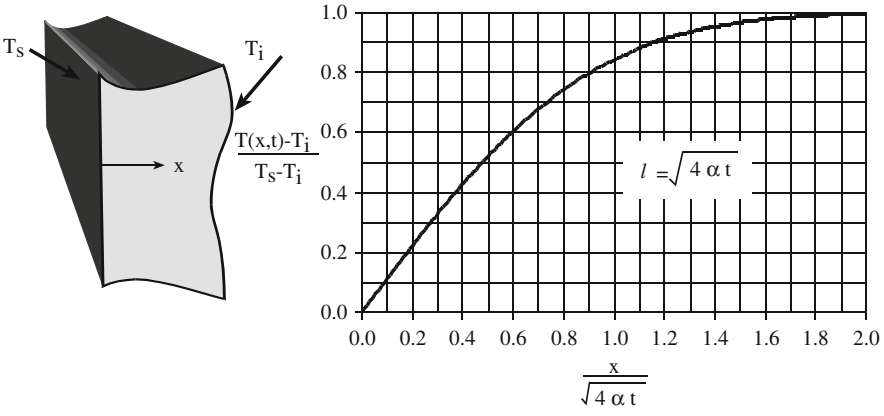


Fig. 12.2 Transient response of a semi-infinite object with fixed surface temperature

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-u^2} du \tag{12.9}$$

The solution in Eq. (12.8) motivates the concept of characteristic length, $l = \sqrt{4\alpha t}$. In fact, if the thickness of a finite-sized object is large compared to its characteristic length, then it can be considered semi-infinite, and Eq. (12.8) applies.

The second situation in which thermal diffusivity is important is the estimation of thermal time constant, i.e., how fast will an object respond to an external thermal stimulus. In this example, a spherical object (with radius a) is initially T_i. At time zero, the temperature on the surface of the sphere is fixed at T_s. For this situation, time constant, τ, is defined as the time it takes for the average temperature in the sphere to reach 63% ($1–e^{-1}$) of the final temperature. Using numerical simulations described in Section 11.5, the time constant for a sphere can be determined from its surface area and its thermal diffusivity.

$$\tau \;=\; 0.00433\, \frac{4\pi\, a^2}{\alpha} \tag{12.10}$$

It is important to realize that the time constant, also called diffusion time, is highly dependent on the geometry of a problem. For a one dimensional semi-infinite geometry, the diffusion time is given by

$$\tau \;=\; \frac{l^2}{4\alpha} \tag{12.11}$$

12.1.4 Specific Heat

Specific heat, c, determines the amount of thermal energy that will be stored in a material at a certain temperature. From thermodynamics, specific heat is defined as the change in enthalpy, h, with respect to temperature, T, at constant pressure. Specific heat at constant pressure is also related to its thermal conductivity, thermal diffusivity and density.

$$c \;\equiv\; \left.\frac{\partial h}{\partial T}\right|_p \;=\; \frac{k}{\rho\alpha} \tag{12.12}$$

For an object with homogeneous properties at constant temperature, the total thermal energy is a function of its density, specific heat, volume, and absolute temperature.

$$Q \;=\; \rho\, c\, V\, T \tag{12.13}$$

where V is the volume, and T is the temperature in Kelvin.

12.1.5 Tissue Perfusion

Blood flow is the volume of blood moving per unit of time. Blood flow is usually measured across a surface and includes both a direction and amplitude. For example, a cardiac output of 5 L/min means the heart will pump about 83 mL of blood per second across the aortic valve. Significant heat transfer occurs at the terminal arterioles, which are vessels with diameters of about 50–100 μm. Unfortunately, there are millions of terminal arterioles in a typical tissue of interest. Consequently, the thermal effects of these many blood vessels are typically lumped into a single nondirectional parameter called perfusion, as illustrated in Fig. 12.3. Clinically, perfusion is often measured as volume flow rate of blood per mass of tissue, having units of mL/100 g-min. However, when analyzing heat transfer, we sometimes define perfusion as the mass flow rate of blood per volume of tissue:

$$w \equiv \frac{\text{mass}_{\text{blood}}}{\text{volume}_{\text{tissue}}\text{time}} \ (\text{g cm}^{-3}\text{s}^{-1}) \quad \text{or} \quad w \equiv \frac{\text{volume}_{\text{blood}}}{\text{mass}_{\text{tissue}}\text{time}} \ (\text{mL/100 g-min}) \tag{12.14}$$

Blood perfusion through the vascular network and the local temperature distribution are interdependent. Many environmental (e.g., heat stress and hypothermia), pathophysiologic (e.g., inflammation and cancer), therapeutic (e.g., heating/cooling pads) situations create a significant temperature difference between the blood and the tissue through which it flows. The temperature difference causes convective heat transport to occur, altering the temperatures of both the blood and the tissue. Perfusion based heat transfer interaction is critical to a number of physiological processes such as thermoregulation and inflammation.

The convective heat transfer depends on the rate of perfusion and the vascular anatomy, which vary widely among the different tissues, organs of the body,

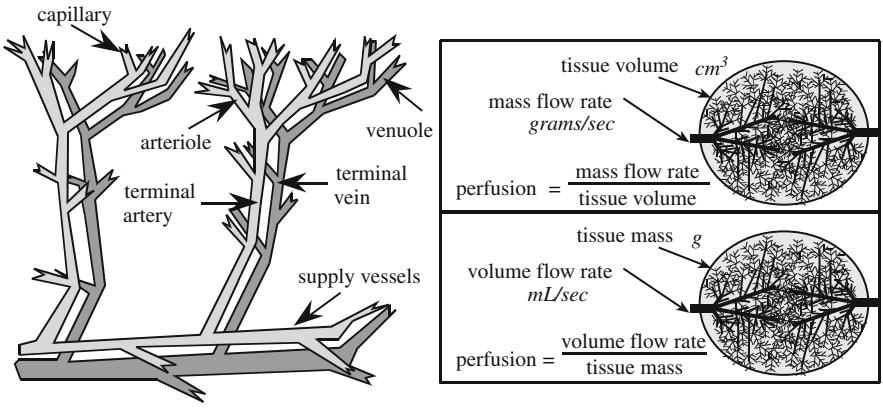


Fig. 12.3 Perfusion is the nondirectional volumetric blood flow rate in tissue

and pathology. Diller et al. published an extensive treatment of bioheat transfer phenomena [2]. In living tissue, blood perfusion provides a significant contribution to heat transport that must be adequately modeled in order to make accurate thermal predictions.

12.1.6 Reviews of Thermal Measurements

There are many good reviews of techniques to measure thermal properties [3–6]. *Thermophysical Properties of Matter* is a ten volume set which catalogs thermal properties. Volumes 1 and 3 contain thermal conductivity data, and volume 10 contains thermal diffusivity data. Extensive reviews of measurement techniques exist as a preface to each volume of the set. Additional thermal property data can be found in Eckert and Drake [7]. John Chato has written an excellent chapter in *Heat Transfer in Medicine and Biology*, edited by Shitzer and Eberhart, which reviews techniques to measure thermal properties of biologic materials [3]. Duck has also written an excellent review chapter on this subject [8].

12.2 Measurement of Thermal Properties

12.2.1 Overview

Rather than review all the methods possible for measuring thermal properties, this section presents one method the reader can employ if they wish to measure thermal properties on their own. Thermal probe techniques are used frequently to determine the thermal conductivity and the thermal diffusivity of biomaterials [9]. Common to these techniques is the use of a thermistor bead either as a heat source or a temperature sensor. Various thermal diffusion probe techniques have been developed from Chato's first practical use of the thermal probe. Physically, for all of these techniques, heat is introduced to the tissue at a specific location and is dissipated by conduction through the tissue and by convection with the blood perfusion.

Thermal probes are constructed by placing a miniature thermistor at the tip of a plastic catheter. The volume of tissue over which the measurement occurs depends on the surface area of the thermistor. Electrical power is delivered simultaneously to a spherical thermistor positioned invasively within the tissue of interest. The tissue is assumed to be homogeneous within 1 mL around the probe. The electrical power and the resulting temperature rise are measured by a microcomputer-based instrument. When the thermal probe is inserted into living tissue, the thermistor heat is removed both by conduction and by heat transfer due to blood flow near the probe. In vivo, the instrument measures effective thermal properties, which are the combination of conductive and convective heat transfer. Thermal properties are derived from temperature and power measurements using equations that describe heat transfer in the integrated probe/tissue system.

The following five complexities make the determination of thermal properties a technically challenging task. First, tissue heat transfer includes conduction, convection, radiation, metabolism, evaporation, and phase change. It is difficult but necessary to decouple these different heat transfer mechanisms. Second, the mechanical and thermal interactions between the probe and tissue are complex, and must be properly modeled to achieve accurate measurements. When the probe is inserted into living tissue, a blood pool may form around the probe because of the mechanical trauma. Because the probe is most sensitive to the tissue closest to it, the presence of a pool of blood will significantly alter the results. Tissue damage due to probe insertion may also occur in vitro. Third, the tissue structure is quite heterogeneous within each sample. Thus, the probe (which returns a single measurement value) measures a spatial average of the tissue properties surrounding the active elements. Unfortunately, the spatial average is very nonuniform. The probe is most sensitive to the tissue immediately adjacent to it. It is important to control this effective measurement volume. If the effective volume is too small, then the measurement is highly sensitive to the mechanical/thermal contact between the probe and tissue. If the effective volume is too large, then the measurement is sensitive to the boundary conditions at the surface of the tissue sample. Fourth, there are significant sample to sample and species to species variabilities. One must be careful when extrapolating results obtained in one situation to different situations. Fifth, tissue handling is critical. Thermal properties are dependent on temperature and water content [1–8]. Blood flow, extracellular water, and local metabolism are factors which strongly affect heat transfer in living tissue, but are difficult to experimentally determine or control. Once a tissue dies, if handled improperly there will be significant water fluxes which will affect tissue thermal properties. Tissues should be stored in a slightly hypertonic saline buffer to minimize tissue mass transfer.

Self-heated thermistors have been used to measure perfusion [1, 10–12]. Effective thermal conductivity, k_{eff}, is the total ability of perfused tissue to transfer heat in the steady state. k_{eff} is the combination of conduction (due to intrinsic thermal conductivity, k_m) and convection (due to perfusion). Measurements of k_{eff} are very sensitive to perfusion. The limitation of most techniques is that empirical calibrations must be performed in order to accurately measure perfusion.

12.2.2 Constant Temperature Heating Technique

To measure the baseline tissue temperature, T_0, the instrument first places all three relays into the opposite position as shown in Fig. 12.4. This creates a resistance bridge similar to Fig. 11.4. Then, all three relays are activated and an electronic feedback circuit applies a variable power, $P(t) = V_0^2/R_{\text{set}}$, in order to maintain the average thermistor temperature at a predefined constant, T_h. The power required to maintain a constant temperature heating includes a steady state term and a transient term. Theoretically, the transient response is defined with an $\text{erf}(t^{-1/2})$, similar to Eq. (12.8), but after the first 1 second of heating, we can accurately approximate the applied thermistor power as

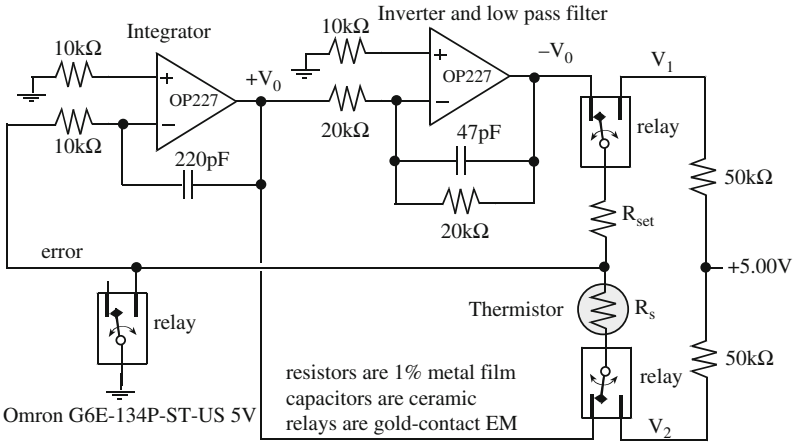


Fig. 12.4 Electronic controller used to measure thermal properties and perfusion

$$P(t) = A + Bt^{-1/2} \tag{12.15}$$

The erf($t^{-1/2}$) behavior, approximated by $t^{-1/2}$, arises from the spherical geometry, and incorporates the fact that the electronic controller must deposit more power during the initial transient phase to maintain the average temperature rise in the thermistor. The feedback circuit will produce whatever power is needed to maintain the temperature rise; steady state power, A, is used to measure thermal conductivity, while the transient response, B/A, measures thermal diffusivity. The R_{set} resistor in Fig. 12.4 determines the heated temperature. Since the feedback circuit drives the thermistor resistance equal to R_{set}, the heated temperature, T_h, will be equal to the equivalent thermistor temperature when its resistance is R_{set}.

In order to measure thermal conductivity, thermal diffusivity, and tissue perfusion the relationship between applied thermistor power, $P(t)$, and resulting thermistor temperature rise, $\Delta T(t)$ equals $T_h - T_0$ must be known. In the constant temperature method, the ΔT is constant because of the feedback circuit in Fig. 12.4. The thermistor bead is treated as a sphere of radius "a" embedded in an infinite homogeneous medium.

Since all media are considered to have constant parameters with respect to time and space, the initial temperature will be uniform when no power is supplied to the probe. Assume both T_b and T_m equal the initial temperature, T_0, at time 0. Assuming the venous blood temperature equilibrates with the tissue temperature and that the metabolic heat is uniform in time and space, the Pennes' bioheat transfer equation in spherical coordinates is given by [12, 13].

$$\rho_b c_b \frac{\partial V_b}{\partial t} = k_b \frac{1}{r^2} \frac{\partial}{\partial r}\left(r^2 \frac{\partial T_b}{\partial r}\right) + \frac{A + Bt^{-1/2}}{\frac{4}{3}\pi a^3} \qquad r < a \tag{12.16}$$

$$\rho_m c_m \frac{\partial V_m}{\partial t} = k_m \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial T_m}{\partial r} \right) - w\, c_{bl} (T_m - T_0) \qquad r > a \tag{12.17}$$

where subscript b refers to the thermistor bead, subscript m refers to the tissue medium, and subscript bl refers to the blood. Perfect thermal contact is assumed between the finite-sized spherical thermistor and the infinite homogeneous perfused tissue. At the interface between the bead and the tissue, continuity of thermal flux and temperature leads to the following boundary conditions:

$$T_b = T_m \qquad \text{at } r = a \tag{12.18}$$

$$k_b \frac{\partial T_b}{\partial r} = k_m \frac{\partial T_m}{\partial r} \qquad \text{at } r = a \tag{12.19}$$

The other boundary conditions are necessary at positions $r \to 0$ and $r \to \infty$. Since no heat is gained or lost at the center of the thermistor, the center temperature must be finite:

$$k_b \frac{\partial T_b}{\partial r} = 0 \quad \text{as } r \to 0 \tag{12.20}$$

Because the thermistor power is finite and the tissue is infinite, the tissue temperature rise at infinity goes to zero:

$$T_m \to T_0 \text{ as } r \to \infty \tag{12.21}$$

It is this last initial condition that allows the Laplace transform to be used to solve the coupled partial differential equations. The Laplace transform converts the partial differential equations into ordinary differential equations which are independent of time t. The steady state solution allows for the determination of thermal conductivity and perfusion [12].

$$T_b(r) = T_0 + \frac{A}{4\pi a k_b} \left\{ \frac{k_b}{k_m(1+\sqrt{z})} \right\} + \frac{1}{2} \left[1 - \left(\frac{r}{a} \right)^2 \right] \tag{12.22}$$

$$T_m(\mathrm{r}) = T_0 + \frac{A}{4\pi r k_m} \left(\frac{e^{(1-r/a)\sqrt{z}}}{1+\sqrt{z}} \right) \tag{12.23}$$

where z is a dimensionless Pennes' model perfusion term ($w\, c_{bl} a^2 / k_m$). The measured thermistor response, ΔT, is assumed be the simple volume average of the thermistor temperature:

$$\Delta T = \frac{\int_0^a T_b(r)\, 4\pi\, r^2\, dr}{\frac{4}{3}\pi\, a^3} - T_0 \tag{12.24}$$

Inserting Eq. (12.22) into (12.24) yields the relationship used to measure thermal conductivity assuming no perfusion [1].

$$k_m = \frac{1}{\frac{4\pi a?T}{A} - \frac{0.2}{k_b}} \tag{12.25}$$

A similar equation allows the measurement of thermal diffusivity from the transient response, again assuming no perfusion [12].

$$\alpha_m = \left[\frac{a}{\sqrt{\pi} B/A \left(1 + 0.2\frac{k_m}{k_b}\right)} \right] \tag{12.26}$$

Rather than using the actual probe radius (a) and probe thermal conductivity (k_b), the following empirical equations are used to calculate thermal properties.

$$k_b = \frac{1}{\frac{c_1 \Delta T}{A} + c_2} \tag{12.27}$$

$$\alpha_b = \left[\frac{c_3}{\sqrt{\pi} B/A \left(1 + 0.2\frac{k_m}{c_4}\right)} \right] 2 \tag{12.28}$$

The coefficients c_1, c_2, c_3, and c_4 are determined by operating the probe in two materials of known thermal properties. Typically agar-gelled water and glycerol are used as thermal standards. This empirical calibration is performed at the same temperatures at which the thermal property measurements will be performed.

It is assumed that the baseline tissue temperature, T_0, is constant during the transient heating interval. Patel has shown that if the temperature drift, dT_0/dt, is larger than 0.1°C/min, then significant errors will occur [15]. The electronic feedback circuit forces T_h to a constant. Thus, if T_0 is constant then ΔT does not vary during the transient heating interval.

First, the initial temperature T_0 is measured with the three relays in the opposite position as shown in Fig. 12.4. The heated temperature, T_h, is a fixed value determined by the R_{set} resistor. Thus, the fixed temperature rise, ΔT, is calculated as $T_h - T_0$. At time = 0, the three relays are flipped to the position shown in Fig. 12.4, which activates the feedback controller self-heating the thermistor. Applied power, $P(t)$, is measured during the heating interval. At the end of the heating interval the three relays are flipped back, placing the system back in temperature sensing mode. Figure 12.5 shows some typical responses for a 30 s heating interval. Linear regression is used to calculate the steady state and transient terms in Eq. (12.15). The steady state response (time equals infinity) is a measure of the thermal conductivity. The transient response (slope) indicates the thermal diffusivity.

The choice of how long to self-heat the thermistor involves a tradeoff. Shorter heating times are better for small tissue samples and for situations where there

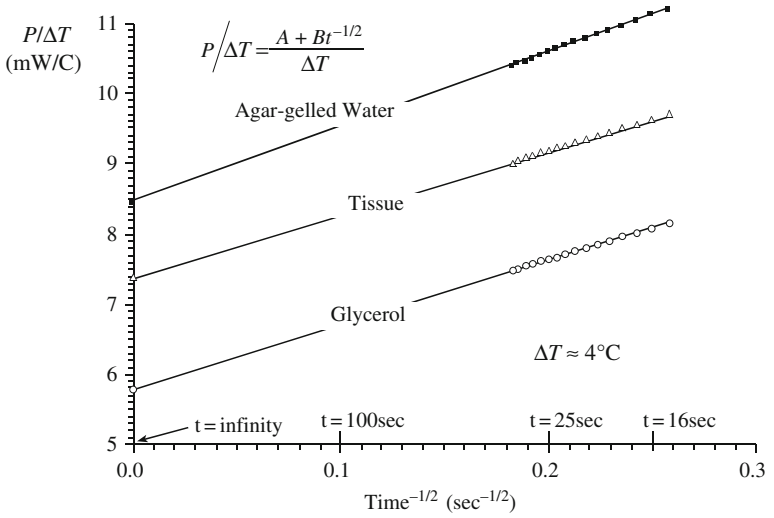


Fig. 12.5 Typical $P/\Delta T$ versus time $^{-1/2}$ data for the constant temperature heating technique

is baseline tissue temperature drift. Another advantage of shorter heating times is the reduction in the total time required to make one measurement. Longer heating times increase the measurement volume and reduce the effect of imperfect thermistor/tissue coupling. Typically, shorter heating times are used in vivo because it allows more measurements to be taken over the same time period. On the other hand, longer heating times are used in vitro because accuracy is more important than measurement speed. For thermistors with diameters of 0.5–1 mm, heating intervals of 10–60 s obtain good results.

12.2.3 Probe Design

Thermal probes must be constructed in order to measure thermal properties. The two important factors for the thermal probe are thermal contact and transducer sensitivity. The shape of the probe should be chosen in order to minimize trauma during insertion. Any boundary layer between the thermistor and the tissue of interest will cause a significant measurement error. The second factor is transducer sensitivity, which is the slope of the thermistor voltage versus tissue thermal conductivity. Equation (12.25) shows for a fixed ΔT k_m and k_b, the thermistor power (A) increases linearly with probe size (a). Therefore larger probes are more sensitive to thermal conductivity. It is important to use thermistors with a resistance of 1 $k\Omega$ at 25°C, so the electronics will be able to self-heat it easily. GE P60BA102M and Fenwal 121-102EAJ-Q01 are glass probe thermistors that make excellent transducers (diameter is about 1.5 mm). The glass coated spherical probes provide a large bead size and a rugged, stable transducer. The GE BR42KA102M bead thermistors also provide excellent results (diameter is about 1.1 mm). For large tissue samples

multiple thermistors can be wired in parallel, so they act electrically and thermally as one large device. There are two advantages to using multiple thermistors. The effective radius, $a = c_1/4\pi$, is increased from about 0.1 cm for a typical single P60DA102M probe to about 0.5 cm for a configuration of three P60DA102M thermistors. The second advantage is that the three thermistors are close enough to each other that the tissue between the probes will be heated by all three thermistors. This cooperative heating tends to increase the effective measurement volume and reduce the probe/tissue contact error. Good mechanical/thermal contact is critical. The probes are calibrated after they are constructed, so that the thermistor geometry is incorporated into the coefficients c_1, c_2, c_3, and c_4. The same water bath and probe configuration should be used during the calibration and during the tissue measurements.

12.2.4 Calibration

Calibration is a critical factor when using an empirical technique. For temperatures below 0°C, ice and ethylene glycol are used as thermal standards. For temperatures between 0 and 15°C, agar-gelled water and ethylene glycol can be used as thermal standards. For temperatures between 15 and 75°C, agar-gelled water and glycerol are suitable. One gram of agar per 100 mL of water should be added to prevent convection. For temperatures up to 150°C, ethylene glycol and glycerol are suitable. The instrument has been used to measure k_m and α_m of various concentrations of agar-gelled water at 37°C. Table 12.1 shows that the agar concentration does not affect thermal properties.

Each probe is heated 10 times in each medium at the temperatures of interest. The average $\Delta T/A$ and B/A for each set is recorded. Equations (12.15) and (12.16) are used to calculate the coefficients c_1, c_2, c_3, and c_4 knowing the true thermal properties of water and ethylene glycol. Appendix 1 contains thermal properties of the reference media.

A mixture of water and glycerol can be used to estimate the accuracy of the technique. The mass fraction, m, can be used to determine the true thermal properties of the mixture [16].

$$k_m = m\,kg + (1-m)k_w + 1.4\,m\,(m-1)(k_w - k_g - 2) - 0.014\,m(m-1)(\mathrm{T} - 20°\mathrm{C}) \tag{12.29}$$

Table 12.1 Thermal properties of various concentrations of agar-gelled water

Grams of agar/100 mL of water	k_m(mW/cm-°C)	α_m(1000*cm^2/ sec)
0.5	6.16 ±0.01	1.53 ±0.02
0.5	6.13 ±0.01	1.54 ±0.02
1.0	6.15 ±0.02	1.54 ±0.03
1.0	6.15 ±0.02	1.54 ±0.02
1.5	6.16 ±0.01	1.55 ±0.03
2.0	6.11 ±0.03	1.54 ±0.02

Table 12.2 Measured conductivity versus true conductivity for 5 media with known thermal conductivity

True k_m(mW/cm-°C)	k_m(mW/cm-°C)	Error (%)
6.23	6.20	–0.5
5.52	5.39	–2.2
4.41	4.48	1.4
3.78	3.85	1.9
2.89	2.87	–0.9

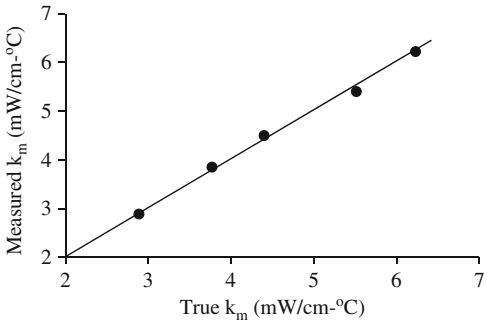


Fig. 12.6 Measured versus true thermal conductivity for 5 media with known thermal conductivity using the constant temperature heating technique

$$\alpha_m = m\,\alpha_g + (1 - m)\,\alpha_w \tag{12.30}$$

Equations (12.29) and (12.30) were empirically determined, and the nonlinear terms result from the chemical interactions between the water and glycerol molecules. Agar-gelled water (m=0), glycerol (m=1) and 3 mixtures ($0 < m < 1$) of the two were used to estimate the accuracy of the thermal conductivity measurement. The average conductivity measurement error was 1.4%, and the average diffusivity measurement error was 4%. The results are shown in Table 12.2, and plotted in Fig. 12.6.

12.3 Temperature Dependent Thermal Properties

12.3.1 Temperature Dependence of Organ Tissue

When modeling heat transfer in situations where the temperature range exceeds 10°C, it is important to consider the temperature dependence of the tissue thermal properties. Valvano measured tissue thermal properties as a function of temperature using the constant ΔT thermistor heating technique [9, 17]. The results shown in Table 12.3 were derived from in vitro measurements taken at 3, 10 17, 23, 30, 37, and 45°C.

The animal tissues were measured from freshly sacrificed dogs, rabbits, and pigs. The normal human tissues were obtained from autopsy. The human cancers

Table 12.3 Parameters of Eqs. (12.31) and (12.32) for estimating thermal properties as a function of temperature [9]

		k_0	k_1	α_0	α_1
Tissue	Species	mW/cm°C	mW/cm-°C^2	cm^2/s	cm^2/s-°C
Adenocarinoma of the Breast	Human	4.194	0.03911	0.001617	–0.000049
Cerebral Cortex	Human	5.043	0.00296	0.001283	0.000050
Colon Cancer	Human	5.450	(at 19°C)	0.001349	(at 19°C)
Fat of Spleen	Human	3.431	–0.00254	0.001321	–0.000002
Liver	Human	4.692	0.01161	0.001279	0.000036
Liver	Pig	4.981	0.00800	0.001240	0.000053
Liver	Rabbit	4.668	0.02601	0.001370	0.000178
Lung	Human	3.080	0.02395	0.001071	0.000082
Lung	Human	4.071	0.01176	0.001192	0.000031
Lung	Pig	2.339	0.02216	0.000695	0.000080
Myocardium	Dog	4.869	0.01332	0.001296	0.000058
Myocardium	Human	4.925	0.01195	0.001289	0.000050
Myocardium	Pig	4.841	0.01333	0.001270	0.000051
Pancreas	Dog	4.790	0.00849	0.001287	0.000062
Pancreas	Human	4.365	0.02844	0.001391	0.000084
Pancreas	Pig	4.700	0.00194	0.001530	0.000130
Renal Cortex	Dog	4.905	0.01280	0.001333	0.000039
Renal Cortex	Human	4.989	0.01288	0.001266	0.000055
Renal Cortex	Pig	4.967	0.01176	0.001284	0.000039
Renal Cortex	Rabbit	4.945	0.01345	0.001311	0.000027
Renal Medulla	Dog	5.065	0.01298	0.001305	0.000063
Renal Medulla	Human	4.994	0.01102	0.001278	0.000055
Renal Pelvis	Dog	4.930	0.01055	0.001334	0.000052
Renal Pelvis	Human	4.795	0.01923	0.001329	0.000011
Spleen	Human	4.913	0.01300	0.001270	0.000047
Spleen	Rabbit	4.863	0.01267	0.001257	0.000042

were freshly excised. The k_0, k_1, α_0, and α_1 values are the linear fit of the thermal properties as a function of temperature.

$$k = k_0 + k_1 T \tag{12.31}$$

$$\alpha = \alpha_0 + \alpha_1 T \tag{12.32}$$

where temperature T is in °C. The average k_1 is +0.014 mW/cm-°C per °C. Similarly, the average α_1 is +0.000053 cm^2/s per °C.

12.3.2 Temperature Dependence of Human Arterial Tissue

Aortic tissue was obtained from a local pathology lab. Surface probes were constructed by placing a spherical thermistor at the edge of an insulating surface (rubber

Table 12.4 Thermal conductivity (mW/cm-°C) of human aorta and atherosclerotic plaque

Tissue	*n*	at 35°C	at 55°C	at 75°C	at 90°C
Normal aorta	12	4.76 (0.41)	5.03 (0.60)	5.59 (0.37)	6.12 (0.12)
Soft plaque	25	4.84 (0.33)	5.02 (0.40)	5.42 (0.45)	5.83 (0.69)
Hard plaque	17	5.02 (0.59)	5.26 (0.73)	5.81 (0.82)	6.19 (0.85)
Water	[4, 25]	6.21	6.46	6.66	6.76

Table 12.5 Thermal diffusivity (*1000 cm^2/s) of human aorta and atherosclerotic plaque

Tissue	*n*	at 35°C	at 55°C	at 75°C	at 90°C
Normal aorta	12	1.27 (0.07)	1.33 (0.11)	1.44 (0.10)	1.56 (0.05)
Soft plaque	25	1.28 (0.05)	1.34 (0.07)	1.41 (0.11)	1.48 (0.15)
Hard plaque	17	1.32 (0.07)	1.37 (0.12)	1.53 (0.17)	1.66 (0.20)
Water	[4, 25]	1.50	1.57	1.61	–

cork). The thermistor protruded slightly below the surface so that about 30–50% of the thermistor surface came in direct contact with the tissue. The thermal probes were placed on the endothelial surface of the aortic wall, and the tissue/probe combination was wrapped in plastic. The tissue surface was kept wet to improve the thermal contact and to prevent drying. The samples were placed in a breaker of saline and the breaker was put into a temperature controlled water bath. Thermal conductivity and thermal diffusivity were measured ten times at each temperature 35, 55, 75, and 90°C. The measurement order was varied between 35, 55, 75, 90 95, 75, 55, 35 75, 90, 55, 35 and 55, 35, 90, 75. Measurements were obtained from both normal and diseased tissue. The plaques were categorized by gross visual observation. The "hard" plaques were firm and bony. The "soft" plaques were pliable, loose and buttery. The results from 54 tissues are presented in Tables 12.4 and 12.5. The column *n* refers to the number of tissue samples. The standard deviation is given in the parentheses. The values for water were interpolated from the standard parameters shown in Appendix 1 [4, 25].

The two sample t-test with $p = 0.05$ was used to determine significant differences. The tissue thermal properties increased with temperature. Thermal property values for all tissues were less than water. The measurement order did not affect the measured thermal properties. There was no difference between the thermal conductivity of normal aorta, and soft plaque. The thermal conductivity and thermal diffusivity of hard plaque were slightly higher than normal aorta and soft plaque.

12.3.3 Temperature Dependence of Canine Arterial Tissue

Carotid and femoral arteries were harvested immediately post mortem. The thermal probes were placed on the endothelial surface of the arterial wall. Thermal conductivity and thermal diffusivity were measured ten times at each temperature 25, 35,

45, 55, 65, 75, 85, and 95°C. Measurements were obtained only from normal tissue. The results from 18 tissues are summarized in Eqs. (12.33)–(12.36) [2].

Canine femoral artery

$$k\,(\text{mW/cm-°C}) = 3.688 + 0.0062014\ \text{T(°C)} \tag{12.33}$$

$$\alpha\,(\text{cm}^2/\text{s}) = 0.001003 + 0.000001381\ \text{T(°C)} \tag{12.34}$$

Canine carotid artery

$$k\,(\text{mW/cm-°C}) = 4.480 + 0.0000164\ \text{T(°C)} \tag{12.35}$$

$$\alpha\,(\text{cm}^2/\text{s}) = 0.001159 + 0.00000396\ \text{T(°C)} \tag{12.36}$$

The two sample t-test with $p = 0.01$ shows that both thermal conductivity and thermal diffusivity are larger in carotid versus femoral artery. These results could be explained from the fact that the carotid artery contains more collagen than the femoral artery. A tissue with a higher percentage of collagen will have lower thermal properties because collagen is a thermal insulator.

12.3.4 Temperature Dependence of Swine Myocardial Tissue

Swine myocardial samples were harvested immediately post mortem. The thermal probes were placed on the left ventricular muscle. Thermal conductivity and thermal diffusivity were measured ten times at each temperature 25, 37, 50, 62 and 76°C. Measurements were obtained only from normal tissue. The results are summarized in Tables 12.6 and 12.7.

Table 12.6 Thermal conductivity (mW/cm-°C) of myocardial tissue [2]

Temperature	25°C	37°C	50°C	62°C	76°C
	5.23	5.14	5.17	4.39	5.24
	5.07	5.12	4.75	3.30	4.29
	5.30	5.21	5.61	5.67	4.83
	5.43	5.54	4.22	4.16	5.89
	4.68	5.35	4.93	5.33	5.23
	5.25	5.08	4.84	5.70	5.39
	5.27	5.48	4.42	5.11	4.75
	5.28	4.57	4.93	4.99	3.25
	5.86	5.76	5.52	5.03	2.69
	4.78	5.10	5.88	5.30	5.28
	4.75	5.35	5.35	4.67	5.60
	4.92	6.02	5.60	5.49	4.68
Mean	5.15	5.31	5.1	4.93	4.76
Std. Dev.	0.33	0.37	0.51	0.70	0.95

Table 12.7 Thermal diffusivity (cm^2/s) of myocardial tissue [2]

Temperature	25°C	37°C	50°C	62°C	76°C
	0.00151	0.00170	0.00165	0.00159	0.00167
	0.00154	0.00147	0.00203	0.00235	0.00249
	0.00143	0.00165	0.00151	0.00169	0.00166
	0.00146	0.00143	0.00116	0.00191	0.00229
	0.00159	0.00160	0.00176	0.00167	0.00173
	0.00141	0.00178	0.00179	0.00163	0.00185
	0.00165	0.00149	0.00235	0.00143	0.00185
	0.00132	0.00206	0.00179	0.00170	0.00199
	0.00141	0.00144	0.00147	0.00143	0.00062
	0.00168	0.00179	0.00160	0.00180	0.00167
	0.00154	0.00156	0.00173	0.00161	0.00173
	0.00164	0.00138	0.00171	0.00169	0.00192
Mean	0.00152	0.00161	0.00171	0.00171	0.00179
Std. Dev.	0.00012	0.00020	0.00031	0.00025	0.00047

12.3.5 Thermal Properties of Frozen Tissue

The thermal properties of frozen tissue are significantly different from normal tissue. Valvano measured frozen tissue thermal properties using the constant ΔT thermistor heating technique [18]. The results shown in Table 12.8 were derived from in vitro measurements taken at –18, –5, and +0.1°C.

N is the number of measurements and *M* is the number of tissues. (±0.xx) is the standard deviation of the average [18]

Table 12.8 Average thermal properties

T(°C)	N	M	Species	Tissue	k_m(mW/cm-°C)	α_m(100*cm^2/s)
+0.1	45	6	Bovine Kidney Cortex	4.54 (±0.16)	1.18 (±0.09)	
−5	15	4	Bovine Kidney Cortex	15.35 (±1.09)	4.71 (±0.99)	
−18	18	3	Bovine Kidney Cortex	13.72 (±0.73)	6.84 (±0.83)	
+0.1	66	9	Bovine Liver	4.17 (±0.13)	1.05 (±0.09)	
−5	66	9	Bovine Liver	13.96 (±2.49)	4.77 (±0.58)	
−18	56	8	Bovine Liver	9.89 (±0.44)	5.71 (±0.74)	
+0.1	48	6	Bovine Muscle	4.25 (±0.37)	1.05 (±0.11)	
−5	42	7	Bovine Muscle	13.93 (±1.23)	5.37 (±0.97)	
−18	60	8	Bovine Muscle	10.76 (±1.14)	6.84 (±1.10)	
+0.1	21	3	Bovine Fat	1.93 (±0.12)	0.59 (±0.13)	
−5	32	4	Bovine Fat	2.66 (±0.38)	0.98 (±0.19)	
−18	24	4	Bovine Fat	2.80 (±0.53)	1.54 (±0.57)	

12.4 Thermal Properties as a Function of Water and Fat Content

The most important consideration of thermal properties when developing laser techniques is water content. As tissue is heated, significant changes in thermal properties will occur as water is lost. In a global sense, the thermal properties of tissue

are determined by the relative concentrations of its constituent parts. Spells found a linear relationship between tissue thermal conductivity and water content [19]:

$$k\,(\text{mW/cm-}^\circ\text{C}) = 0.54 + 5.73\,m_{\text{water}} \quad \text{for } m_{\text{water}} > 0.2 \tag{12.37}$$

where m_{water} is the mass fraction of water in the tissue.

Cooper and Trezek found an empirical relationship between thermal conductivity and mass fractions of water, protein and fat [20].

$$k\,(\text{mW/cm-}^\circ\text{C}) = \rho \sum_n \frac{k_n m_n}{\rho_n} = \rho = \left(6.28 m_{\text{water}} + 1.17 m_{\text{protein}} + 2.31 m_{\text{fat}}\right) \tag{12.38}$$

Cooper and Trezek (1971) found similar relationships for specific heat and density.

$$c\,(\text{J/g-}^\circ\text{C}) = \sum_n c_n m_n = 4.2\,m_{\text{water}} + 1.09\,m_{\text{protein}} + 2.3\,m_{\text{fat}} \tag{12.39}$$

$$\rho\,(g/\text{cm}^3) = \frac{1}{\sum\limits_n \frac{m_n}{\rho_n}} = \frac{1}{m_{\text{water}} + 0.649 m_{\text{protein}} + 1.227 m_{\text{fat}}} \tag{12.40}$$

12.5 Measurements of Perfusion

12.5.1 Introduction

Perfusion, the transmission of blood in the microcirculation, is an important factor in surgery, tissue transplants, heart disease and cancer therapy [1–3]. Despite its importance, no clinical method of measuring perfusion is currently available for a majority of applications. One technique that shows considerable promise involves the use of self-heated thermistors [1–3]. In this method, a miniature thermistor (0.5–2.5 mm diameter) is placed invasively in the tissue of interest and heated with a predetermined applied power. Since both tissue conduction and perfusion act to carry heat away from the thermistor, the resulting volumetric-average temperature rise in the thermistor bead, ΔT, is related to both the tissue thermal conductivity and perfusion. By knowing the intrinsic tissue conductivity and the effective conductivity of the tissue (due to both blood flow and conduction), the perfusion rate can be calculated.

At least two difficulties exist with this technique. The first is that the intrinsic tissue conductivity of perfused tissue must be known in order to calculate the perfusion rate. Although one could stop the blood flow to a tissue and measure its conductivity, this is clearly not desirable, nor is it always practical. The second difficulty is that there is considerable debate concerning which heat transfer model best describes tissue blood flow. A model suggested by Pennes has been widely adopted by researchers to model the heat transfer effects of perfusion [13]. k_{eff} is defined to the measured thermal conductivity in the presence of perfusion, and k_m is defined

to the intrinsic thermal conductivity in the absence of perfusion. Assuming Pennes' model, Eq. (12.23) is rearranged to determine perfusion [12].

$$w = \frac{(k_{\text{eff}} - k_m)^2}{k_m c_{bl} a^2} \tag{12.41}$$

The effective thermal conductivity, k_{eff} is measured using constant ΔT heating Eq. (12.14). k_m can be measured at the same location under no flow conditions, or it can be measured in vitro in tissues of the same type and species.

The validity of this model has been disputed. Weinbaum and Jiji proposed a perfusion model in which countercurrent heat exchange between artery-vein pairs plays an important role [2]. Patel and Valvano have developed a model to describe the heat transfer away from a thermal point source due to perfusion [14, 21]. In this model, heat transfer due to blood flow and intrinsic tissue conductivity are lumped together into an effective conductivity term, k_{eff}. Perfusion is approximately linearly related to the difference in effective and intrinsic conductivities:

$$w = c_5(k_{\text{eff}} - k_m) \tag{12.42}$$

where w is perfusion, k_m is intrinsic tissue conductivity, and c_5 is a calibration coefficient found experimentally by operating the probe in a tissue of known perfusion. Anderson and Valvano proposed another model of perfusion in the kidney cortex in which blood flow is modeled with discrete arterioles and venules (50–75 μm in diameter) [22]. With this small artery model, the slope of the k_{eff} curve increases with increasing perfusion rate. As the arteries and veins are spaced closer together in the tissue, perfusion becomes more linearly related to k_{eff}–k_m.

12.5.2 Alcohol-Fixed Canine Kidney Experiments

A microcomputer based instrument was used to measure perfusion in alcohol fixed canine kidneys. Experiments were performed with P60DA102N thermistors. Calibration coefficients were determined by operating the probe in agar-gelled water and glycerol. The applied power resulted in a thermistor surface temperature rise of no more than 5°C and thus prevented thermal damage to the fixed kidney. The thermistor probes were then embedded in alcohol-fixed canine kidneys and the kidneys were perfused with deionized water. k_{eff} were measured at several different perfusion rates.

Perfusion rate was calculated by the dye injection procedure reported by Valvano et al. [14]. "Evans Blue" dye was injected into the renal artery of the kidney. The kidney was then dissected into dyed and undyed components, with the kidney cortex typically being uniformly dyed and the rest of the kidney undyed. The dyed components were assumed to have uniform perfusion. The measured kidney flow rate was then divided by the dyed mass to determine the perfusion rate. Errors in this method

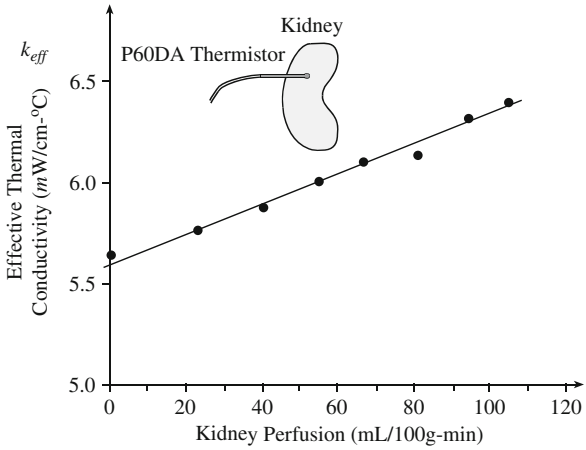


Fig. 12.7 Effective thermal conductivity versus kidney perfusion for a P60 thermistor

occurred mainly due to tissue fluid loss during dissection and difficulties in separating dyed and undyed tissues. A total possible error in the absolute value of perfusion was about 30%. The shape of the k_{eff} versus perfusion curve was not dependent on the uncertainty of perfused mass, but depended only on the uncertainty in the kidney flow. This uncertainty was estimated to be less than 1%.

Figure 12.7 plots tissue conductivity versus perfusion rate for a P60 thermistor. k_{eff} appears almost linearly related to perfusion, with just a slightly increasing slope in the k_{eff} curve with increasing perfusion.

12.5.3 Canine Muscle Experiments

The overall purpose of the experiment was to evaluate surgical and pharmacologic support methods for animals in severe shock. The dog was anesthetized and hypovolemic shock was induced by removing a significant amount of the blood volume. The animal was then supported using an aortic balloon pump. At the end of the experiment, the blood volume was returned to the dog. The animal was sacrificed in order to perform microsphere counting.

A secondary purpose was to evaluate the self-heated thermistor measurements of perfusion. The constant ΔT technique was used. Equations (12.15) and (12.25) were used to calculate effective thermal conductivity, k_{eff}. The linear Eq. (12.42) was used to measure perfusion. The probe was located in the left shoulder muscle. Figure 12.8 plots measured perfusion versus time. Perfusion dropped sharply after shock was induced, and remained low even with the aortic balloon. A large hyperemia was observed after the blood volume was returned. The microsphere perfusion measurements are also shown in Fig. 12.12. This experiment illustrates the differences between microsphere and thermistor perfusion measurements. Microspheres provide considerable information concerning the spatial distribution of perfusion,

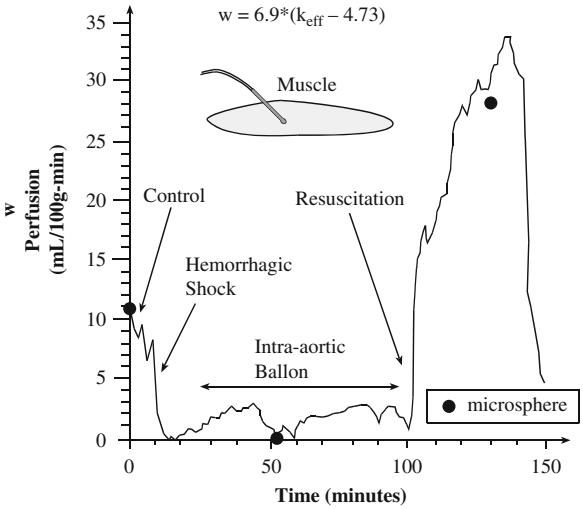


Fig. 12.8 Perfusion measured in canine skeletal muscle

but at only a few times. One the other hand, thermistors provide continuous perfusion measurements, but only at a few locations. The thermistor technique can be used clinically because it is not necessary to harvest the tissue. Another advantage of the thermistor technique is the perfusion values are available in real time, allowing the researcher to evaluate the tissue status during the experimental protocol.

12.5.4 Trauma at Insertion Site will Decouple Probe from Tissue

The measurement volume of a thermistor is a complex function of many factors, including the perfusion rate and vascular anatomy of the tissue of interest. A boundary layer (decoupler) between the thermistor and the tissue causes a significant measurement error. This unwanted boundary layer is often caused by the probe itself during insertion. A 1-D finite element model was used to quantify this relationship [2, 15]. Figure 12.9 plots the conductivity error as a function of the decoupler thickness. In this case, the decoupler conductivity was 4 mW/cm-°C, representing dead (and dry) fibrous tissue that could develop around the probe after long term implantation. Short-term trauma is typified by bleeding at the insertion sight and could have been modeled using blood conductivity. The larger probes exhibit a smaller error, but are likely to cause a larger decoupler because of the increased trauma during insertion.

12.5.5 Temperature Dependence of Perfusion

Perfusion depends on a wide variety of factors, some local to the tissue (pH, temperature, O_2), some which are external but directly control local flow (parasympathetic,

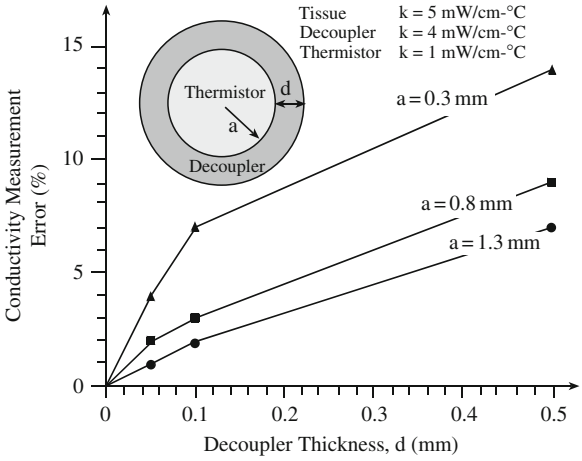


Fig. 12.9 Finite element results of the measurement error versus decoupler thickness [2, 15]

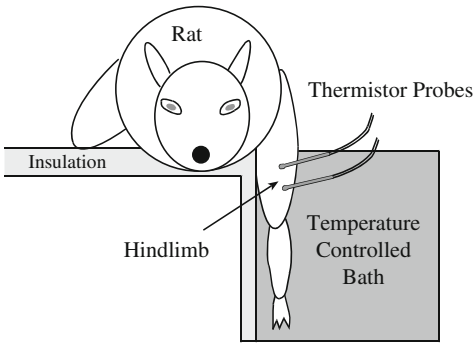


Fig. 12.10 Schematic of the heated rat hindlimb experiment

hormones), and some which indirectly affect local flow (heart rate, blood pressure, skin temperature, needs of other organs. A simple experiment studied the effect of local tissue temperature on muscle perfusion. The constant ΔT method was used to measure perfusion in an anesthetized rat. The muscle temperature was manipulated by placing the hindlimb into a water bath, see Fig. 12.10. Insulation was carefully placed so as to minimize changes to the body temperature.

Figure 12.11 shows the water bath temperatures during a typical experiment. Independent measurements confirmed that the rectal and neck temperatures were constant while the muscle in the hindlimb was heated.

Figure 12.12 shows the perfusion as a function of tissue temperature for a typical experiment. The dip in perfusion as a function of temperature as shown in Fig. 12.12 consistently occurred, but did not always occur at the same temperature. This dip may be due to an anastomotic shunt attempting to regulate the core body temperature. A linear fit to this averaged data gives the following approximation.

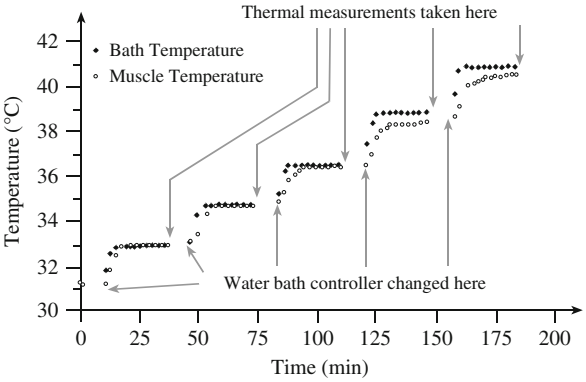


Fig. 12.11 Variation in water bath and muscle temperature for the entire protocol

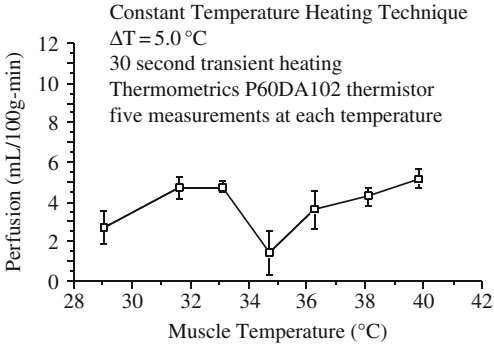


Fig. 12.12 Perfusion versus muscle temperature during a typical experiment

$$w = 1.9720(1 + 0.059\,T) \tag{12.43}$$

where w has units of mL/100 g-min and T has units of °C.

Yuan et al. [23] measured perfusion and temperatures at various locations within each of four canine prostates subjected to a transurethral microwave thermal source. The total number of the perfusion sampling points coupled with temperature is 15. Colored microspheres were used to measure perfusion. Temperatures were measured using miniature thermistors. The prostate temperatures were raised to 40–45°C by 5 W step increments of the microwave power at hourly intervals to 15 W. Temperatures and perfusion were measured at baseline, and at the beginning and end of each heating interval. Thus, the periods between perfusion samples were approximately either 5 or 60 min. Under baseline conditions, the temperature fluctuations within the prostate were approximately ± 0.3°C. A relative dispersion estimate of 15% was derived from one dog for the fluctuations in baseline perfusion. Thus, changes in absolute perfusion and temperature greater than 15% and 0.3°C, respectively, were considered to be substantial changes.

As heating progressed, a variety of substantial changes were observed, but no uniform pattern emerged. However, the measurements included changes typically

expected for hyperthermia: (1) an initial perfusion increase associated with elevating the baseline temperature, (2) a perfusion return towards baseline after this initial increase, and (3) a dramatic increase in perfusion at elevated temperatures. The initial perfusion increases were observed in three dogs when the temperatures exceeded 38 ± 3°C (mean ± 1 s.d., $N = 8$). The perfusion increased 34% from a baseline value of 59 ± 26 mL/100 g-min over a temperature rise of 1.7 ± 1.3°C. Half of the measurements in the three dogs subsequently showed a decrease in perfusion ranging from 16 to 25%. In two dogs, dramatic perfusion increases as high as 364% were observed with a corresponding decrease in tissue temperature. Figure 12.13 shows the changes as the experiment progressed from baseline conditions.

Xu et al. [24] measured perfusion using the pulse-decay self-heated thermistor technique in these dog prostates during the same transurethral microwave hyperthermia treatments. Interestingly in the exact same dogs at roughly the same locations, the perfusion response to temperature measured with the thermal technique was roughly linear with temperature as shown in Fig. 12.14.

The interdependence between perfusion and temperature was observed in these studies. Most notably, a decrease in tissue temperature associated with a dramatic increase of perfusion.

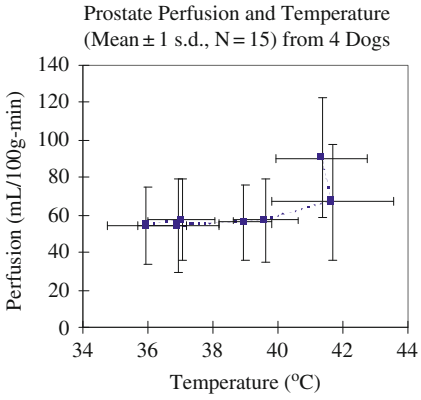


Fig. 12.13 Perfusion measured with microspheres versus prostate temperature [23]

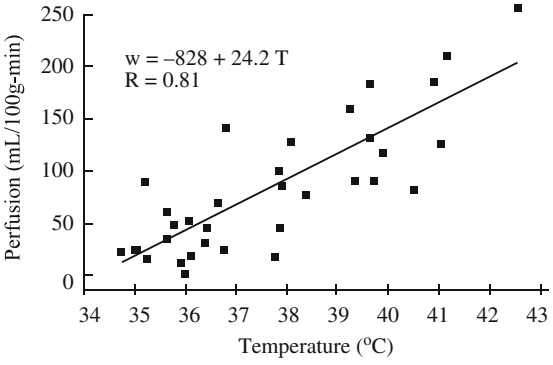


Fig. 12.14 Perfusion versus prostate temperature [24]

12.6 Conclusions

From a first approximation, biologic tissue behaves like a homogeneous solid with thermal properties that depend on its water content (see Eqs. (12.37)–(12.40)). Because water plays a dominant role governing thermal events in tissue, models that do not account for the loss or transport of water will be significantly flawed. Unless the tissue is severely altered, one can estimate water loss by weighing the tissue before, during, and after experimentation.

Temperature and perfusion have important, but secondary, effects on thermal properties. For most tissues, thermal properties increase with temperature. However, if the temperature rise causes a water loss, then thermal properties could significantly drop. When modeling heat transfer in situations where the temperature range exceeds 10°C, it is important to consider the temperature dependence of the tissue thermal properties. The average increase in thermal conductivity (Table 12.3) was +0.014 mW/cm-°C per °C. In this same data, the average increase in thermal diffusivity was +0.000053 cm^2/s per °C.

In living tissue, blood flow dramatically affects heat transfer. The goal of presenting perfusion data in this chapter was not to provide an exhaustive list of possible scenarios, from which you choose the one that best matches your experimental conditions. Rather these data describe the degree of importance and complexity involved in describing thermal effects of blood flow. For example, Fig. 12.7 shows perfusion in the kidney increases the thermal conductivity from about 5.5 mW/cm-°C to about 6.5 mW/cm-°C. Similarly, the data in Fig. 12.8 together with Eq. (12.42) indicate that the presence of normal perfusion in muscle increases the thermal conductivity from about 4.7 mW/cm-°C to about 6.2 mW/cm-°C. In summary, tissue with blood flow can transport more heat than tissue without blood flow.

Blood velocity varies considerably throughout the vascular network. Blood traveling slowly through capillaries has the opportunity to completely thermally equilibrate with the surrounding tissue. Models have shown that for vessels less than 100 μm in diameter significant heat transfer can occur between blood and tissue. It is relevant to consider the time it takes blood to enter an arteriole, pass through a capillary, and return via a venule. It takes about 20 s for blood to make a complete loop in the circulatory system. This means for time scales much less than 1 s, the effects of perfusion can be neglected. However, if one is interested in tissue temperature seconds after a short laser pulse, then perfusion must be considered.

The results in Figs. 12.12, 12.13, and 12.14 illustrate the complex interdependence of perfusion and temperature. For most living tissues, as temperature increases, perfusion will increase as well (possibility as a survival mechanism), which in turn causes the temperature to fall. It takes time for this behavior to develop, and the responses vary with temperature, duration, and tissue type. The perfusion temperature data in this chapter fall into the category of local tissue response, and there is an even more complex behavior when altering temperatures in the thorax or cranium. Equally complex is the interdependence of perfusion, water loss, and temperature occurring in the skin. For more details about these behaviors, see [2].

Appendix 1: Thermal Standards

Table 12.9 Thermal properties of water [4, 25]

Temperature (K)	Temperature (°C)	k (mW/cm-°C)	α(1000*cm^2/s)
280	6.85	5.818	1.35
290	16.85	5.918	1.42
300	26.85	6.084	1.46
310	36.85	6.233	
320	46.85	6.367	1.55
330	56.85	6.485	
340	66.85	6.587	1.61
350	76.85	6.673	
360	86.85	6.743	
370	96.85	6.797	

Table 12.10 Thermal properties of ethylene glycol [4, 7, 25]

Temperature (K)	Temperature (°C)	k(mW/cm-°C)	α(1000*cm^2/s)
260	–13.15	2.522	0.934
270	–3.15	2.536	
280	6.85	2.549	
290	16.85	2.563	
300	26.85	2.576	
310	36.85	2.590	
320	46.85	2.603	
330	56.85	2.616	
340	66.85	2.630	
350	76.85	2.643	
360	86.85	2.657	
370	96.85	2.670	
380	106.85	2.684	
390	116.85	2.697	
400	126.85	2.710	
410	136.85	2.724	
420	146.85	2.737	
430	156.85	2.751	

Table 12.11 Thermal properties of glycerol [4, 25]

Temperature (K)	Temperature (°C)	k(mW/cm-°C)	α(1000*cm^2/s)
294			0.9474
300	26.85	2.880	
310	36.85	2.893	
320	46.85	2.906	
330	56.85	2.919	
340	66.85	2.933	
350	76.85	2.946	
360	86.85	2.959	
370	96.85	2.972	
380	106.85	2.985	
390	116.85	2.999	
400	126.85	3.012	
410	136.85	3.025	
420	146.85	3.038	
430	156.85	3.051	

Appendix 2: Perfusion Values

Table 12.12 Perfusion measurements in mL/100 g-min

Organ	Tissue	Perfusion	$\pm\sigma$	Species
Brain	Mean Cerebral	37.3	6.2	dog
	Cerebral cortex	44.6	6.1	dog
	Cerebral cortex	44.7	3.6	dog
	Cerebral white	24	3.7	dog
	Cerebral white	30.5	2.2	dog
	Corpus callosum	24.9	3.5	dog
	Frontal cortex	106	6	rat
	Parietal cortex	107	5	rat
	Occipital cortex	98	8	rat
	Caudate nucleus	88	3	rat
	Hypothalamus	103	6	rat
	Thalamus	92	7	rat
	Hippocampus	68	6	rat
	Medulla	84	4	rat
	Corpus callosum	62	3	rat
	Cerebrum	48	3	cat
	Cerebrum	53	6	cat
	Cerebral cortex	64	5	cat
	Cerebral cortex	75	11	cat
	Caudate nucleus	65	6	cat
	Caudate nucleus	87	11	cat
	Cerebral white	29	3	cat
	Brain stem	52	3	cat
	Brain stem	58	3	cat
	Cerebellum	63	4	cat

Table 12.12 (continued)

Organ	Tissue	Perfusion	$\pm\sigma$	Species
	Cerebellum	70	7	cat
	Spinal cord	22	2	cat
	Spinal cord	31	3	cat
	Cerebrum	32	6	monkey
	Cerebrum	39	3	monkey
	Cerebral cortex	40	11	monkey
	Cerebral cortex	50	11	monkey
	Caudate nucleus	39	5	monkey
	Caudate nucleus	47	5	monkey
	Cerebral white	20	5	monkey
	Cerebral white	21	6	monkey
	Brain stem	33	6	monkey
	Brain stem	36	3	monkey
	Cerebellum	40	8	monkey
	Cerebellum	50	2	monkey
	Spinal cord	20	3	monkey
	Spinal cord	21	3	monkey
	Mean Cerebral	49.1	7.8	baboon
	Gray matter	56.2	7.7	baboon
	White matter	16.4	2.2	baboon
	Cerebrum	79	12	pig◊
	Cerebellum	78	6	pig◊
	Brain stem	46	6	pig◊
Kidney	Total	432	30	dog
	Outer cortex	632	33	dog
	Outer cortex	756	70	dog
	Outer cortex	462		dog
	Inner cortex	438	119	dog
	Inner cortex	335	42	dog
	Inner cortex	396		dog
	Medulla	77	6.4	dog
	Outer medulla	130		dog
	Outer medulla	210	28	dog
	Outer 1/4 cortex	424	40	dog
	Next 1/4	624	61	dog
	Next 1/4	355	45	dog
	Inner 1/4	173	13	dog
	Cortex	421	64	dog
	Total	320	30	rabbit
	Outer cortex	884	1.4	rabbit
	Middle cortex	953	1.3	rabbit
	Inner cortex	324	0.4	rabbit
	Cortex	400–500		human◊
	Outer medulla	120		human◊
	Inner medulla	25		human◊
	Cortex	363	54	pig
	Whole kidney	280	5	pig◊
	Whole kidney	470	30	rat
	Outer cortex	870	60	rat
	Inner cortex	470	40	rat

Table 12.12 (continued)

Organ	Tissue	Perfusion	$\pm\sigma$	Species
Myocardium	Left ventricle	145	9	pig◊
	Right ventricle	92	12	pig◊
	Left ventricle	158	13	dog
	Left ventricle	109	35	dog
	Left ventricle	61	41	dog
	Left ventricle	88	14	dog
	LV Epicardium	89	14	dog
	LV Epicardium	86	14	dog
	LV Endocardium	114	18	dog
	LV Endocardium	108	18	dog
Pancreas		0.8–1.6		dog
Muscle	gracilis	4.9	6.4	dog
	resting thigh 1.5 cm depth	1.8	1.1	human◊
	resting thigh 3.0 cm depth	2.6	1.3	human◊
	resting thigh	2.0–2.2		human◊
	skeletal, average	2.7		human◊
	vastus medialis and triceps	14	3	pig◊
	skeletal	9.4	5	dog
Intestine		38.9	4	dog
Spleen		146	14	dog
		11		dog
		360	16	pig◊
Liver	hepatic artery	24	4	pig
	hepatic artery	22	4	pig◊
		140	43	rat
		80	31	dog
Bone	Femur (plasma flow)	9.6	3.9	rat
	Tibia (plasma flow)	8.3	3.2	rat
	Femur (plasma flow)	14	1	pig◊
Skin (forearm)	in the cold	≈2		human◊
	thermoneutral	4 to 5		human◊
	hyperthermic	>20		human◊
	thermoneutral	90	9	pig◊
Fat		21	6	pig◊
Prostate	10 mm depth	58	3	dog
	10 mm depth	38	4	dog
	9 mm depth	19	4	dog
	9 mm depth	55	11	dog
	5 mm depth	15	9	dog
	4 mm depth	88	6	dog
		47	24	dog

All data except where noted with ◊ was taken under anesthesia. Most of this information was compiled by K.R. Holmes of the University of Illinois and distributed at the 1990 Allerton Workshop on Biothermal Engineering [2]

Acknowledgments Experiments shown in this chapter were funded in part by grants from the Whitaker Foundation, the NIH (GM31931-01A3), and the Texas Advanced Technology Program (#003658-570). Bob Dhaens performed the canine muscle experiments, Anil Paranjape performed the rat hindlimb experiments, Gary Anderson designed and executed the canine kidney protocol, Cheyenne Wu took the data in Figs. 12.4, and 12.5, and David Yuan performed the prostate experiments.

References

1. Bowman HF. Estimation of tissue blood flow. In: A Shitzer and RC Eberhart (eds) *Heat transfer in medicine and biology.* Plenum, New York pp. 193–230 (1985).
2. Diller KR, Valvano JW, and Pearce JA. Bioheat transfer. In: Kreith F and Y Goswami (eds) *CRC handbook of mechanical engineering*. CRC Press, Boca Raton, FL (2004).
3. Chato JC. Measurement of thermal properties of biological materials. In: A. Shitzer and RC Eberhart (eds) *Heat transfer in medicine and biology*. Plenum, New York pp. 167–192 (1985).
4. Touloukian YS et al. *Thermophysical properties of matter: Thermal conductivity*, Vo1. 3. IFI/Plenum, New York, pp. 120, 209 (1970).
5. Touloukian YS et al. *Thermophysical properties of matter: Thermal conductivity*, Vo1. 3. IFI/Plenum, New York, Preface (1970).
6. Touloukian YS et al. *Thermophysical properties of matter: Thermal diffusivity*, Vo1. 10. IFI/Plenum, New York, pp. 15a–42a (1973).
7. Eckert ERG and Drake RM Jr. *Heat and mass transfer*. McGraw-Hill, New York, 2nd edition (1959)
8. Duck FA. *Physical properties of tissue: A comprehensive reference book*. Academic, London (1991).
9. Valvano JW, Cochran JR, and Diller KR. Thermal conductivity and diffusivity of biomaterials measured with self-heated thermistors. *Int. J. Thermophys*., 6:301–311 (1985).
10. Chen M, Holmes KR, and Rupinskas V. Pulse-decay method for measuring the thermal conductivity of living tissue. *J. Biomech. Eng.*, 103:253–260 (1981).
11. Holmes KR and Chen MM. Local tissue heating, microbead pulse decay technique for heat transfer parameter evaluation. In: W Muller-Schenburg (ed) *Measurement of blood flow and local tissue energy production by thermal methods*. Thieme-Stratton, New York, pp. 50–56 (1983).
12. Valvano JW, Cochran JR, and Diller K. The simultaneous measurement of thermal conductivity, thermal diffusivity and perfusion in small volume of tissue. *J. Biomech. Eng.*, 106:192–197 (1984).
13. Pennes HH. Analysis of tissue and arterial blood temperature in the resting human forearm. *J. Appl. Physiol*., 1:93–102 (1948).
14. Valvano JW, Badeau AF, and Pearce JA. Simultaneous measurement of intrinsic and effective thermal conductivity. In: Chato CH, Diller KR, Diller TE, and Roemer RB (eds) *Heat transfer in bioengineering and medicine*, ASME Winter Annual Meeting, Boston, HTD-Vol. 95, BED, Vol. 7, pp. 31–36 (1987).
15. Patel PA et al. A finite element analysis of a surface thermal probe. In: KR Diller (ed) *Thermodynamics, heat, and mass transfer in biotechnology*, ASME Winter Annual Meeting, Boston, HTD-Vol. 90, BED, Vol. 5, pp. 95–102 (1987).
16. Rastorguev YL and Ganiev YA. Thermal conductivity of aqueous solutions or organic materials. *Russian J. Phys. Chem*., 40:869–871 (1966).
17. Valvano JW and Chitsabesan B. Thermal conductivity and diffusivity of arterial wall and atherosclerotic plaque. *Lasers Life Sci*., 1:219–229 (1987).
18. Valvano JW. Low temperature tissue thermal properties. *Low Temperature Biotechnol.*, American Society of Mechanical Engineers Heat Transfer Division, 98:331–346 (1988).
19. Spells KE. The thermal conductivities of some biological fluids. *Phys. Med. Biol*., 5:139–153 (1960).

20. Cooper TE and Trezck GJ. Correlation of thermal properties of some human tissues with water content. *Aerospace. Med*., 42:24–27 (1971).
21. Patel PA, Valvano JW, Pearce JA, Prahl SA, and Denham CR. A self-heated thermistor technique to measure effective thermal properties from the tissue surface. *J. Biomech. Eng.*, 109:330–335 (1987).
22. Anderson GT, Valvano JW, and Santos RR. Self-heated thermistor measurements of perfusion. *IEEE Trans. Biomed. Eng*., 39(9):877–885 (1992).
23. Yuan DY, Xu LX, Zhu L, Holmes KR, and Valvano JW. Perfusion and temperature measurements in hyperthermic canine prostates. *17th Southern Biomedical Engineering Conference*, p. 85, Feb. 7–8 (1998).
24. Xu LX, Zhu L, and Holmes KR. Thermoregulation in canine prostate during transurethral microwave hyperthermia, part II: Blood flow response. *Int. J. Hyperthermia.,* 14(1):65–73 (1998).
25. Touloukian YS et al. *Thermophysical properties of matter: Thermal diffusivity*, Vo1. 10. IFI/Plenum, New York, pp. 390, 589 (1973).

Chapter 13
Thermal Damage and Rate Processes in Biologic Tissues

Sharon Thomsen and John A. Pearce

13.1 Introduction

Heat is generated in laser irradiated tissues by absorption and transformation of the light energy into heat. Once generated within tissues, heat is heat no matter what original energy source is used to produce it. Heating of cells and biological tissues can produce reversible injury and dysfunction that can be repaired by innate cellular and host mechanisms. However, more severe, irreversible damage leads to death immediately (primary thermal effects) or after (delayed secondary effects) the heating event. Sometimes, the dividing line between reversible (repairable) and irreversible (lethal) damage in living, surviving tissues is not easily observed at the time of heating. Therefore, to determine accurately the extent of effective (killing) thermal treatment, the observer has to wait for all the moribund, injured cells to die and undergo post-mortem necrosis (two-four days) [1–6]. In addition, low temperature heat can directly trigger cell death by apoptosis, a sequential series of complicated processes that requires energy, signal transduction, enzyme activation and time (hours to days) to develop [6–9].

The effects of cellular and tissue heating can be studied using both qualitative and quantitative gross and microscopic pathologic methods. Some of these effects are characteristic of direct thermal damage, but others are delayed non-specific responses and thus are attributable to thermal injury only when a history of thermal damage is provided.

In this chapter, the qualitative and quantitative pathologic markers of thermal damage will be identified, described and illustrated. In addition, experimental and computational methods will be described for estimating kinetic coefficients for those particular markers that develop as a result of physical and chemical reactions that can be described by first-order kinetics. The experimental results can be obtained under constant temperature or transient thermal history situations simulating continuous and pulsed laser irradiations. To be useful for the prediction

S. Thomsen (✉)
500 Discovery View Drive, Sequim, WA, USA
e-mail: slthomsenmd@yahoo.com

A.J. Welch, M.J.C. van Gemert (eds.), *Optical-Thermal Response of Laser-Irradiated Tissue*, 2nd ed., DOI 10.1007/978-90-481-8831-4_13,

of these temperature/time treatment dosimetries, the kinetic model must be coupled to quantitative pathological analyses of well-defined pathologic end points showing thermal effects. For many heat treatment clinical goals that involve cell killing, the more useful endpoints are the spatial boundaries of (1) cell and tissue death (effective treatment volume) and (2) full extent of the lethal and non-lethal thermal effects (total treatment volume). In the case of in vivo experiments with transient thermal histories, the experimental data yield a noisy kinetic plot; however, good estimates of the appropriate rate coefficients can be made using a relatively new method described in Section 13.3.

13.2 Pathophysiology and Pathogenesis of Photothermal Lesions

13.2.1 General Principles

13.2.1.1 Primary Thermal Injury

Pathologic mechanisms of primary thermal injury are those that produce acute structural and functional abnormalities in cells and tissues due to the *direct physical interaction* of the heat with cellular and tissue proteins, lipoproteins and water. These pathologic effects can be detected immediately, during, or just after the heating episode in either living tissues (in vivo), living cells in culture (in vitro), or non-living tissues or organs recently removed from animals (ex vivo). The most characteristic and reliable diagnostic histopathophysiologic markers (indicators) of thermal injury are those produced by primary mechanisms of pathologic effects (see Table 13.1).

13.2.1.2 Secondary (Delayed) Thermal Injury

Pathologic mechanisms of secondary (delayed) thermal injuries are based on *delayed pathophysiologic cellular, tissue or host responses* triggered by cellular and tissue injuries and/or death. These secondary pathologic responses can occur soon (several seconds, minutes or hours) or long (days, weeks or months) after the heating event. Therefore, the full story of the development, progression and resolution (healing) of photothermal lesions can only be studied in living and surviving cells, tissues and organisms (see Table 13.2). These delayed responses are not unique to thermal lesions but are characteristic of the cellular and host reactions to any chemical, mechanical or physical insult that results in reversible and irreversible (lethal) cellular and tissue injury [6].

13.2.1.3 Identification of Thermal Lesions

The determination that a lesion is due to thermal injury requires (1) the observation of characteristic (diagnostic) pathologic markers of thermal damage (mainly

primary effects) in the tissues, (2) having the history of a heating episode and/or (3) the documentation of the heating parameters (tissue temperature and time at temperature) to show that significant heat injury could have been produced in that tissue or organ. Satisfaction of at least two or all three of these criteria is necessary to establish the role of heat in the creation of the lesion.

Table 13.1 Primary thermal effects in non-viable and viable tissue targets: low to high temperature sequence

In Vitro/Ex Vivo/In Vivo
I. *Thermal Activation of Increased Mitochondrial Membrane Permeability of Ions*
II. *Thermal Dissociation (Melting) of Phospholipid Cellular Membranes*
• Rupture of cellular and intracellular membranes • Loss of transmembrane electrical potentials • Leakage of cellular contents including water, ions and micro and macromolecules • Increasing disruption of functional arrays of membrane-associated enzymes & proteins
III. *Thermal Intracellular Protein Denaturation*
• Vital (Respiratory) mitochondrial enzymes • Membrane-associated receptor, transport and signal processing proteins • Metabolic and synthetic enzymes • Protein secretory products • Other intracellular and extracellular non-structural proteins, lipoproteins and glycoproteins • Intracellular nuclear and cytoplasmic structural proteins
IV. *Thermal Extracellular Stromal Protein Denaturation*
• Loss of tertiary and secondary molecular structure and associated physiologic functions • Severe intracellular and extracellular protein denaturation and precipitation in situ (heat fixation)
V. *Water Vaporization*
• Tissue desiccation • Accumulation of water vapor in expanding tissue vacuoles
VI. *Tissue Caramelization and Carbonization*
VII. *Tissue Ablation*
• Rupture of thin-walled vacuoles with explosive fragmentation of desiccated tissue • Tissue combustion • Tissue vaporization • Tissue plasma formation

Modified from [4]

Table 13.2 Secondary (delayed) thermal effects: pathologic responses to cell and tissue injury and death

In vivo and survival only: local host responses	In vivo and survival only: cellular and tissue responses	In vivo and survival only: host responses
I. *Red Thermal Vascular Damage Zone (Transition Zone*)* • Hemostasis • Thrombosis • Hemorrhage • Hyperhemia	II. *Cellular Repair (Non-lethal Injury)* • Reconstitution of cellular membranes • Heat shock protein production • Resumption of cellular synthetic and physiologic functions III. *Traumatic Cell and Tissue Death* • Lytic necrosis due to release of lytic enzymes from thermally damaged intracellular organelles (lysosomes) • Coagulative cell necrosis due to loss of energy production in thermally damaged mitochondria • Ischemic (coagulative) necrosis due to regional blood flow blockage (infarction) IV. *Cell Death Due to Apoptosis*	V. *Inflammatory Response to Tissue Death and Necrosis* • Release of local cytokines, cytotactic factors and vascoactive factors • Activation of cell attachment receptors • Activation of immune responses VI. *Organization of Necrotic Debris* VII. *Tissue Regeneration and Repair* • Parenchymal cellular regeneration • Wound healing: vascular and fibrous granulation tissue formation Fibrous scar maturation Fibrous scar contraction

Modified from [4]
*Transition Zone. Described by J. Coad in [10]

13.2.1.4 Factors Associated with Thermal Lesion Size, Development and Progression

The extent and severity of the pathologic effects produced by the interactions of heat with biologic tissue are determined by (1) the geometry of the heat source volume and thermal gradients created in the tissue over time, (2) tissue optical and thermal properties before heating and their dynamic changes during and after heating, (3) tissue temperatures and time at temperature (thermal history) at any point along the thermal gradients, (4) target tissue composition and structure including protein, lipid and water content and their structural anatomy, (5) target tissue viability, (6) mechanisms of primary (direct) and secondary (delayed) thermal injury, and (7) the various cellular, tissue and host responses to the injury over time in surviving organisms. The signals for these responses involve the release of several soluble

tissue cytokines, cytotactic factors and growth factors from (1) injured, dying and/or necrotic cells, (2) invading inflammatory cells, and (3) surrounding viable parenchymal and connective tissue cells. Usually, these factors are various polypeptides that can trigger (1) local vascular and cellular inflammatory events, (2) immune responses, and (3) control regional cell proliferation and protein synthesis over time [6, 10–13].

13.2.1.5 Cells: Basic Units of Life and Death

In general, when considering lethal thermal injury of tissues, it is good to remember that individual cells are the basic unit of life. All tissues and organs are a mixture of *living* cells residing in a more heat resistant, *non-living* extracellular matrix. The extracellular matrix is a structurally and functionally complex tangle of formed and unformed proteins such as collagens, elastins, glycoproteins, globular proteins and polypeptides suspended in a solution of water, sugars and salts. The more formed matrix elements such as the collagens, elastin and basement membranes are arranged into a scaffolding of support platforms and compartments to provide the functional organization for parenchymal cells and tissues such as glands, skin, nerves, muscle and bone. The matrix components are produced and maintained by living cells within the tissue including parenchymal cells (epithelial, muscle and nerve cells) and connective tissue cells such as fibroblasts, lipocytes and myofibrocytes [14–18].

At low temperatures and heating times, individual tissue cells can die, but the extracellular matrix remains intact leaving a preformed scaffolding to guide regenerating parenchymal cells into reconstituted, functioning structures such as glandular ducts and liver lobules. However, when higher temperatures and longer heating times are applied, then the scaffolding collapses and the whole necrotic tissue is replaced by scar tissue. Also, if the rate of formation of scar tissue is greater than that of parenchymal cell regeneration, the volume of the necrotic tissue will be filled by scar rather than regenerating parenchyma. These rates vary greatly among different species and among different tissues within the same species [19].

13.2.2 Zones of Thermal Damage: Gross and Histologic Pathologic Features

13.2.2.1 Heat Source Volumes and Thermal Gradients

The geometry of the heat source volume in laser irradiation is determined by the volume distribution of the absorbed light and the transformation of the absorbed light energy to heat [20, 21]. Initially, heat energy is confined to the light absorbing tissue volume. But, over time, temperature gradients between the heated source volume and the cooler peripheral tissues will be produced by the radial and axial transport of the heat energy away from the original heat source volume. Heat produces many

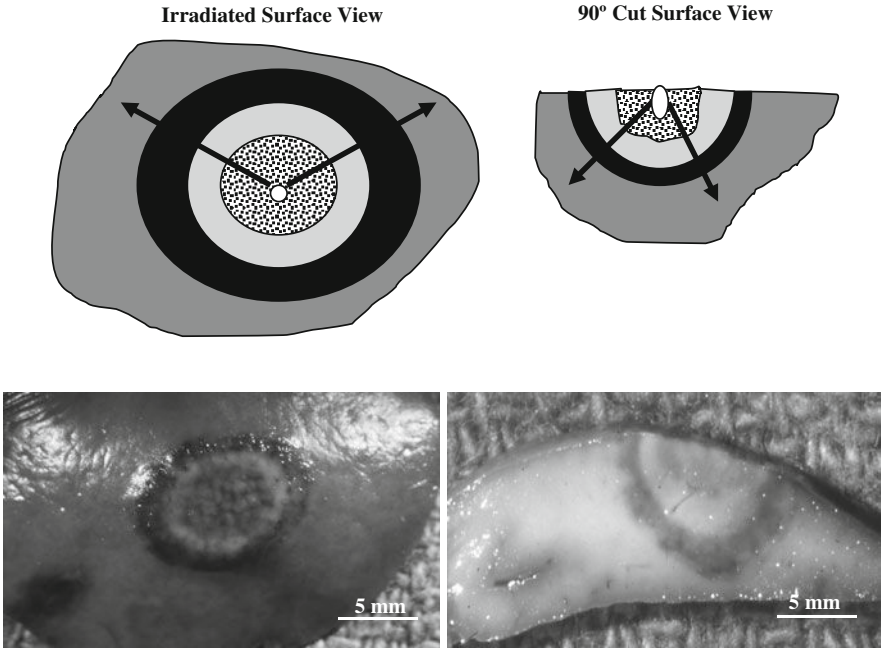


Fig. 13.1 Thermal lesions in living and surviving biologic tissue. *Top*: Schematic views of acute thermal lesions. Heat gradients (*arrows*) extend radially from a single volume heat source ◯ into the peripheral tissues. The different bands represent the concentric zones of thermal damage that form at the hottest point in the center to the cooler periphery. *Bottom*: Similarly oriented views of acute photothermal lesions in rat liver exposed in vivo to cw Nd:YAG laser irradiation (λ =1.064 μ; beam spot size, 2 mm; 9 W, 10 s). Harvested 30 min after irradiation and fixed in 10% buffered formalin. The boundaries of the different concentric thermal damage zones are distinct. The inner mottled zone and the more peripheral homogenous light zone represent different severities of thermal protein denaturation (thermal coagulation), i.e. primary thermal injury. The most peripheral darker zone represents vascular-based responses to heat, i.e. secondary thermal injury

pathological lesions within biologic tissues depending on the tissue temperature $T(r, t)$ and time at temperature $t(r, T)$ at any point along the gradients. These pathologic lesions will form concentric zones of thermal damage around the single volume heat source with the more severe effects found closer to the hot center and less severe in the cooler periphery (Fig. 13.1). Since many of these zones are separated by distinct boundaries, the pathologic effects of heating can not only be recognized and described (qualitative pathology), but also their boundaries can be measured and mapped (quantitative pathology). These data can be used to generate mathematical models of thermal effects in tissues as discussed in Section 13.3 [3, 10, 19, 22, 23].

13.2.2.2 Ex Vivo and In Vivo Heating: No Survival

Beginning at the hot center and progressing to the cooler periphery, the thermal damage zones include (1) tissue defects due to tissue ablation, (2) carbonization, (3) water vaporization, (4) thermal tissue and cellular protein coagulation or denaturation, and (5) cellular and intracellular organelle lipoprotein membrane rupture and collapse. All of these primary pathological effects reflect heating conditions sufficient to cause lethal thermal injury, that is, injury leading to cell death in living tissues and organs. All these pathologic zones can be created in recently extracted, fresh (ex vivo) and living (in vivo) tissues and organs.

13.2.2.3 In Vivo Heating with Short Term Survival

However, other pathologic lesions resulting from heating require that the host animals are alive, have an intact blood supply and blood pressure and survive at least for several seconds to a few hours after the heating event (Fig. 13.2). Thermal protein denaturation, cellular membrane damage and endothelial cell death by apoptosis produces blood vessel damage that results in a red thermal zone of vascular damage at the interface of the thermal lesion and the surrounding normal tissue.

Depending on the irradiation/heating conditions, the outer boundary of this red damage zone usually marks the outer boundary of cell/tissue death [2, 3, 19]. However, the repair mechanisms of some viable, yet heat-injured cells located at the cooler, outer edges of the red zone may remain intact. These individual cells can repair themselves and survive. Therefore, for heat treatments of invasive cancers, standard practice involves creating an additional therapeutic margin of 5–10 mm around the tumor to make sure that the thermal treatment volume encompasses as many invading cancer cells as possible.

13.2.2.4 In Vivo Heating with Prolonged Survival

Upon prolonged survival (hours to months), the host response to lethal thermal injury will involve a series of pathophysiological events that are triggered by any injurious insult to living cells and tissues, not just heat. Initially, these non-specific, general responses include post-mortem cell and tissue necrosis (Fig. 13.3). Then, resolution and repair of the tissue structural and functional void left by the necrotic process involves (1) organization, that is, clean up and removal of necrotic debris soon followed by (2) tissue repair that includes tissue regeneration and wound healing (scar formation) (see Figs. 13.6, 13.7, 13.8, and 13.9).

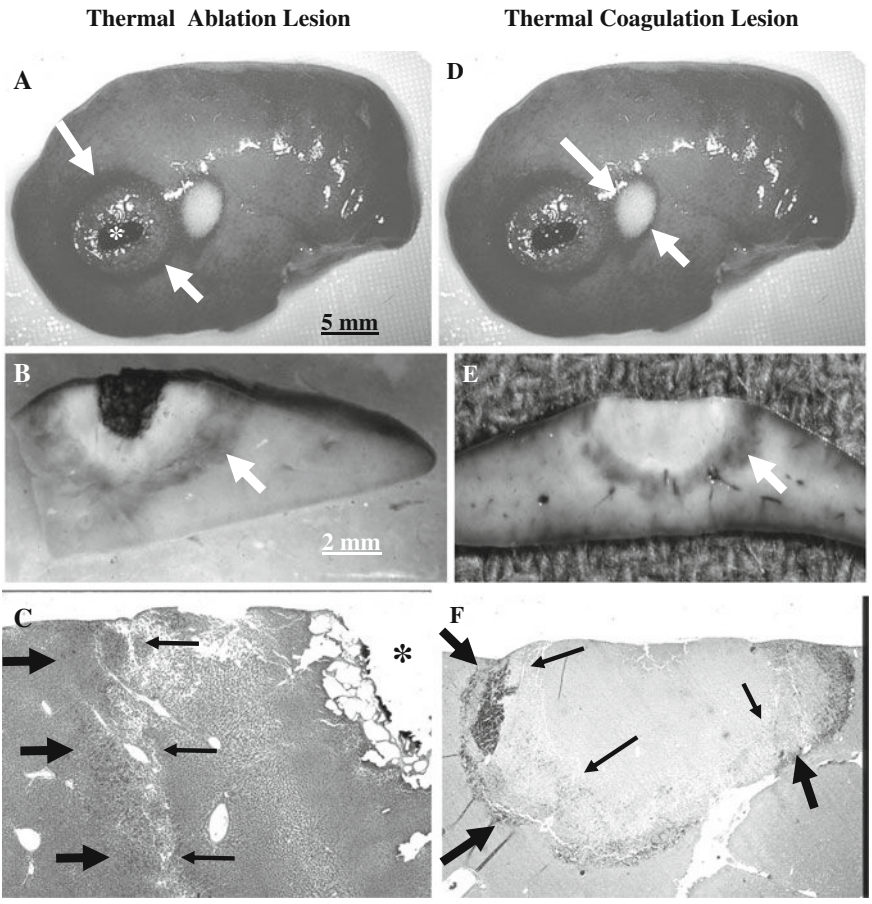


Fig. 13.2 Photothermal lesions in rat liver surviving 30 min after irradiation (cwNdYAG laser; $\lambda = 1.064$ μ: beam spot size, 2 mm; 9 W, 10 s). (**a**) Ablation lesion: irradiated surface, fresh specimen, gross appearance: black carbon deposits line the ablation hole (*asterisk*) that is surrounded by mottled, thermally denatured (coagulated) tissue. A distinct darker rim of thermal vascular damage (*arrows*) defines the periphery of the heat lesion. (**b**) Ablation lesion. cut surface, 10% formalin fixed gross specimen: The ablation hole is lined by a thin layer of dark carbonized tissue and surrounded by a prominent zone of lightly colored, thermally coagulated liver tissue and, more peripherally, a darker band of tissue (*arrow*) distorted by vascular thermal damage. (**c**) Ablation lesion. Microscopic section: a distinct zone of vacuoles formed by entrapment of water vapor in desiccated tissue is located just below the thin black carbon deposits that line the ablation hole (*asterisk*). The next zone is formed of thermally denatured (coagulated) tissues and, more peripherally, the outer zone is the band of red vascular thermal damage (*arrows*). The tissue beyond the red zone is normal, viable liver tissue. (Hematoxylin and eosin, H and E, stains. Original Magnification 25×). (**d**) Thermal coagulation lesion. Gross appearance: thermal denaturation (thermal coagulation) of the numerous tissue proteins and other macromolecules produce a white, opaque, firm central lesion surrounded by a distinct dark band of thermal vascular damage (*arrows*). (**e**) Thermal coagulation lesion. Cut surface, 10% buffered formalin fixed gross specimen: The zone boundaries demarcating the central zone of light tan, thermally coagulated liver, the more peripheral dark band of thermal vascular damage (*arrow*) and the outer non-damaged

Fig. 13.2 (continued) viable liver are distinct. (**f**) Thermal coagulation lesion. Microscopic appearance: The entire central zone of thermal denaturation (thermal coagulation) appears homogeneous at this low magnification. The more peripheral thermal vascular damage zone (*arrows*) appears more spotty and irregular in this 4 μ thick tissue section than in the gross lesion. The thermal vascular damage zone is dark because of primary heat fixation of red blood cells within their blood vessels, hemorrhage (blood escaping through the walls of heat damaged blood vessels, thrombosis (intravascular blood coagulation) and hyperhemia (increased blood flow through dilated blood vessels). (H and E stains. Orig. Mag. 16×)

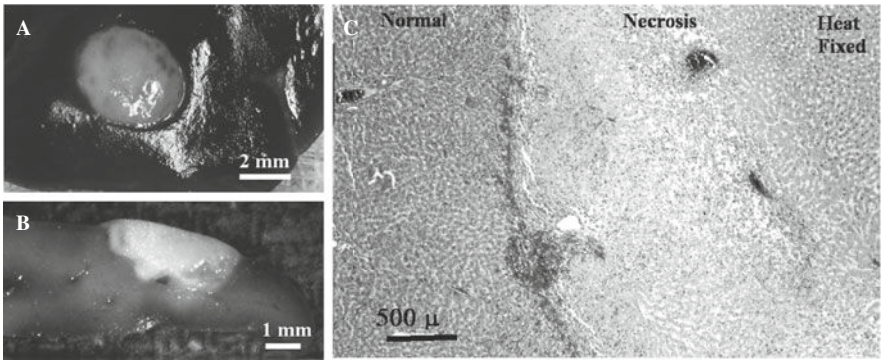


Fig. 13.3 Necrosis of photothermal coagulation lesion in rat liver surviving 3 days after irradiation (cw NdYAG laser; $\lambda = 1.064$ μ; beam spot size, 2 mm; 9 W, 10 s). (**a**) Necrotic lesion, irradiated surface, gross appearance of fresh specimen: the boundary of the opaque, lightly colored necrotic photothermal lesion with the surrounding dark viable liver tissue is very distinct. (**b**) Necrotic lesion, cut surface 10% buffered formalin-fixed gross specimen: the peripheral vascular thermal damage zone is no longer apparent since all damaged cells and tissue in that zone have died and undergone necrosis. (**c**) Necrotic lesion, microscopic appearance: the higher magnification offered by microscopy allows the observation of different types of tissue damage within the necrotic photothermal lesion. The viable, normal liver (*left*) is separated from the dead, necrotic liver (*right*) by a thin band of hemorrhage not visible in the gross specimen. The zone of necrosis includes regions of post-mortem lytic necrosis and post-morterm coagulative necrosis. [H and E stains.]

13.2.3 Mechanisms of Thermal Injury

13.2.3.1 Primary (Direct) Thermal Injury

The mechanisms of primary direct thermal injury of cells and tissues involve the numerous processes of tissue ablation, carbonization, water vaporization, thermal denaturation of proteins and cell membrane rupture (see Table 13.3). The first three mechanisms are very complicated physiochemical reactions that are not explained by simple first-order kinetics. On the other hand, heat-induced intracellular and extracellular protein molecular denaturation and cellular lipoprotein membrane dissociation "melting" and disruption are primary processes that have been shown to be first-order kinetic processes that can be described by the Arrhenius equation as will be discussed in Section 13.3.

Table 13.3 Treatment effects produced along heat gradients extending from hot center of single volume heat source to cooler periphery in vivo and in vitro heating

Treatment effect	Mechanism of treatment effect	Pathologic marker of treatment effect
Tissue ablation	Removal of tissue mass by explosive fragmentation, tissue vaporization and combustion	Defect or hole in tissue
Tissue carbonization	Thermal reduction of organic tissue components to carbon	Thin line of carbon lining ablation defect
Tissue desiccation and vacuolization	Tissue water vaporization and expansion of steam vacuoles entrapped in tissue	Damage zone of vacuolated, shrunken, brittle tissue beneath carbon layer
Thermal tissue coagulation	Thermal denaturation of extracellular matrix proteins Thermal denaturation of cytoskeletal and cell proteins Thermal denaturation of other cellular and tissue proteins	Heat fixation of cells and tissues Loss of native birefringence of collagen, elastin and muscle Hyalinization of collagen Spindling of epithelial cells
Rupture of cellular membranes	Thermal denaturation of vital respiratory enzymes Molecular dissociation of cell Membranes and collapse of cell organelles	Hyperchromasia and spindling of cell nuclei Hyperchromasia of cell cytoplasm

13.2.3.2 Tissue Mass Ablation

For the purposes of this chapter, *tissue* ablation is defined as removal of tissue mass to create a space within the tissue. *Functional* ablation is the removal or modification of a tissue function due to actual removal of mass or lethal injury of tissue. "Ablation" as used in the cardiac clinical literature usually means *functional* ablation. For example, creation of a thermal coagulation lesion in heart muscle is done to "remove" a functional source of cardiac arrhythmias. The *treatment endpoint* is the "ablation" or disappearance of the arrhythmia as monitored by an electrocardiogram and the *pathologic endpoint* is a volume of thermally coagulated (dead) myocardium.

Lists of the mechanisms of laser-mediated tissue ablation can be arranged by temperature gradients beginning at high temperature mechanisms requiring huge amounts of heat energy delivered in very short pulses then proceeding to those associated with lower temperatures and longer irradiation times. These are (1) plasma formation, the ionization of molecules at the irradiated surface; (2) tissue vaporization, the phase changes of solid, organic tissue components to liquids and gases; (3) combustion, rapid oxidation of organic tissue components with production of flames and smoke comprised of various gases and very small carbonized particles

and inorganic ashes; and (4) explosive tissue fragmentation due to rapid expansion and rupture of water vapor (steam) vacuoles in the heated tissues ("popcorn" effect).

Pathologic evaluation of *tissue ablation* involves observation and measurement of what is *not* there, that is, a hole or defect in the tissue at the irradiation site. Also, the defect has to be associated causally with a reasonable thermal history and/or surrounded by other tissue abnormalities that are characteristic of thermal damage and tissue ablation. Otherwise, the naïve (and not-so-naïve) observer could misidentify any old tissue defect as being the "laser photothermal cut" (S. Thomsen and J. Coad, personal experiences, 1979–2008). With proper orientation, measurements of the maximum dimensions the tissue defect and the surrounding zones of thermal tissue damage in the tissue can be done using gross and microscopic observations [19].

13.2.3.3 Carbon Formation and Carmelization of Tissue

Carbon formation is the heat-mediated reduction of tissue and cellular organic molecules to elemental carbon that forms a very thin black membrane (5–20 microns) covering the walls of the defect or the irradiated surface of the tissue (see Figs. 13.2 and 13.4). The clumped black carbon particles are easily seen and measured in properly stained microscopic tissue sections. Carmelization is the melting, partial reduction and resolidification of tissue sugars. It occurs at slightly lower temperatures than carbonization but, in light microscopic histological sections, is frequently overshadowed by carbonization and staining artifacts induced by tissue desiccation. Carbonization is typically associated with local temperatures in excess of 200°C.

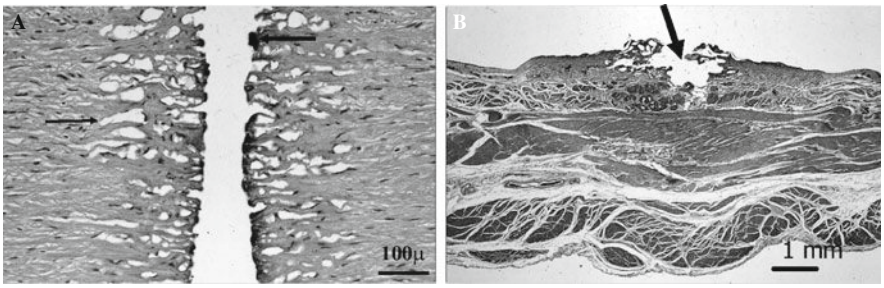


Fig. 13.4 Water vaporization mechanisms: tissue dessication, water vapor vacuole formation and explosive fragmentation. (**a**) Thermal damage zones: carbon formation, caramelization, water vacuole formation and tissue dessication in argon laser irradiated dog aorta. The ablation hole was formed by high energy mechanisms indicated by black carbon granule formation (*thick arrow*) and caramelization of surface tissues and the even sides of the ablation hole. A uniform band of water vapor vacuoles (*thin arrow*) has formed in the hyperchromatic, hyalinized, dessicated, thermally denatured cells and connective tissue located just below the ablation surface. The boundaries of this water vapor vacuolization zone are measurable and can be predicted in rate process mathematical models. (H and E stains. Orig. Mag. 400×) (**b**) Thermal tissue ablation by explosive fragmentation in dog urinary bladder mucosa: argon laser irradiation in vivo. The very irregular ablation defect (*arrow*) was produced by rapid expansion and rupture of large and small, thin-walled water vapor vacuoles. The thin vacuolar walls consisted of desiccated, brittle, thermally denatured tissues. The boundaries of the more peripheral thermal damage zones are not easily recognized at this low magnification. (H and E stains. Orig. Mag. 6.25×.)

13.2.3.4 Water Vaporization

Water vaporization, vapor diffusion and tissue desiccation appear as the tissue temperatures approaches 100°C at standard pressure (Fig. 13.4). When generation of water vapor is faster than its rate of diffusion out of the tissue, the vapor can be trapped within the tissue, forming steam vacuoles. If the incident laser irradiation energy continues to be converted to heat energy, pressure rapidly increases, the steam expands within the vacuoles stretching their now desiccated, brittle, thin walls until they rupture either creating bigger vacuoles below the surface or explosions throwing tissue fragments of the vacuole wall from the surface ("popcorn effect") [20, 24–26].

The temperature history of the "popcorn effect" and explosive vacuole rupture at the tissue surface can be followed using a calibrated thermal camera and a microphone (or ears of the investigators). As irradiation is delivered into the tissues, the surface temperature rises (sometimes up to 130–160°C), the water vapor expands within the vacuoles until they rupture with a loud "pop." The "pop" marks the rapid escape of the steam into the atmosphere and a marked temperature drop to around 100°C, the temperature of water vapor in the remaining tissues. Depending on the tissue temperatures, tissue water content and heating times, a band-like zone of variably sized, intact and ruptured vacuoles can be seen just below the carbon layer in histologic sections [27]. This zone tends to have distinct and measurable borders with the adjacent thermally coagulated (denatured) tissues when viewed with a light microscope and represents a major zone of water-mediated thermal effects in tissues (Fig. 13.4). Although water vaporization is a complicated thermodynamic process, the pathological effects of tissue vacuole formation, expansion, vacuole size and tissue desiccation can be modeled and incorporated into mathematical models of thermal damage in biological tissues [28].

13.2.3.5 Thermal Protein Denaturation (Thermal Protein Coagulation)

General Comments About Tissue Proteins

Proteins are ubiquitous tissue components that include a myriad of pure proteins, lipoproteins and glycoproteins. These molecules form solids, gels, membranes and dissolved materials within (intracellular) and outside (extracellular) the cells that comprise any tissue [15–18, 29, 30]. Cellular enzymes are proteins that act as catalysts needed for proper, coordinated cell functions required for life. They are critical components of systems responsible for (1) energy production, (2) protein, carbohydrate, lipid and DNA/RNA synthesis and degradation, (3) cell repair, (4) transport of nutrients, water, salts, metabolic wastes and secretory products across cell membranes and (5) signal transduction across membranes inside and outside cells.

Proteins also form the important structural components within the cell which are the intracellular cytoplasmic and nuclear skeletal proteins and the cell motility proteins [17, 31, 32]. The intracellular structural and motility proteins are necessary for (1) organelle and cell support, (2) intracellular and transmembrane material transport, (3) cellular movement and (4) reproduction and regeneration with formation of the mitotic spindle for chromosome sorting and movement and the mechanisms for cell division.

The extracellular structural proteins form intricate scaffolding networks and environments necessary for integration of cells into tissues and for coordinated structural and functional organization of tissues to form organs and the whole organism. The most prevalent extracellular structural proteins are (1) the formed elements, collagens, elastins, fibronectin and neutral, basic and acidic aminoglycans and (2) the various amorphous proteins, aminoglycans and polypetides including mucins, globular carrier proteins, and other cell secretions and products. These elements are mixed with water, salts, sugars and polypeptides to form the semi solid gel in which cells reside, function and travel [15, 16, 18, 33–36].

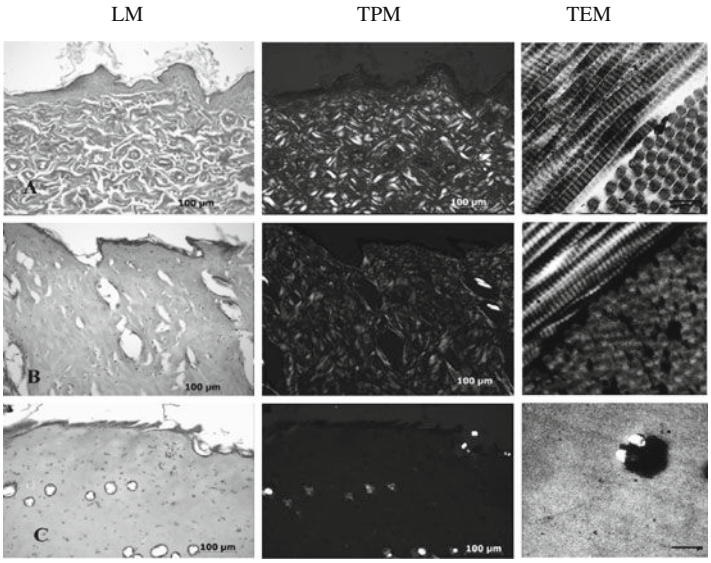


Fig. 13.5 Loss of Type I collagen birefringence image intensity: rat skin heated in vitro in water bath for 20 min at various temperatures. [*Left photos*: H and E stained. Light Microscopic images (LM); center photos: H and E stained. Transmission Polarized Microscopic images (TPM) Orig. Mag. 100×. *Right photos*: Lead citrate and uranyl acetate stains Transmission Electron Microscopic images (TEM). Bar = 200 nm)] (**a**) (*top row*) Rat skin at room temperature: Large concentration of native state relative to concentration of non-native state. The Type I collagen of the skin dermis retains its distinct fibrous pattern and bright birefringence image as seen in the LM and TPM images. The characteristic, semi-crystalline striated arrangement of the collagen molecules forming the collagen fibril demonstrated in this TEM image is responsible for Type I collagen birefringence, the optical property of rotating transmitted polarized light in a dark field. (**b**) (*middle row*) Rat skin heated to 60°C: Decreased concentration of native state relative to concentration of non-native state. The dermal collagen is hyalinized ("glassy") when viewed with LM. However, the residual fibrous pattern is revealed in the TPM of the heated collagen which has undergone a partial loss of birefringence image intensity. TEM shows that collagen fibrils retain their regular longitudinal molecular arrays (striations) responsible for the residual birefringence image. However, the fibrils have swollen, their distinct borders are obscured and are surrounded by darkly stained electron-dense partially denatured proteins and glycoproteins thus producing the homogeneous image of hyalinization in the LM. (**c**) (*bottom row*) Rat skin heated to 90°C: Low – essentially zero – concentration of native state relative to concentration of non-native state. The dermal collagen is now totally homogeneous and the birefringence image is totally lost. TEM shows only finely distributed granules of thermally denatured collagen and considerally larger grains of contaminating dirt

13.2.3.6 Thermal Denaturation of Proteins

The direct, primary thermal effect on the intracellular and extracellular proteins, glycoproteins and lipoproteins of cells and tissues is thermal denaturation of these macromolecules [37–40]. Thermal denaturation is the accumulated structural and functional change in heated protein that result from the physical and chemical molecular disequilibrium and phase changes introduced initially by heating and subsequent cooling. Usually, thermally denaturized macromolecules lose their characteristic quaternary, tertiary and secondary molecular configurations due to disruptions of hydrogen and ionic bonds as well as sulfide and other covalent bonds formed among primary amino acids chains, the building blocks of proteins. In the denaturation process, the molecules are reduced from highly organized, characteristic structures with specific functions to amorphous granules of collapsed and ruptured primary chains with no biological function (see Fig. 13.5). Accumulation of these highly scattering granules is responsible for the generation of the usually opaque, firm white or light tan volumes observed in heated target tissues. The transformation of fried egg whites from clear, viscous fluid rich with the protein, albumin, to a white, opaque solid is the classic example of thermal denaturization of protein [40]. The same denaturation processes are responsible for the loss of native birefringence of Type I collagen and the motor proteins, actin and myosin, in the sarcomeres of skeletal and smooth muscle cells, as discussed in Section 13.3 [41–48].

On the other hand, molecular protein denaturation and dehydration can be associated with transparent clearing of highly scattering, opaque white collagen-rich (Type I collagen) tissues such as skin dermis or muscle tendon [49]. Unfortunately, the temperatures/time at temperature thresholds of thermal clearing are much higher than those associated with cell killing; therefore heat cannot be used to "make skin transparent" since second and third degree burns would be produced.

13.2.3.7 Heating of Cellular Membranes

Cells are membrane-bound bags of proteins, carbohydrates, lipids, water and salts in which are suspended smaller, membrane-bound organelles such as nuclei, endoplasmic reticulum, mitochondria, the Golgi apparatus, secretory vacuoles and phagosomes [15–18, 32, 50]. The organelles compartmentalize, organize and separate different cellular metabolic functions, including energy generation (mitochondria), synthesis of DNA and mRNA (nuclei) and proteins (endoplasmic reticulum), packaging (Golgi apparatus) and storage of secretory products (secretory vacuoles) and degradation products (secondary lysosomes).

Cell membranes are composed of phospholipids oriented into a polar bilayer so that the hypdrophobic portions (fatty acid chains) of the molecule form the inner layers of the membrane and the hydrophilic portions (phosphate or organic esters) form the outer layers to face the watery environments of the extracellular and intracellular spaces. Embedded within and on both surfaces of the cellular membranes are important protein and glycoprotein structures including specific receptor proteins,

water pores, ion channels and enzymes. These proteins are important for transport of nutrients, metabolites, salts and water and signal transduction (stimulating and inhibiting) across the membranes.

When heated above physiological temperatures, the hydrogen bonds between the hydrophobic lipids become weaker and the individual phosholipid molecules separate creating defects in the membranes. The water and solutes within and outside the membrane barriers can leak in and out thus disrupting the intracellular and extracellular functional equilibria of cells and the surrounding the environments. If the heating continues and/or the temperatures increase, the defects enlarge, overwhelming the usual repair mechanisms, and the cell dies.

13.2.4 Mechanisms of Heat-Induced Cell Death

13.2.4.1 "Heat Sensitive" and "Heat Resistant" Cells and Tissues

Thermal denaturation of intracellular and extracellular proteins and molecular dissociation (melting) of cellular membrane lipids result in changes of molecular structure and loss of cellular function that, if sufficiently severe and/or prolonged, will produce lethal thermal damage and cell death. The tissue temperatures and times at temperature at which these changes occur differ among proteins, lipoproteins, glycoproteins, cell membranes and tissues. Some tissue components are more sensitive to heat elevation than others and denature or dissociate at lower temperatures, thus are considered relatively "heat sensitive." Those proteins, glycoproteins and lipoproteins that denature at higher temperatures are considered "heat resistant." A very general order of the range of more heat sensitive to the more heat resistant tissue proteins would be the following: (1) respiratory, energy-generating enzymes of the mitochondria, followed by the (2) intra- and extra-cellular non-structural tissue proteins, glycoproteins and lipoproteins, and lastly, the (3) intra- and extracellular structural tissue proteins. Cell membranes dissociate at low temperatures; therefore, the lipoprotein membranes are more heat sensitive than other cellular proteins [50–52].

13.2.4.2 Minimal Thermal Injury that Leads to Cell Death

To produce the high-energy molecule, adenosine triphosphate (ATP), the respiratory enzymes have to be arranged in a specific order on the inner membrane of the mitochondrion [7, 16, 17]. Therefore, the critical minimal lethal thermal injury to a cell is the rupture and fragmentation of the mitochondrial membrane along with the thermal denaturation of the respiratory enzymes sufficient to produce cell death. However, if the resultant membrane defects are small, they can be repaired and, if few enzymes are denatured, they can be replaced on the membrane by synthesis of new ones. Thus, if sufficiently small, a thermal injury can be considered to be reversible and the cell will survive the injury.

Cell death produced by lethal thermal injury can be due to (1) abrupt physical disruption of cell structure and function and structure by "heat fixation" of proteins in situ, a primary thermal effect, (2) gradual loss of vital cellular functions including disruption and loss of energy production, gradual loss of repair enzyme function and gradual loss of membrane integrity and function, and (3) induction of apoptosis as discussed below. The damage involving gradual loss of vital cellular functions produces severely injured (moribund) cells that can continue to have regions of cytoplasm and nucleoplasm that function for a while, but not at full capacity. These injuries will eventually lead to cellular system failure and death. Therefore, some tests for viability such as cellular dye exclusion and vital stains that rely on the oxidation/reductions functions of certain mitochondrial respiratory enzymes can indicate false "viability" for a time in injured and dying cells. The enzymes may be working for the time being, but the cells and tissues are moribund and eventually will die. Depending on the tissue, organ and animal species, cell and tissue death will occur immediately or within hours, but may take a day or two to become fully manifest by the development of post-mortem necrosis.

The events of isolated cell death by pure apoptotic mechanisms do not lead to post-mortem necrosis [6]. However, in real life and death situations, both post-mortem necrosis and apoptosis can occur in the same tissues and, indeed, may share similar triggering mechanisms [7, 53–56].

The exact temperature/time thresholds for cell death have to be determined for each cell and tissue type. However, some cells are more heat sensitive than others. In general, cancer cells are killed more easily than their normal counterparts. The basic theory of clinical low temperature hyperthermic cancer treatments exploit this temperature/time differential to kill off the cancer cells while sparing the normal ones [57, 58].

13.2.4.3 Post-Mortem Necrosis

Necrosis, by definition, occurs after cell (and tissue) death and involves the degeneration and disintegration of cells and tissues after death [6]. In laymen's terms, necrotic tissues and organs are "rotten" tissues and organs. There are generally two histopathologic types of post-mortem necrosis: lytic necrosis and coagulative necrosis; and, frequently, they are intermixed in the same dead tissue. *Lytic necrosis*, the chemical fragmentation (lysis) of cells, cellular membranes and organelles and extracellular matrix, is due to the extensive and overwhelming effects of proteolytic and lipolytic enzymes [6, 59]. These are enzymes that have either leaked from the ruptured lysosomes of the dead or dying cells themselves or have been secreted by infiltrating inflammatory cells (polymorphonuclear cells). The second type of post-mortem necrosis is *coagulative necrosis*, the gradual degeneration and collapse of cells and cellular organelles [6]. This necrosis is usually associated with loss of energy production by direct mitochondrial dysfunction or by lack of oxygen delivery to the tissue by blocked blood flow (ischemia). Usually, coagulative necrosis occurs in cells that contain relatively few lysosomes and relatively large numbers of mitochondria such as striated muscle cells of the heart and muscloskeletal system. The two types of post-mortem necrosis occur within thermal lesions with the zone of

lytic necrosis being more peripheral to the zone of coagulative post-mortem necrosis since the more peripheral zones receive the infiltrating inflammatory cells with their lytic enzymes first [10, 23] (see Fig. 13.6).

Coagulative post-mortem necrosis is associated with the loss of energy production. The major source of metabolic energy generation occurs in the mitochondria where high energy phosphate bonds of adenosine triphosphate (ATP) are created by chemical reactions catalyzed by the respiratory enzymes. These protein enzymes are found on the inside surface of the folded inner membrane of the mitochondrion arranged in a regular sequential order that allows the efficient functioning of the chain oxidative/reduction reactions that produce ATP [6, 53–56]. Heat-induced mitochondrial membrane rupture will disrupt this highly efficient order and, along with the thermal denaturation of the respiratory enzymes, will lead to insufficient energy production and cell death to be followed by necrosis. However, loss of oxygen due to stoppage of blood flow in the red thermal damage zone can also

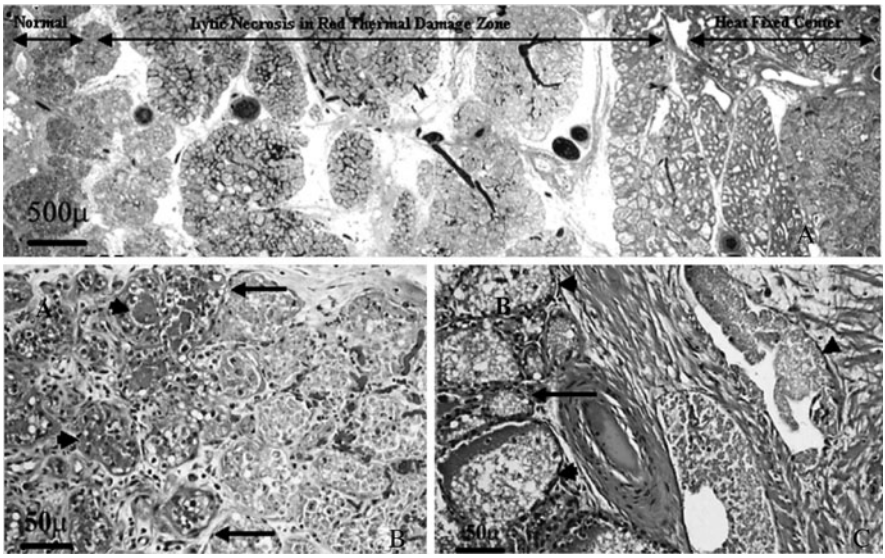


Fig. 13.6 Host responses to lethal thermal damage in post-lactational pig breast at three days survival. (**a**) Microscopic overview of thermal damage zones in pig breast. The boundary between the normal tissue and tissue undergoing lytic necrosis marks the full extent of the thermal treatment effect. This boundary is essentially coincident with the outer boundary of the vascular changes seen in the red thermal damage zone of the acute lesion. (H and E stains. Orig. Mag. 60×) (**b**) Boundary of normal post-lactational breast tissue and lytic necrosis. The boundary (*arrows*) between the viable and necrotic breast tissue is distinct. The viable tissue (left of boundary) consists of glands is composed of the degenerating vacuolated epithelial cells and residual milk proteins in the duct lumens (*arrow heads*) characteristic of the normal physiologic involutional changes in breast after weaning. The necrotic cells of the tissues to the right of the boundary are fragmented and many have no nuclei. H and E stains. Orig. Mag. 640×) (**c**) Heat fixed breast tissue in center of lesion. The hyperchromatic, shrunken nuclei and cytoplasm (*arrows*) are characteristic of thermally denatured (heat fixed) glandular cells. The lumens of the dilated glands are filled with foamy material (*arrow heads*) characteristic of milk production that was thermally fixed in place at the time of treatment three days before. (H and E stains. Orig. Mag. 640×)

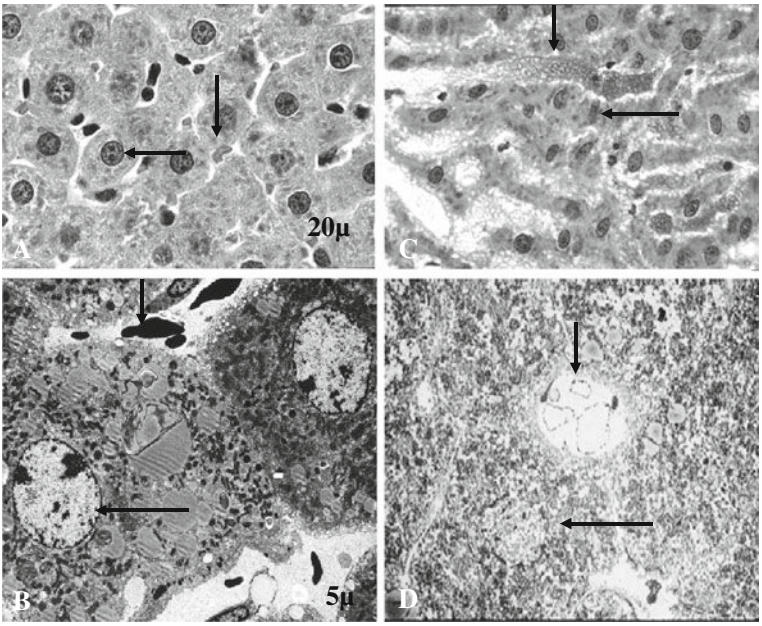


Fig. 13.7 Heat fixation in cw NdYAG (0.9 W, 6 s) laser irradiated rat liver: acute thermal coagulation lesion. (**a** and **b**) Normal liver: the liver cells (*horizontal arrow*) contain sharply demarcated nuclei and distinct cytoplasmic granular structures. The plasma membranes of the red blood cells (*vertical arrows*) are intact. ((**a**) LM H and E stains. Orig Mag. 640×; (**b**) TEM Orig. Mag. 5000×) (**c** and **d**) Heat fixation: in the LM image (**c**), the heat fixed liver cells and their nuclei are shrunken and dark (*horizontal arrow*) with fairly distinct boundaries. But in the TEM image (**d**), the nuclei are obscured (*horizontal arrow*). The thermally denatured nucleoproteins have entrapped the DNA that still binds with the hematoxylin stain in the LM but does not bind the heavy metal salt stains in the TEM. The TEM image shows the empty ghosts of the red blood cells outlined by their ruptured membranes (*vertical arrow*). ((**c**) LM H and E stains. Orig. Mag. 640×. (**d**) TEM Org. Mag. 5000×)

be an important factor in the generation of coagulative post-mortem necrosis due to ishchemia.

13.2.4.4 Cell Death Due to Rapid Thermal Coagulation (Heat Fixation)

Rapid elevations of tissue temperature in a short time can produce sudden death by immediate thermal coagulation and precipitation of cellular and extracellular proteins in situ ("heat fixation"). Concomitant extensive membrane rupture leads to immediate membrane depolarization and leakage of water, salts and small molecules out of the cells and their organelles, but their denatured protein contents stay in place. After this catastrophic death, none of the cells are alive, none of their intrinsic enzymes are functional and the heat fixed cells and extracellular matrix will not undergo post-mortem necrosis. Heat fixed tissues are found in the hottest portion

of the thermal lesion which usually is the center of the lesion. Because heat fixed tissues resemble "living tissues" in routinely-prepared and stained histopathologic sections viewed with a light microscope, numerous investigators (including some pathologists not aware of this phenomenon) have mistakenly reported the centers of thermal lesions as "alive" (see Figs. 13.6 and 13.7).

Eventually, small volumes of heat fixed dead tissue will be liquefied, organized and removed by inflammatory cells directly invading from viable tissue found at the periphery of the thermal lesion or from blood flowing in open lumens of necrotic blood vessels or in new blood vessels formed in wound granulation tissue. However, if the heat fixed volume is sufficiently large and communication cannot be established with the surrounding viable tissue, the heat fixed tissues will slowly fade away into granules, lose water and frequently will undergo calcification (dystrophic calcification) forming a hard mass surrounded by fibrous scar tissue [10].

13.2.4.5 Heat Fixation Vs. Pathophysiologic Post-Mortem Coagulation Necrosis: A Diagnostic Conundrum

In photothermal lesions, coagulative post mortem necrosis is to be distinguished from heat fixation which is the immediate (ms to s) result of direct thermal denaturation (thermal coagulation) and fixation of tissue proteins, glycoproteins and lipoproteins in situ. Heat fixed tissues tend to be in the hottest center of the thermal lesion while the post-mortem coagulative necrosis zone is more peripheral [10, 60]. Heat fixed tissues retain patterns of their cellular structures longer than tissues undergoing post mortem necrosis. Heat fixed cells retain distinct differential staining of the cytoplasmic proteins and nuclear DNA for days to weeks while in most cases of post mortem coagulative necrosis, the nuclear DNA disintegrates and disappears within a week. Ultimately, small volumes of heat fixed tissues will undergo lytic necrosis and organization by inflammatory cells invading from the peripheral viable tissues. However, large volumes of heat fixed tissues in some organs will never be penetrated by inflammatory cells or blood vessels and will never undergo lysis or organization. These volumes will be isolated by scar tissue and persist for several months to years [10].

13.2.4.6 Programmed Cell Death (Apoptosis)

The most critical eukaryotic cell organelles in terms of cell life and death are the mitochondria, intracytoplasmic organelles bound by two membranes that are considered to be the main arbiters of cell life and death [7–9, 53–55, 61]. Mitochondria are not only the sites of the energy production required to sustain life but also they contain the receptors and reactants required for apoptosis. Apoptosis is a physiologic mechanism of cell death in which a series of signal transductions at the mitochondrial or cell membranes initiate an activation cascade of specific mitochondrial and cytoplasmic proteins. These proteins or their precursors have to be in place ready to trigger the energy-requiring chemical reactions that produce characteristic cellular

morphologic and functional alterations over time (hours to 2–3 days) that terminate in fragmentation of the cells with the formation of apoptotic bodies, small bits of cell cytoplasm and nucleus. These apoptotic bodies disappear within a few hours after their formation by being ingested (by phagocytosis) into adjacent living cells where the bodies are digested within secondary lysosomes. Theoretically, no cellular contents are released from the cells dying by apoptosis; therefore there are no cytokines released to promote cellular or vascular inflammatory responses by the host. Some cells are genetically programmed to undergo apoptosis more readily than others. These include endothelial cells, lymphocytes, cells of embryologic tissues and some neoplastic cells. To make things even more confusing, the signals that initiate apoptosis can also activate mechanisms of cell death followed by necrosis. Most thermal lesions contain morphological and chemical markers of both death processes.

13.2.5 Wound Healing

Now that the different mechanisms of cell death and tissue necrosis have been discussed, the next steps the host will initiate will be the resolution of the pathologic effects of the injury, that is, the processes and events of wound healing. These include the activation of the mechanisms of wound healing including (1) organization (2) tissue regeneration, (3) repair and (4) scar tissue formation [6, 62–65].

13.2.5.1 Organization

Inflammatory cells enter the necrotic volumes upon the resumption of blood flow in damaged yet still open blood vessels or by migration from the surrounding viable tissues beginning about 8–12 h after cell and tissue death. Certain inflammatory cells, the macrophages, function as garbage collectors to clean up (organize) the necrotic debris in preparation for the later stages of wound healing: tissue regeneration, repair and scar formation. Depending on the species, organ and tissue, the organization will begin at about 2–7 days after injury. The debris-collecting macrophages phagocytose (“ingest”) the necrotic debris and carry it around in their cytoplasm in specialized intracellular membrane-bound organelles called secondary lysosomes (phagosomes). The “loaded” macrophages then (1) are incorporated into the wound healing tissues and remain in the local tissues, (2) migrate into the lymphatic vascular channels to land and accumulate in lymph nodes (regional “garbage dumps”), or (3) migrate back into blood vessels to be delivered, metabolized, excreted and/or stored in more distant tissues.

13.2.5.2 Tissue Regeneration and Tissue Repair

Some parenchymal cells and tissues such as liver cells, the epidermis of the skin and the glandular epithelium of gastrointestinal tract and breast among others are

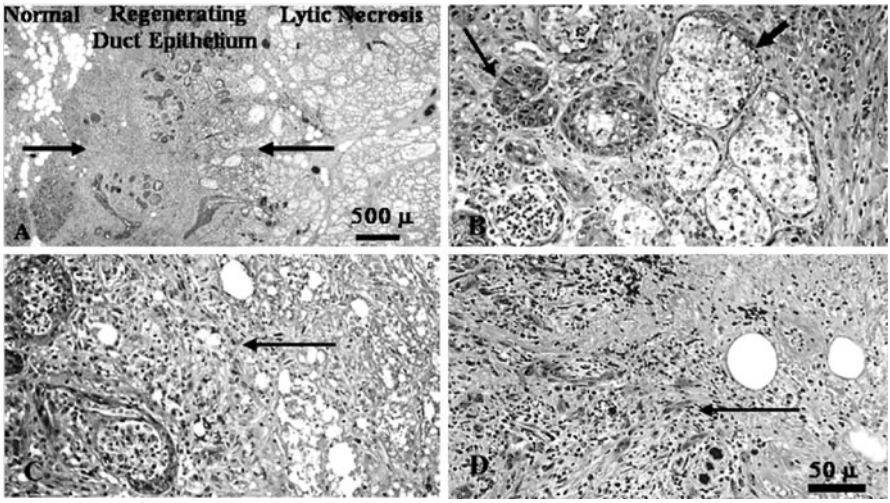


Fig. 13.8 Host responses to lethal thermal damage in pig breast at ten days survival. (**a**) Microscopic overview of healing of thermal lesion. A broad zone of scarring fibrosis and partial epithelial regeneration (*arrows*) separates the now physiologically quiescent, normal fibroglandular-adipose breast tissue from the more central portion of the lesion that is now undergoing lytic necrosis. Although some parenchymal glandular epithelial regeneration has occurred, the functional, lobular glandular pattern has not been reconstituted and the glandular elements are irregularly separated by fibrous scar tissue. (H and E stains. Orig. Mag. 25×) (**b**) Epithelial regeneration into remaining intact glandular connective tissue scaffolding. Regenerating glandular epithelial cells grow along the still-intact basement membrane lining of the breast ducts (*fat short arrow*) to form a lobular gland structure. The epithelium continues to proliferate to fill the duct lumen (*thin long arrow*). The epithelial cells are large with relatively abundant cytoplasm. (H and E stains. Orig. Mag. 640×) (**c**) Epithelial regeneration into region of collapsed necrotic connective tissue. Without the guidance of an intact connective tissue scaffolding, the regenerating epithelial cells grow in a more haphazard pattern. The boundary (*arrow*) between the viable (*left*) and necrotic (*right*) tissue is filled with infiltrating inflammatory cells and proliferating faintly spindled cells that could be activated macrophages, endothelial cells of new blood vessels (vascular granulation tissue) and/or myofibroblasts of new scar tissue (fibrous granulation tissue). (H and E stains. Orig. Mag. 640×) (**d**) Granulation tissue formation at edge of necrotic tissue. The new blood vessels characteristic of early wound healing form irregular, linear structures (*arrow*) in this example of early vascular granulation tissue. No epithelial regeneration is seen in this field. (H and E stains. Orig. Mag. 640×)

capable of regeneration after necrosis (see Fig. 13.8). As long as the usual tissue connective architecture is intact, the regenerating cells will be guided into the connective tissue scaffolding to form normally functional tissue structures such as liver lobules and digestive mucosal glands. The new functional cells are derived from immature, precursor cells in the surrounding viable tissues. These precursor cells are stimulated to proliferate, migrate and mature in response to numerous growth factors produced by the residual, living connective tissue and inflammatory cells at the periphery of the lesion. However, if the necrotic connective tissue scaffolding collapses, the proliferating parenchymal cells pile up into dysfunctional, disorganized

strands or nodules buried in scar tissue. Some adult parenchymal tissues, such as nerve cells of the brain and cardiac muscle cells, do not have readily available, dedicated precursor cells (or very few stem cells that are not easily stimulated to grow and mature) and, in general, are not capable of tissue regeneration. Usually, large volume parenchymal cell regeneration does not occur after large volume necrosis and the tissue repair is mainly by fibrous scar formation to fill in the void.

13.2.5.3 Granulation Tissue and Scar Formation

Whenever a defect is created by the organization of the necrotic tissues, the void is filled with either regenerating tissues or vascular and fibrous granulation tissue (early wound healing tissues) or a mixture of both (see Fig. 13.9). These granulation repair tissues are the earliest forms of scar tissue and will appear in the wound anywhere from 3–5 to 10 days, again, depending on the species and the tissue type within a species. Over the next days to months the scar tissue will be remodeled (regression of the vascular components and thickening of the collagen fibers) and the resultant maturing fibrous scar will contract and become smaller than the original thermal lesion.

13.2.6 Role of Heat Shock Proteins (HSP) in Thermal Lesions

HSP include a group of intracellular proteins that, among many other functions, facilitate the assembly of component peptides into the final configuration of functioning proteins [6, 17, 29, 61]. To initiate heat shock protein synthesis, the cell cultures or tissues are exposed to low temperatures (or other low level injury) for relatively short times. This heating can stimulate specific DNA expression and RNA production for heat shock protein synthesis while not killing the cells by direct injury or by triggering apoptosis. Once the DNA expression/RNA molecules are synthesized, they can remain active for some days so that, if the same cultures or tissues are reheated, the heat shock protein synthetic apparatus is in place and ready to produce new facilitators of enzymatic or structural protein repair. Thus, the cultures/tissues "preconditioned" by the initial heating are more "heat resistant" or protected from thermal injury from subsequent heating episodes because they can institute repair processes more readily than "unconditioned" cultures/tissues.

13.2.7 Mathematical Modeling of Pathophysiologic Thermal Effects

As discussed in great detail in the following Section 13.3, mathematical models estimating pathophysiologic markers (pathologic end points) of thermal damage in cells and tissues that follow first order kinetics can be developed. These can be used

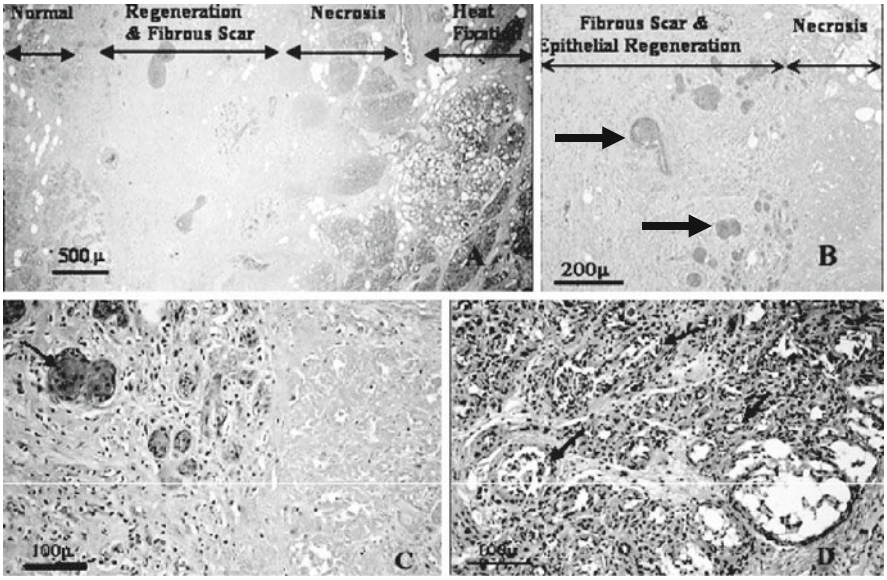


Fig. 13.9 Host responses to lethal thermal damage in pig breast at 28 days survival. (H and E stains) (**a**) Overview of healing thermal lesion. A relatively thick band of fibrous scar tissue that contains a few irregular, variably sized glandular duct structures separates the lobular glandular normal breast tissue from the more central necrotic and heat fixed breast. (Orig. Mag. 6.25×) (**b**) Prominence of fibrous scar tissue formation relative to glandular epithelial regeneration. A higher magnification of the fibrous scar and regenerating glands illustrates the abundance of fibrous scar tissue relative to the paucity of regenerated glandular ducts (*arrows*) in the breast. Fibrous scar tissue formation usually is more dominant relative to regeneration of parenchymal glands in many species. (Orig. Mag. 16×) (**c**) Metaplasia of regenerating glandular epithelium. The regenerating new epithelial cells have changed (metaplasia) from low columnar ductal epithelial cells to larger, more deeply staining stratified squamous epithelial cells (*arrow*). This squamous metaplasia is frequently seen in regenerated tissues of solid glandular structures such as prostate and breast no matter the method of tissue injury is used. The boundary between the viable tissues and necrotic tissues is distinct. (Orig. Mag. 500×) (**d**) Persistence of heat fixed tissues. The central heat fixed tissues have disintegrated and shrunken to some extent but still maintain their glandular configuration (*arrows*). No viable cells including infiltrating inflammatory cells have penetrated into the heat fixed tissue volume. (Orig. Mag. 500×)

to describe the temperature/time thermal history, geometry and extent of clinically significant photothermal effects in laser irradiated cells, tissues and organs. Most of these markers or end points are directly related to primary thermal effects such as protein denaturation and cellular membrane rupture ("melting"). They are created at relatively low temperatures (<100°C) but at variable heating times. And, because some of these pathologic markers form distinct lesions with measurable boundaries, they can be described by rate process models whose kinetic coefficients are experimentally derived. Each end point has a unique set of coefficients; therefore, the end points have to be precisely defined to determine the rate coefficients

specific for that particular effect. For example, given equal heating times, the rate coefficients obtained for collagen hyalinization, a protein denaturation process that requires higher temperatures, will be different than the coefficients for cell death due membrane rupture, a lower temperature effect.

13.3 Thermodynamics and Kinetics of Thermal Damage Processes

The thermal aspects of laser-tissue interaction consist of two governing events: (1) the thermodynamics of temperature elevation, and (2) the kinetics of resulting irreversible thermal alterations in tissues. Kinetic and thermodynamic analyses are very different in style in classical treatments. Kinetic analysis concerns the rate at which processes occur; as, for example, reaction rates in chemical transformation of materials. Thermodynamic analysis is usually concerned only with the thermodynamic state change during a process – that is, the rate at which the change takes place is normally not of interest, only the initial and final values of the thermodynamic state variables. There are six thermodynamic state variables, any two of which fully specify the thermodynamic state – i.e. if any two of them are known, the other four may be found uniquely. Consequently, the several state variables comprise a redundant set. Nevertheless, they each provide a somewhat different view of the thermodynamic state, and are thus useful in turn for studying state changes in different ways. The six thermodynamic variables are: temperature, pressure, density, density's inverse specific volume, enthalpy and entropy.

13.3.1 Foundations of Thermal Analysis

The most fundamental of the thermodynamic state variables is temperature; so much so that a whole chapter in this book has been dedicated to its measurement. Two objects are in thermal equilibrium if they are at the same temperature – that is, there will not be a *net* exchange of heat if the objects are placed in contact with each other. Heat is "energy in transit." The so-called "Zeroth Law of Thermodynamics" states that: "if object A is in thermal equilibrium with object B, and object C is also in thermal equilibrium with object B, then object A and object C are in thermal equilibrium with each other."

In the case of laser-tissue interaction, we are interested in the rate of rise of temperature, so the approach in this chapter will be a slightly modified form of thermodynamic analysis: modified to include transient solutions. The thermodynamic processes will be taken to occur under equilibrium or near-equilibrium conditions – that is, state changes occur slowly enough that all thermodynamic variables have the same inter-relationship they would if the rate of change were slowed down even further. Thermal damage kinetics are decidedly non-equilibrium in nature with their rate of change determined by local temperature.

13.3.1.1 Thermal Energy and Enthalpy

Absorption of laser energy results in an increase in the internal thermal energy, U [J] or u [J · kg^{-1}], of the tissue:

$$U = mcT, u = cT \tag{13.1}$$

where m = mass [kg], c = specific heat [J · kg^{-1} · K^{-1}] and T = temperature [K or °C]. This concept is here formulated in terms of a finite "control volume" of material that consists of enough molecules and/or atoms that a spatial and temporal mean value makes sense. Internal thermal energy in this form describes the mean value of the thermal energies of all of the constituent particles in the control volume at the time it is evaluated. Equation (13.1) is not descriptive in a control volume containing only a very few particles that are widely dispersed with an extremely low probability (i.e. rate) of collision with other particles. A statistical (i.e. quantum mechanical) formulation is required to describe those cases. Our case is for aqueous solutions / suspensions, and tissues where the mean time between collisions is vanishingly small; and in the case under study such mean descriptions are extremely useful.

Internal thermal energy is a form of "potential," similar to gravitational potential energy or to electrical potential, so it needs a zero reference point. Thermodynamic tables typically assume zero internal thermal energy at a convenient reference temperature. For example, in the classic "Steam Tables" [66] the reference temperature for the internal thermal energy of liquid phase water is 0°C (273.16 K), and the same reference point for water is used for the enthalpy and the entropy, which remain to be discussed. For most refrigerants (freon, etc.) the reference point is often taken to be −40°C (which just happens to be −40°F as well), the spontaneous freezing point for liquid phase water.

If mass crosses the control volume boundary and/or the pressure or density changes inside the control volume, the total stored (i.e. "available") thermal energy must include "Flow Work" (the pressure-volume product). A more useful description of the available thermal energy is the "enthalpy," H (J) or "specific enthalpy," h [J·kg^{-1}]:

$$H = U + PV;\ h = u + Pv \tag{13.2}$$

where P=pressure [Pa=N · m^{-2}] and V=volume [m^3], v=specific volume [m^3 · kg^{-1}] = $1/\rho$, where ρ = density [kg · m^{-3}]. In studying the units of the equation here, remember that 1 [J] = 1 [N·m]. Enthalpy may properly be thought of as the total thermal energy stored in the volume that is available to do useful work – i.e. it describes the control volume as a thermal engine.

Example 13.1: One gram of incompressible tissue absorbs a total of 2 W for 30 s. If the initial temperature is 37 C, calculate the final temperature after 30 s of absorption. Assume that the effective tissue density is 1050 (kg·m^{-3}) and specific heat is 4050 (J·kg^{-1}). No mass crosses the tissue boundaries in this example.

Solution: All of the energy absorbed goes into raising the temperature – Δv, $\Delta m = 0$:

$$\Delta U\,(\mathrm{J}) = \dot{Q}_{\mathrm{gen}}\left(\mathrm{J}\cdot s^{-1}\right)\Delta t = mc\Delta T$$

$$\Delta U = [2]\,[30] = \left[10^{-3}\right][4050]\,\Delta T$$

$$\Delta T = \frac{60}{4.05} = 14.8\ (^{\circ}\mathrm{C})$$

When mass crosses the volume boundary, it imports thermal energy according to the *net* mass inflow rate, and loses thermal energy proportional to the net mass outflow, or efflux:

$$\frac{\Delta H}{\Delta t} = \frac{\Delta m_{\mathrm{in}}}{\Delta t} c_{\mathrm{in}} T_{\mathrm{in}} - \frac{\Delta m_{\mathrm{out}}}{\Delta t} c_{\mathrm{out}} T \tag{13.3}$$

where m_{in} and m_{out} refer to the mass inflow and outflow, respectively [kg], c_{in} and c_{out} are the respective specific heats [$\mathrm{J\cdot kg^{-1}\cdot K^{-1}}$] and the rate of change in enthalpy is in units of watts [$\mathrm{W = J\cdot s^{-1}}$]. In this formulation we have included only the internal thermal energy contribution, such as would be due to tissue blood perfusion and lymphatic flow. A potential contribution to the total thermal energy, the enthalpy, due to density or pressure change has not been included.

Example 13.2: For the tissue in Example 13.1, assume that a net blood flow of [$0.1\mathrm{g\cdot s^{-1}}$] crosses the boundaries at an inlet temperature of 37°C. Inlet and outlet blood flow are equal in the tissue. Assume that the blood specific heat is the same as water, 4,186 [$\mathrm{J\cdot kg^{-1}}$]. Calculate the tissue temperature at 30 s, as before.

Solution:

(1) We will have to make some assumption about the outlet temperature of the blood. The problem has not been specific about how long the blood remains resident in the tissue. However, the reasonable assumption is that the blood leaves the tissue in thermal equilibrium with it (this is equivalent to assuming that the blood circulation is in "thermally small" vessels).
(2) The temperature rise with no blood flow is substantial, 14.8°C; so, while an accurate solution to this problem requires integrating the transient heat transfer case, we can get a working estimate with a simplified thermodynamic analysis.

$$\Delta U\, mc\,[T - T_0] - \dot{m}_b c_b\,[T - T_0]\,\Delta t = \dot{Q}_{\mathrm{gen}}\Delta t$$

$$\Delta U = [1\times 10^{-3}]\,[4050]\,[T - 37]$$

$$+\,[0.1\times 10^{-3}]\,[4186]\,[T - 37]\,[30] = [2]\,[30]$$

$$\Delta T\,[4.05 + 3.0\times 4.186] = 60\,(\mathrm{J})$$

$$\Delta T = \frac{60}{16.6} = 3.61\ (^{\circ}\mathrm{C})$$

(3) Note the strong effect of the addition of blood flow. (See Fig. 13.2. The regional blood flow was stopped by the initial lesion, the thermal coagulation lesion. Therefore, the adjacent tissue was not cooled during the laser irradiation that resulted in more severe damage, the thermal ablation lesion.)

The effect of heat transfer within the volume is to distribute the internal thermal energy uniformly. Of course, in spatially-distributed laser irradiation patterns, adjacent volumes of tissue will have differing temperatures. Consequently, conduction heat transfer can be a dominant thermal phenomenon, especially in long-term low fluence rate laser activation. Conduction heat transfer is a boundary (surface area) phenomenon:

$$\dot{\boldsymbol{q}}_{\text{cond}} = -k\frac{\partial T}{\partial n} \qquad \dot{\boldsymbol{q}}_{\text{cond}} = -k\nabla T \tag{13.4}$$

where $\dot{q}_{\text{cond}}$ = convection heat flux [W · m^{-2}] in scalar form, k = thermal conductivity [W · m^{-1} · K^{-1}] and n is the normal direction for the surface over which the heat transfer is calculated; $\dot{\boldsymbol{q}}_{\text{cond}}$ is the vector heat flux, and ∇T is the temperature gradient (the gradient of a scalar field is a vector). The negative sign reflects the direction of conduction heat transfer: from higher temperature to lower temperature, i.e. anti-parallel to the temperature gradient (a vector).

Similarly, if there is heat transfer by convection from an exposed control volume surface, it may be found from:

$$\dot{q}_{\text{cond}} = h[T - T_{\text{atm}}] \tag{13.5}$$

where $\dot{q}_{\text{cond}}$ = convection heat flux [W · m^{-2}], h = convection heat transfer coefficient [W · m^{-2} · K^{-1}], and T_{atm} is the atmospheric, or surrounding fluid, temperature [K]. There is also the possibility of thermal radiation from the surface. The direction of convection and radiation heat transfer is normal to the surface ($\boldsymbol{n}$), so the appropriate surface boundary condition is:

$$\mathbf{n} \cdot (k\nabla T) = h[T - T_{\text{atm}}] + \sigma_B\, \varepsilon \left[T^4 - T_{\text{atm}}^4\right] \tag{13.6}$$

where we have now included the contribution of thermal radiation from the surface, as described in Chapter 11 (i.e. σ_B = the Stefan - Boltzmann constant [W · m^{-2} · K^{-4}], and ε = the surface emissivity, and T is in [K]).

Convection is a special case of conduction heat transfer, where the medium is in motion. Of course, estimation of the convection coefficient becomes quite complex when the fluid near the surface is compressible since the nature of the fluid flow boundary layer (turbulent or laminar) has a very strong influence on the effective convection coefficient. Also, proper formulation of the radiation boundary conditions is even more complex, depending as it does on the shape and coupling factors of surrounding thermal radiation objects, and their temperatures.

Example 13.3: For the tissue under the same conditions as in Example 13.1, assume that the tissue volume is contained in a 1 cm cube. Assume that the tissue volume is at the surface, and has 5 neighbor tissue volumes below and around it. The thermal conductivity is 0.46 [W m^{-1} K^{-1}]. Calculate the net conduction heat transfer out of the tissue volume at the end of the activation (30 s) if the final temperature in the 5 adjacent tissue volumes has risen to half of the temperature in this control volume (i.e. 37 + 14.8 = 51.8°C). Assume that the temperatures represent the values at the centers of the respective control volumes. Neglect inlet and outlet blood flow in the tissue.

Solution:

$$\dot{q}\left(\mathrm{W}\cdot\mathrm{m}^{-2}\right) = k\frac{\Delta T}{\Delta x}$$

(1) Each adjacent control volume (CV) has half the temperature rise of this volume, i.e. 7.4°C. The temperature difference between the adjacent CVs is therefore 7.4°C.
(2) For volumes 1 cm on a side, the distance between adjacent centers is also 1 cm. This means that the temperature gradient is $7.4\,(^{\circ}\mathrm{C}\cdot\mathrm{cm}^{-1}) = 0.074\,(^{\circ}\mathrm{C}\cdot\mathrm{m}^{-1})$.
(3) For the 5 sides the net conduction heat transfer out of the CV is:

$$\dot{Q}\,(\mathrm{W}) = [5]\,[10^{-4}\mathrm{m}^2]\,[0.46]\,[0.074] = 17\,(\mu W)$$

Example 13.4: For the same control volume (CV) as in Example 13.3, calculate the heat transfer by convection from the exposed surface. Assume that $h = 25$ $(\mathrm{W}\cdot\mathrm{m}^{-2}\cdot\mathrm{K}^{-1})$ and the room temperature is 23°C.

Solution:

$$\dot{q}\left(\mathrm{W}\cdot\mathrm{m}^{-2}\right) = h\,[T - T_{\mathrm{atm}}]$$

$$\dot{q}(W) = [1\times 10^{-4}\mathrm{m}^2][25][51/8 - 23^{\circ}\mathrm{C}] = 72(\mathrm{mW})$$

Much more significant than conduction heat transfer.

13.3.1.2 Entropy: The Second Law of Thermodynamics

The energy balance in the first law of thermodynamics (Chapter 10) represents the "break even" thermal point of view. It assumes that all processes are reversible and that all required energy can be accounted for in terms of the thermal energy. That is, there are no "preferred directions" for energy change or for material transformation. It would be equivalent to considering that all chemical reactions are equally likely to occur in either direction. We need some method for identifying preferred pathways

for material transformations, so that we may describe changes that are "more likely," and the concept of "entropy" fills that need.

Classical Thermodynamic Description

As a concept, entropy often gives trouble but it really shouldn't. In perhaps its simplest form, the specific entropy, s ($\mathrm{J{\cdot}kg^{-1}{\cdot}K^{-1}}$), is a measure of randomness or disorder in a system. It is easy to see that a crystalline solid is more ordered than the same substance in liquid or gas phase. As a substance transitions from solid to liquid to gas its entropy increases. It takes energy to create the orderliness of a crystal, so both the enthalpy and the entropy change as a crystal is formed from, say, the liquid phase. Similarly, in chemical reactions the reactants form products because that is more likely than having the products decompose into the reactant atoms or molecules. Hydrogen and oxygen mixtures combine explosively when stimulated to form water; water is an extremely stable molecule that only decomposes to hydrogen and oxygen when considerable energy is added to the system to favor that transition.

Thermodynamic systems (i.e. "heat engines" in the classical sense) are constrained to extract energy from a "thermal energy reservoir" and reject energy to a "thermal energy sink." That is, there will always be "waste heat" for any thermal engine. This is one of the many statements of the "Second Law of Thermodynamics." Others include: (1) the entropy of the universe is always increasing, (2) a perpetual motion machine of the second kind is impossible, and (3) any process that would decrease the entropy of a completely isolated system is impossible (a completely isolated system cannot exchange either heat or mass with the outside world). There are others, as well, but this is enough for illustration. All of them boil down to the simple truth that no thermal system can be 100% efficient.

Heat exchange, for example, is always entropy increasing. When heat, dQ [J], is extracted from a thermal energy reservoir at temperature T_1, the entropy decreases by: $dS_1 = -dQ/T_1$ [$\mathrm{J \cdot K^{-1}}$]. The entropy gained by the extracting object at temperature T_2 is: $dS_2 = +dQ/T_2$. For conduction heat transfer, it must always be true that $T_1 > T_2$, so the net entropy in the universe, $S_u = dS_1 + dS_2$, has increased because of the heat transfer. This leads to the "Clausius Inequality" [67]:

$$\oint_{\Sigma} \frac{dQ}{T} \leq 0 \tag{13.7}$$

Here the closed surface integral is used to denote the *net* entropy change for the *system* (that is, the system must reject heat to the universe in order to function), and Σ is the closed outer boundary of the thermodynamic system.

A single DNA molecule is very well-ordered, and thus at low entropy compared to the constituent molecules required to form it. This might, at first glance, seem to violate the Second Law, but it doesn't because the relative "orderliness" of the unused fractions of the reactants is so much less than their original form

that the randomness in the universe has increased because of the formation of the DNA molecule. The cell is an open system, and the entropy of the byproducts of metabolism is greater than that of the organized molecules created with them, such as DNA, RNA, ATP and, of course, many others. In addition, every biologic cell generates excess heat of metabolic reactions, the waste heat of the cell as a thermal engine, that must be rejected to the environment. All cells obey the Second Law, just as any heat engine must. If a cell is placed in a completely insulated container, it will die as its temperature increases (essentially without bound) in order to satisfy the Second Law.

Perhaps the most pertinent example for this chapter is the change from liquid phase to vapor phase water (boiling) at constant temperature and pressure – this requires the addition of sufficient thermal energy to increase the entropy of the water to that of the vapor phase. At saturation both the pressure and temperature are constant and both the change in enthalpy, $\Delta h\,[\mathrm{J}\cdot\mathrm{kg}^{-1}]$, and the resulting change in entropy, $\Delta s\,[\mathrm{J}\cdot\mathrm{kg}^{-1}\cdot\mathrm{K}^{-1}]$, depend on the temperature at which the boiling takes place:

$$\Delta h = T\Delta s \tag{13.8}$$

In the Steam Tables [66] the somewhat arbitrary reference point for entropy is $s = 0$ at $T = 0°\mathrm{C}$ (273.16 [K]) – this is the triple point for water at pressure $P = 1$ atmosphere (i.e. 101.3 [kPa]). At the triple point, the liquid, vapor and solid phase are in thermal equilibrium, so it makes a convenient reference for both the enthalpy and the entropy of water.

Example 13.5: Calculate the thermal energy ($\Delta h = h_{fg}$) required to boil 10 g of water at 1 atmosphere pressure and saturation temperature (i.e. 100°C) if you know that the change in entropy is $\Delta s = s_{fg} = 6036$ (J/kg/K). Here the "*fg*" subscript indicates the change in phase from fluid to gas. Compare the answer with the tabulated phase change enthalpy at saturation, $h_{fg} = 2.257 \times 10^6$ (J/kg).

Solution:

$\Delta h = T\Delta s = (100+273.16)^*(6036) = 2.252\times10^6$ (J/kg) . . . within round-off error of h_{fg}
(Note: Don't forget to use absolute temperature here!)
10 g = 0.01 kg => Energy required: $\Delta h = 2.252 \times 10^4$(J)

A similar change in entropy accompanies chemical reactions. The enthalphy change includes Gibb's Free Energy of Formation, $\Delta g\,[\mathrm{J}\cdot\mathrm{kg}^{-1}]$, and the entropy change of the reaction:

$$\Delta h = \Delta g + T\Delta s \tag{13.9}$$

Remember that some reactions are exothermic and others are endothermic as a result of the relative Gibb's energy of formation. We will make use of this relation in a later section of this chapter.

Statistical Thermodynamic Description

Up to this point the discussion of thermodynamics has been limited to the macroscopic (i.e. classical thermodynamics) point of view: quantities such as enthalpy and internal thermal energy are deterministic macroscopic properties and considered continuous. In classical thermodynamics they may be traded between control volumes in infinitessimally small amounts with no lower limit. The classical thermodynamic relationships are routinely used to describe systems as diverse as steam engines, gas turbines, refrigerators, subsonic and supersonic nozzle expansions and living cells.

But entropy is really a stochastic (random) quantity, as described above, and not deterministic (non-random). We should therefore expect a stochastic (microscopic) description, then, in terms of random variables. This point of view was originally proposed by Boltzmann in the middle of the 19th century, and was quite controversial at the time. Very briefly, the entropy is an "extensive" property; that is, it is the sum of the entropies of all of the particles (molecules or atoms) of a system. Further, the permitted "states" of individual particles of matter are not continuous, but are limited to discretely quantized values. This is a fundamental tenet of quantum mechanics. Systems of many interacting molecules undergo continuous change in their quantum state, exchanging energy in numbers of finite, discrete "quanta," and in no smaller amount.

If we follow the thermal energy of an individual molecule for a period of time, we would expect to observe a mean value with discrete (i.e. quantum) fluctuations about the mean described by the various statistical moments: mean, standard deviation, variance, and so on. A very highly ordered ensemble of molecules will have a very small standard deviation, and a more disordered ensemble will have a large standard deviation (looking at just the first and second statistical moments for this discussion). The Quantum State Probability, $p_i(t)$ can be used to describe this, where $p_i =$ the fraction of total molecules in state "i" at any time. If the total number of possible quantum states is low, then the p_i values will be high (and the standard deviation low); and the converse. At all times, the total probability $= 1$, so for N total states:

$$\sum_{i=1}^{N} p_i = 1 \tag{13.10}$$

The macroscopic property used in the previous discussion is the probability-weighted (expected) value, which is called the mean value – for example, the macroscopic internal thermal energy, u, is the mean value, $<u>$, of the allowed quantum state molecular thermal energies, u_i. Note that $<u>$ can be continuous even though u_i is discrete since the mean value of a random process need not be an observable value. If there are N total allowed quantum states, then at any instant:

$$u = \sum_{i=1}^{N} p_i u_i [\mathrm{J} \cdot \mathrm{kg}^{-1}] \tag{13.11}$$

If a complete thermal system, C, contains constituents A and B, where there are N allowed states for A molecules and M allowed states for B molecules, then the complete system, C, is described by their joint probabilities, $p_{ij} = p_i \cdot p_j$. That is, the joint probability is the product of the individual probabilities when the two random processes are statistically independent. To construct a useful model for the entropy of the thermodynamic system, we use the joint probability in the form of an equivalent function, f (as yet undefined):

$$s = \sum_{i=1}^{N}\sum_{j=1}^{M} p_i\, p_j\, f\left(p_i\,,\ p_j\right) = \sum_{i=1}^{N}\sum_{j=1}^{M} p_i\, p_j\, f\left(p_i\ \cdot\ p_j\right) = \sum_{i=1}^{N} p_i f\left(p_i\right) + \sum_{j=1}^{M} p_j f\left(p_j\right) \tag{13.12}$$

The function f must be such that statistical independence – i.e. Eq. (13.12) – holds irrespective of the values of p_i and p_j. One such function that always satisfies this requirement is the natural logarithm, $f(\arg) = \ln(\arg)$ – i.e. $\ln\{AB\} = \ln\{A\} + \ln\{B\}$. It can be shown after some development (see, for example, [68], Chapter 6) that the general solution to the governing second order differential equation is $f = C\ln\{p_i\}$, and since p_i is always ≤ 1, the constant C must be less than zero so that f is positive-definite. Boltzmann demonstrated that the definition of temperature in the statistical and classical domains would be the same if:

$$s = -k\sum_{i=1}^{N} p_i\ \ln\{p_i\}[\mathrm{J}\cdot\mathrm{kg}^{-1}\cdot\mathrm{K}^{-1}] \tag{13.13}$$

where k = Boltzmann's constant$(1.380 \times 10^{-23}\mathrm{J}\cdot\mathrm{K}^{-1})$ and the negative sign is necessary since the quantum state probabilities are constrained to $0 < p_i \leq 1$.

If all of the allowed quantum states are equally likely – that is, if p_i is uniformly distributed – and there are W allowed states, then Boltzmann demonstrated that:

$$s = k\ln\{W\} \tag{13.14}$$

A uniformly distributed random variable is one that maximizes the entropy in a system of molecules. Equation (13.14) is the Boltzmann definition of entropy, while Eq. (13.13) is the Gibbs definition of entropy. When first introduced, Eq. (13.14) was an extremely controversial construction; so much so that Boltzmann was vilified by the thermodynamic classicists of the day. His final word on the matter was to use Eq. (13.14) as his epitaph.

In summary, while the detailed thermodynamic state must be described by quantum mechanical statistical thermodynamics, the mean thermodynamic properties are adequate for laser heating calculations unless the distribution of states deviates significantly from the equilibrium distribution. That is to say, classical thermodynamic variables adequately describe the processes unless all of the events occur in *extremely* short time intervals (i.e. time intervals approaching the mean time between collisions).

The Zeroth Law of Thermodynamics states that "there is a thermal game;" the First Law says that "the best that you can do is to break even," and the Second Law says that "you can never do that well." An unstated Third Law is that "you can *never* get out of the game."

13.3.2 Modeling Tissue Thermal Events

13.3.2.1 Sub-Coagulation Thermal Models

An energy balance for tissue under laser irradiation where metabolic heat and phase change are negligible can be assumed to follow the traditional Bioheat Equation as previously described, and repeated here:

$$\rho c \frac{dT}{dt} = \dot{Q}_{\text{gen}} + \nabla \cdot (k \nabla T) + \dot{w} \rho c_b \, [T_b - T] + \dot{Q}_{\text{met}} [\text{W} \cdot \text{m}^{-3}] \qquad (13.15)$$

where the "b" subscript refers to blood properties, and $\dot{w}$ is the tissue perfusion $[\text{kg}_\text{b} \cdot \text{kg}_\text{tissue}^{-1} \cdot \text{s}^{-1}]$. In Chapter 3, the notation for laser heat generation is S; but since S is used for entropy in this chapter we have used $\dot{Q}$.

13.3.2.2 Evaporation of Surface Water

Surface water evaporation typically dominates other surface loss terms (convection and radiation) above about 60°C [69]. A useful model for evaporation using Stelling's Formula has been derived from experimental solar pond data, as described in [70]:

$$W_{\text{evap}} = [A_s + B_s U_{\text{inf}}] \; (P_{\text{sat}}(T) - E_{\text{atm}}) \, [\text{m} \cdot \text{s}^{-1}] \qquad (13.16)$$

where W_{evap} = surface evaporation rate $[\text{m} \cdot \text{s}^{-1}]$, $A_s = 7.31 \times 10^{-11}$ $[\text{m} \cdot \text{Pa}^{-1} \cdot \text{s}^{-1}]$, $B_s = 1.2 \times 10^{-11}$ $[\text{Pa}^{-1}]$, U_{inf} = free stream air velocity above the surface $[\text{m} \cdot \text{s}^{-1}]$, $P_{\text{sat}}(T)$ = the saturation pressure at tissue temperature, T, [Pa], and E_{atm} = atmospheric water vapor pressure = $(Hu) P_{\text{sat}}(T_{\text{inf}})$ [Pa], where Hu = relative humidity (expressed as a fraction). To include this process in model calculations, the evaporation rate, W_{evap}, is multiplied by the control volume surface area, A $[\text{m}^2]$, liquid-phase density ρ $[\text{kg} \cdot \text{m}^{-3}]$, phase change enthalpy h_{fg} $[\text{J} \cdot \text{kg}^{-1}]$, and divided by the control volume, V $[\text{m}^3]$, to get the equivalent volumetric power, Q_{evap} $[\text{W} \cdot \text{m}^{-3}]$:

$$\dot{Q}_{\text{evap}} = \frac{A h_{fg} \rho W_{\text{evap}}}{V} [\text{W} \cdot \text{m}^{-3}] \qquad (13.17)$$

Here "f" indicates the liquid state, and "g" the gas state, as above.

13.3.2.3 Water Vaporization

When liquid-phase water is vaporized (boiled) we must add the required thermal energy at the boiling point temperature and pressure, T_{sat} and P_{sat}, in the form of the vaporization enthalpy, Δh_{fg}, and as a result the entropy will change, Δs_{fg}, as well as the density (or specific volume). The density/specific volume change is actually included in the phase change enthalpy:

$$\Delta h_{fg} = T\Delta s_{fg} \ = \Delta u_{fg} + P_{\text{sat}}\Delta v_{fg} \tag{13.18}$$

where Δu_{fg} is the internal thermal energy of phase change $= [c_g - c_f]T_{\text{sat}}$, and Δv_{fg} is the change in specific volume at the boiling point pressure, P_{sat}. For example, at 100°C water has (see Example 13.4):

$$\Delta h_{fg} = 2.257 \times 10^6\,[\text{J} \cdot \text{kg}^{-1}],\ \rho_{\text{sat}} = 958\,[\text{kg} \cdot \text{m}^{-3}],\ \Delta s_{fg} = 6036\,[\text{J} \cdot \text{kg}^{-1} \cdot \text{K}^{-1}]$$

Aside from the effect vaporization has on tissue optics, expansion of the tissue water, whether or not phase change takes place, accounts for the internal stresses applied to tissue. In order to include these effects in the energy balance, the left hand side of the energy balance, Eq. (13.15), must be modified. Neglecting perfusion and metabolic heat for the moment:

$$\frac{\partial\,(ph)}{\partial t} \ = \ \dot{Q}_{\text{gen}} \ + \ \nabla \ \cdot \ (k\nabla T) \tag{13.19}$$

The enthalpy is the sum of internal thermal energy and the pressure-volume product, as in Eq. (13.2). Note that the thermal model now includes phase change since the pressure-volume product has been included in the enthalpy.

Numerical implementation of this model may follow the same philosophy as that of the low temperature case if two thermodynamic parameters can be evaluated, $\partial P/\partial T$ and $\partial \rho/\partial T$:

$$\left\{\rho c \ + \ \frac{\partial P}{\partial T} \ - \ \frac{P}{\rho}\frac{\partial \rho}{\partial T}\right\}\frac{\partial T}{\partial t} \ = \ \dot{Q}_{\text{gen}} \ + \ \nabla \ \cdot \ (k\nabla T) \tag{13.20}$$

In free liquid phase water that is not constrained by rigid or visco-elastic tissue structures, $\partial P/\partial T$ is usually negligibly small. However, a fourth order polynomial fit ($r^2 = 1.000$) derived from "The Thermodynamic Properties of Steam" [66] for the saturation pressure is:

$$\begin{aligned} P_{\text{sat}}\,(T) \ = & 1.686 \times 10^{-9}T^5 \ + \ 6.681 \times 10^{-8}T^4 \\ & + \ 2.411 \times 10^{-4}T^3 \ - \ 0.03623T^2 \ + \ 2.389T \ - \ 0.432 \end{aligned} \tag{13.21}$$

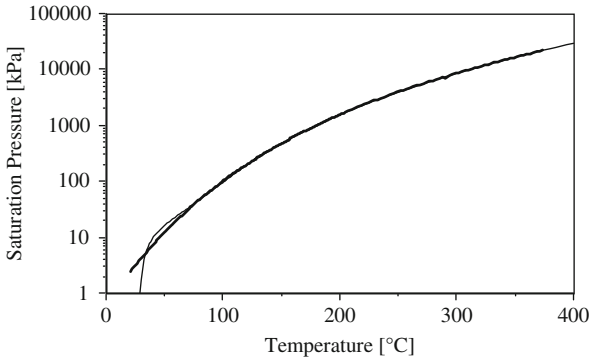


Fig. 13.10 Saturation pressure, P_{sat} [kPa] vs. temperature, T [°C]. *Thick line* plot is experimental data (—), and polynomial fit of Eq. (13.21), *thin line* (—)

where P_{sat} is in [kPa = kN·m^{-2}] and T is in Celsius. In Fig. 13.10 this function deviates from the tabulated data only at temperatures near the freezing point, 0°C. The $\partial P/\partial T$ term may be quickly derived from this polynomial fit.

The density change term cannot be neglected in free water. This additional term (when phase change is included) accounts for the kinetic energy imparted to expelled water and describes the origin of acoustic waves in tissue, whether or not phase change is significant. For no vaporization, an estimate of $\partial\rho/\partial T$ can be obtained from the saturation density vs. temperature curve under the valid assumption that the liquid phase water density is primarily determined by temperature rather than local pressure. The saturation density curve, a fifth order polynomial fit from 20 to 374 C ($r^2 = 0.998$) [66] is:

$$\begin{aligned}\rho_{sat}(T) = &6.918 \times 10^{-10}T^5 + 5.594 \times 10^{-7}T^4 - 1.650 \\ &\times 10^{-4}T^3 + 0.01912T^2 - 1.413T + 1026\end{aligned} \tag{13.22}$$

where ρ_{sat} is in [kg·m^{-3}] and, again, T is in Celsius. In Fig. 13.11 the polynomial fit only deviates significantly from the tabulated data at temperatures near the critical point, 375°C.

At temperatures above about 90°C water behavior dominates the observed photo-thermal phenomena. Below the saturation temperature, 100°C at atmospheric pressure, water vaporization is a diffusion-limited surface loss phenomenon (see Eq. (13.16)) – and depends primarily on surface characteristics such as local humidity and temperature dependent mass diffusion coefficients. Above the saturation temperature, the vaporization process becomes a volumetric energy sink in which phase changes generate high local pressures and concomitant tissue stresses. However, the more moderate heating case is of immediate interest in this discussion.

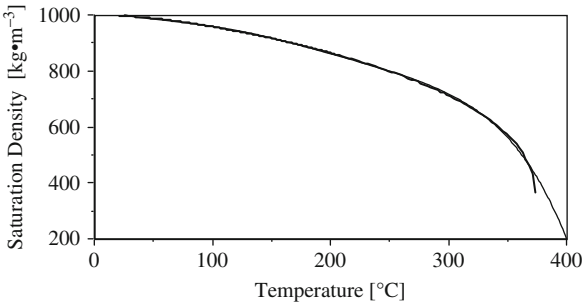


Fig. 13.11 Density of saturated liquid water, ρ_{sat} [kg·m^{-3}], vs. temperature, T [°C]. *Thick line* plot is experimental data (—), and polynomial fit of Eq. (13.22), *thin line* (—)

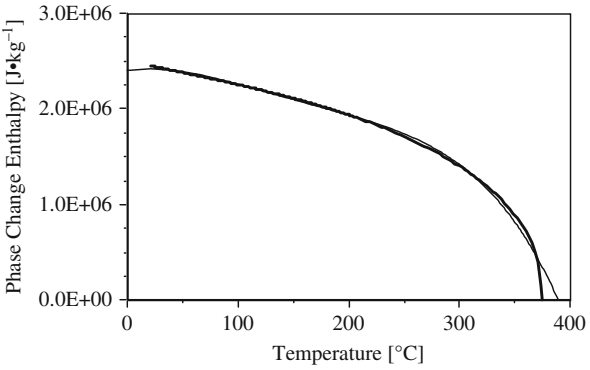


Fig. 13.12 Phase change enthalpy, ρ_{sat}[kg · m^{-3}], vs. temperature, T [°C]. *Thick line* plot is experimental data (—), and polynomial fit of Eq. (13.24), *thin line* (—)

In Fig. 13.10 the saturation pressure at 120°C has increased to 2 atmospheres (200.2 kPa) and climbs precipitously above that. Water vaporization may be included in the modified energy balance of Eq. (13.20):

$$\left\{\rho c + \frac{\partial P}{\partial T} - \frac{P}{\rho}\frac{\partial \rho}{\partial T}\right\}\frac{\partial T}{\partial t} = \dot{Q}_{\text{gen}} + \nabla \cdot (k\nabla T) - \frac{\partial m}{\partial t}\Delta h_{fg}(T) \qquad (13.23)$$

where $\partial m/\partial t$ is the rate of mass vaporization of water per unit volume (kg·s^{-1}·m^{-3}) and $\Delta h_{fg}(T)$ is the temperature dependent phase change enthalpy [J·kg^{-1}] (Fig. 13.12 and Eq. 13.24).

The temperature dependence of saturation density and saturation pressure may be included by incorporating the appropriate correlation relation from Eqs. (13.21) and (13.22). The temperature dependence of the phase change enthalpy may be included using a similar correlation ($r^2 = 0.995$) derived from the same tabular data [66]:

$$\Delta h_{fg} = -4.143 \times 10^{-4} T^4 + 0.247 T^3 - 54.3 T^2 + 1975 T + 2.402 \times 10^6 \quad (13.24)$$

where Δh_{fg} is in [J·kg^{-1}] and T is in Celsius. The polynomial fit only deviates significantly from the tabulated data at temperatures near the critical point.

An alternate, and in some ways equivalent, formulation for vaporization was recently presented by [71] in which phase change processes are included in the left hand side of Eq. (13.15) to obtain an effective specific heat, C', where C' combines the first term on the left hand side with the last term on the right hand side of Eq. (13.23):

$$\rho C' \frac{\partial T}{\partial t} = \rho \left\{ c - \Delta h_{fg} \frac{\partial \rho_w}{\partial T} \right\} \frac{\partial T}{\partial t} = \dot{Q}_{\text{gen}} + \nabla \cdot (k \nabla T) + W \rho_b c_b [T_b - T] \quad (13.25)$$

where W is described in their paper as the "density" of the tissue water, and the perfusion heat has been included. This formulation includes the latent heat of vaporization, but not the changes in tissue density or pressure. However, it has to be used with some caution. First, "W" includes much more than the density of water – in fact, W includes the mass fraction of water in the tissue, so $W = m$ in Eq. (13.23). Second, applying the relation directly as given leads to a decidedly aphysical relationship. That is, for equilibrium vaporization at the saturation temperature $\partial T / \partial t = 0$; consequently, the left hand side of Eq. (13.25) is zero for equilibrium boiling at saturation. The right hand side is certainly not zero – imagine tissue with zero blood flow at uniform temperature (so that both the Fourier conduction term and the blood flow term are zero); $\dot{Q}_{\text{gen}}$ is not zero. The energy balancing effect of the $\partial m / \partial t$ on the right hand side of Eq. (13.23) has been eliminated, leading to an invalid conclusion from Eq. (13.25).

In a numerical model, one may assume that vaporization occurs at local saturation conditions (equilibrium thermodynamics applies) with rate determined by laser excess energy (that is, energy in excess of that consumed by conduction heat transfer, perfusion heat and other losses) and thermodynamic properties. This is equivalent to assuming that evolved water vapor diffuses readily through the tissue structure, and the pressure does not rise significantly. Under these conditions the left hand side of Eq. (13.23) would be zero – i.e. constant temperature boiling at the saturation temperature.

When laser energy, $\dot{Q}_{\text{gen}}$, is added faster than water vapor can diffuse through tissue to maintain liquid saturation at atmospheric pressure, the tissue temperature readily exceeds 100°C when the local tissue pressure rises (and thus lowers the phase change enthalpy). In that case, Eq. (13.23) would have to be used to construct the model. An additional relation would have to be included to complete the model: the rate of pressure rise with water vaporization. The constraints on intracellular water are more significant than on extra-cellular water. That is, extra-cellular expanded water and water vapor need only to "percolate" through spaces in the gross structure while intracellular water vapor expands cell volume putting stress on the cell membrane. For either the intra- or extra-cellular tissue water compartment one would need mass diffusion relations to formulate this term.

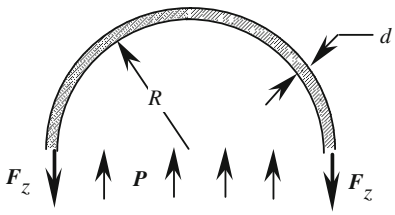


Fig. 13.13 Free-body force diagram of spherical "cell" of radius R with membrane of thickness d stressed by internal pressure, P

The applied stresses, both intra-cellular and extra-cellular, create what can be described as a "thin wall pressure vessel problem." Imagine a spherical pressurized vessel (i.e. cell), such as shown in cross-section in Fig. 13.13, with internal pressure "P" (relative to the external pressure), interior radius "R" and wall thickness, "d." If we cut the spherical vessel at $z = 0$, a free-body diagram reveals that the pressure force, $F_p = \pi R^2 P$, must be balanced by the total Z-direction force, $F_z[N] = 2\pi R d\,\sigma$, where σ [N · m^{-2}] is the wall stress, and so:

$$\sigma = \frac{P\,R}{2d} \tag{13.26}$$

In many cases the local pressure increases cause rupture in the tissue structure, histologically observable as the "popcorn" effect described in Section 13.2.3.4. Elevated tissue pressures also explain the readily observed tearing phenomena in highly layered tissues, such as arterial vessel wall, when irradiated with highly scattered wavelengths. The effects of high local tissue pressures have been observed in many studies when, for example, highly-scattered argon laser irradiation is used to ablate aorta. The scattering events result in high fluence rates just below the tissue surface and corresponding thermally-induced stresses when the tissue water vaporizes.

Example 13.6: This example will adapt "thin wall pressure vessel" analysis to calculate the internal pressure of small bubbles generated in liquid phase water. In this case the "surface tension" or surface free energy, γ(N · m^{-1}) has the role of the "applied stress," σ in Eq. (13.26). The surface tension of water is approximately linearly related to the temperature by γ(dynes · cm^{-1}) = 76.42 − 0.1731 T(°C). Calculate the internal pressure of a water vapor bubble 0.1 mm in diameter in water at 90°C.

Solution:

(1) The pressure forces, F_p, are balanced by the surface tension forces around the circumference of the bubble$_\gamma$, similar to the wall stresses in Fig. 13.13. So:

$$F_p = P\pi R^2 = F_\gamma = \gamma 2\pi R$$

$$P = \frac{2\gamma}{R}\left[\text{N} \cdot \text{m}^{-2}\right]$$

(2) At $T = 90°\text{C}$, $\gamma = 76.4 - 0.1731\,(90) = 60.9\,(\text{dynes} \cdot \text{cm}^{-1}) = 60.8 \times 10^{-7}$ $(\text{N} \cdot \text{m}^{-1})$

$$P = \frac{2[60.8 \times 10^{-7}]}{[0.05 \times 10^{-3}\text{m}]} = 0.243\left(\text{N} \cdot \text{m}^{-2}\right) = 0.243\,(Pa)$$

(3) This is a very small internal pressure.

A model based on the energy balance of Eq. (13.23) should use a small time step so that the contributions of local heat transfer may be calculated and summed with the laser power term. Excess energy above that required to maintain constant temperature may be imagined to go into the vaporization process at the local temperature and pressure. The effect of vaporization on the local pressure may be calculated if an acceptable model for tissue mechanical response is available (i.e. the stress-strain relation, and vapor diffusion coefficients).

At extreme heating rates in tissue and/or free water, such as are realized by pulsed excimer laser activations, the water may become metastable. If so, the vaporization is a rate-limited nucleate boiling phenomenon with rate coefficients determined by the supersaturation ratio, P/P_{sat}, (where P_{sat} is the saturation pressure, Fig. 13.10) and by the surface free energy of nucleated vapor bubbles in the liquid phase. Nucleate boiling only dominates equilibrium boiling at temperatures approaching the critical temperature. In the excimer laser rapid-ablation case the energy powering the debris plume probably comes from nearly instantaneous water phase changes at highly elevated pressures.

13.3.2.4 Coagulation Processes

At higher temperatures the desired thermal end point is coagulation, or a significant irreversible alteration in tissue structural proteins – that is, thermal damage. We often use kinetic models in a "volume fraction" or "probability" form (% of tissue damaged) to predict or describe the evolution of thermal damage. Typical damage processes for which the volume fraction kinetic damage model has been used successfully include: (1) tissue necrosis, (2) loss of birefringence in muscle and collagen, and (3) collagen shrinkage. The damage model is based on the assumption that a single damage process is active, which may not hold in the general case where multiple simultaneous processes occur. Still, the model can be used if the damage processes under study have different basic mechanisms and threshold temperatures and can therefore be considered thermodynamically independent. For example, it is hard to imagine how red blood cell coagulation could affect collagen fiber unraveling, so many common damage processes are quite obviously independent in the thermodynamic sense. In practical calculations, we wish to identify particular damage processes that can act as hallmarks of desired therapeutic end points. As we frequently do not have data for the damage process enthalpy, we

generally do not include the damage process in the thermodynamic calculations of tissue temperature.

Damage processes in tissue coagulation and necrosis typically follow first order unimolecular reaction kinetics. This is such an important topic that it is reserved for detailed discussion in the following section of the chapter, Section 13.3.3. It is very important to note at this point that tissue damage processes are not phase change processes, despite their rather unfortunate previous description in that fashion [37]. It is easy to see why that is so: all phase changes are reversible, all tissue thermal damage processes are irreversible. Repairing tissue thermal damage requires tremendous amounts of energy during the healing (repair and replacement) process.

13.3.3 Kinetics of Thermal Damage Processes

Damage processes can often be modeled as first-order rate processes for which two experimentally derived coefficients are sufficient. Thermal damage in this formulation is exponentially dependent on temperature and linearly dependent on time of exposure. The rate process models apply well to the prediction of damage thresholds, and less well as the damage becomes complete since several of the fundamental assumptions are violated. This section reviews the basis for kinetic models of tissue thermal damage.

13.3.3.1 Theoretical Foundations

The original work on the application of rate process models to thermal damage was reported by Moritz and Henriques in a series of seminal papers entitled "Studies of Thermal Injury" in 1947 [72–75]. In their work, the damage was quantified using a single parameter, Ω, which ranges on the entire positive real axis and is calculated from an Arrhenius integral:

$$\Omega(\tau) = \int_0^{\tau} A\, e^{\left[\frac{-E_a}{R\,T(t)}\right]}\, dt \tag{13.27}$$

where A is a frequency factor [s^{-1}], τ the total heating time (s), E_a an activation energy barrier [$J \cdot mole^{-1}$], R the universal gas constant, 8.3143 [$J \cdot mole^{-1} \cdot K^{-1}$], and T the absolute temperature [K]. One difficulty with this description is that a single damage parameter inherently lumps all damage processes into one global measure, so in multiple damage process cases the lowest temperature process saturates the damage measure early during a laser activation. In this discussion we review the underlying assumptions and origin of the terms in Eq. (13.27) and recast the traditional thermal damage parameter, Ω, into a form suitable for evaluation of multiple-process laser thermal damage effects and for comparison between numerical models and histologic results.

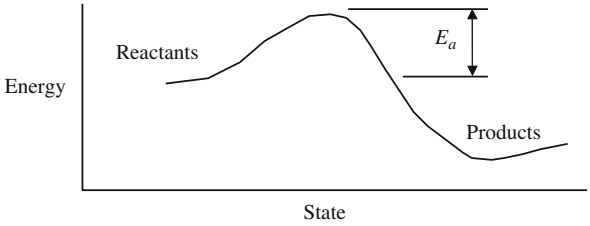


Fig. 13.14 Energy-state diagram. Reactants surmount energy barrier E_a to transition to products

Reaction Product Formation Rates

The basis for rate process models of thermal damage may be obtained from chemical reaction kinetics. In a typical reaction process, thermally active reactants jump over an energy barrier to form products, as illustrated in Fig. 13.14. In the figure, E_a is the energy barrier in Eq. (13.27). The collision theory description of ordinary first order bi-molecular reaction kinetics (see, for example, [76]) holds that the reactants are activated by collisions; n^* are activated out of n total molecules, and the probability of activation is:

$$\frac{n^*}{n} = e^{\left[\frac{-E_a}{RT}\right]} \tag{13.28}$$

In such a process, activated reactants are considered to form an activated "complex" which may either relax to inactivated single reactants or progress to form product molecules. The complex has some of the properties of an ordinary molecule and is at least temporarily stable. For reactant molecules A and B the activated complex is $[AB]^*$, and the sequence of formation is:

$$\text{Reactants} \quad A + B \underset{k_b}{\overset{k_a}{\rightleftarrows}} [AB]^* \xrightarrow{k} \text{Products} \tag{13.29}$$

The overall reaction velocity, $k\,[\mathrm{s}^{-1}]$, determines the rate of formation of product and is related to the equilibrium constant for formation of the activated complex, K^*, by:

$$k = \frac{RT}{Nh_P} K^* \tag{13.30}$$

where

$$K^* = e^{\frac{-\Delta G^*}{RT}}$$

and here N is Avogadro's number (6.023×10^{23}), h_P is Planck's constant (6.627×10^{-34} [J - s]), and ΔG^* is the Gibb's free energy of formation of activated complex [J·mole^{-1}]. As mentioned in the previous section, the free energy of formation is given by:

$$\Delta G^* = \Delta H^* - T\Delta S^* \tag{13.31}$$

where ΔH^* is the enthalpy of activation [J · mole^{-1}] and ΔS^* is the entropy of activation [J · mole^{-1} · K^{-1}]. The activation entropy is not calculable except for the simplest possible reactions in rarified gases, and is therefore usually determined from experimental measurements of the reaction velocity and activation enthalpy. The activation enthalpy, ΔH^*, is determined from the observed activation energy, E_a by:

$$\Delta H^* = E_a - i\,RT \tag{13.32}$$

where i is 1 for first-order reactions in solution and gases, 2 for second order, and 3 for third order reactions.

Unimolecular Process Descriptions

Thermal damage in tissue is a unimolecular process – tissue constituents transition from the native state to the damaged state. Absolute reaction rate theory can also be used to explain the rate of formation for this process if we assume that a time lag exists between molecular activation and denaturation [76]. During this time lag, the molecules may either denature or relax back to the native state, as illustrated in Fig. 13.15. Here, ΔH is the enthalpy (i.e. internal thermal energy) difference between native state and denatured molecules. The relative barriers are such that in the thermal damage of tissue, ΔH^*, is almost always smaller than ΔH. So, the activation process may be regarded as reasonably likely, and the probability of denatured tissue relaxing back to native state tissue is near enough to zero that it may be regarded as the impossible event in the absence of an energy-consuming healing process. The rate of damage formation is then proportional to only those molecules that remain activated. For a unimolecular process in the native state C, having an activated state, C^*, with velocity constants k_a, k_b, and k_3:

$$C + C \underset{k_b}{\overset{k_a}{\rightleftarrows}} C + C^* \tag{13.33a}$$

$$C^* \xrightarrow{k_3} \text{Damaged Molecules} \tag{13.33b}$$

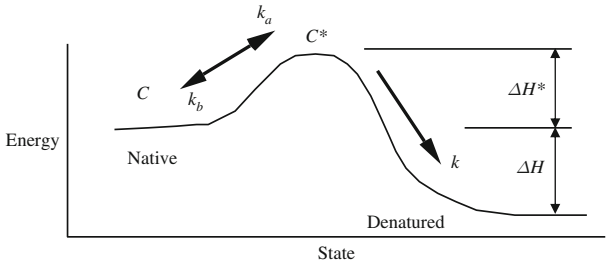


Fig. 13.15 Energy-state diagram for unimolecular reaction with energy barrier ΔH^*. Transition to denatured (damaged) tissue constituents occurs at overall reaction speed k

For this process, the rate of disappearance of native state molecules, $[C]$, is given by:

$$-\frac{d[C]}{dt} = k_3 \left[C^*\right] \tag{13.34}$$

where the bracket is used to indicate molar concentration. Generally $[C^*]$ is neither known nor calculable; however, at sufficiently low concentrations of C^* the steady state principle asserts that for short-lived activated states the rate of formation can be considered equal to the rate of disappearance. So the activated state, $[C^*]$, forms at a rate $k_a[C]^2$, relaxes back to inactivated at rate $k_b[C][C^*]$, and denatures at rate $k_3[C^*]$. Consequently:

$$k_a [C]^2 = k_b [C]\left[C^*\right] + k_3\left[C^*\right] \tag{13.35a}$$

$$\left[C^*\right] = \frac{k_a [C]^2}{k_b [C] + k_3} \tag{13.35b}$$

Of course, we seek an overall reaction velocity, k, which relates $[C]$ to its rate of disappearance:

$$-\frac{d[C]}{dt} = k [C] \tag{13.36}$$

There are two limiting cases for Eq. (13.35b). First, the concentration of remaining undamaged material, $[C]$, may be large enough that deactivation at k_b dominates the k_3 pathway, so $[C^*] \cong [C]k_a/k_b$ for which the overall formation rate, $k = k_3(k_a/k_b)$ and a first order process results. Second, if the remaining undamaged material concentration, $[C]$, is small, $k_3 >> k_b[C]$ and the process is second order since from Eq. (13.35a) $k \cong k_a/k_3[C]$. In liquid phase systems with appreciable concentrations of native state molecules the first condition should apply, so the first order approximation applies. After a long time of exposure at damaging temperatures such that $[C]$ is very small, $k_3 >> k_b[C]$ and a second order process results:

$$-\frac{d[C]}{dt} = k [C]^2 \tag{13.37}$$

where for simplicity the $[C]$ dependence has been removed from k: $k \cong k_a/k_3$. At such low concentrations the damage process is saturated, so we may ignore that case for the present.

Equation (13.36), then, is a Bernoulli differential equation with the solution:

$$C(\tau) = C(0)\, e^{\left\{-\int_0^\tau k\, dt\right\}} \tag{13.38}$$

Equations (13.30) and (13.31) may be used to relate k to ΔH^* and ΔS^*. It should be noted at this point that the energy barrier, E_a, (Fig. 13.14) is in fact $\Delta H^* + RT$

(Eq. 13.32); however, in practice $\{\Delta H^* \cong 5 \times 10^5\} >> \{RT \cong 3 \times 10^3\}$, so little error results from assuming that $E_a \cong \Delta H^*$. This approximation may be used to obtain:

$$k = \left(\frac{RT}{Nh_P}\right) \quad e^{\left[\frac{\Delta S^*}{R}+1\right]} \quad e^{\left[\frac{-E_a}{RT}\right]} \cong \left(\frac{RT}{Nh_P}\right) \quad e^{\left[\frac{\Delta S^*}{R}\right]} \quad e^{\left[\frac{-\Delta H^*}{RT}\right]} \tag{13.39a}$$

$$k \cong A\, e^{\left[\frac{-\Delta H^*}{RT}\right]} \tag{13.39b}$$

The term in parentheses in Eq. (13.39a) suggests that the pre-exponential factor, A, is not constant but is in fact temperature dependent. However, the linear dependence of A on temperature is extremely weak – especially when compared to the exponential dependence in the final term – and its effect is negligible, so that for all practical purposes A may be treated as a constant over the temperature ranges of interest.

Arrhenius Formulations in Thermal Damage Studies.

A more useful form of Eq. (13.27) may be obtained by recasting the result into a volume fraction model. In this formulation, as above, C signifies the remaining concentration of native state (undamaged) tissue constituent molecules. Therefore, the physical significance of the traditional damage measure, Ω, is the logarithm of the ratio of the original concentration of native tissue to the remaining native state tissue at time τ:

$$\Omega(\tau) = \ln\left\{\frac{C(0)}{C(\tau)}\right\} = \int_0^\tau A\, e^{\left[\frac{-E_a}{R\,T(t)}\right]}\, dt \tag{13.40}$$

where the frequency factor, A, and energy barrier, E_a, are related to the activation enthalpy and entropy, ΔH^* and ΔS^*, by Eq. (13.39a).

This form of the damage integral has the advantage that it is more easily compared to quantitative pathologic endpoints such as birefringence loss, collagen damage, or cell survival in culture. Using this description, direct comparisons can be made between computer models of transient thermal fields, $T(x,y,z,t)$, integrated over time and the measured histologic damage. A set of coefficients, A and $E_a \cong \Delta H^*$, is required for each damage process considered in the computer model. For tissue damage processes studied to date, A varies from about 10^{40} to 10^{105} [s^{-1}] while E_a ranges from about 1×10^5 to 9×10^5 [$J \cdot mole^{-1}$]. Each damage process is then allowed to progress in parallel, driven by the calculated thermal field. This formulation assumes that the individual processes are thermodynamically independent, a reasonable description for identifiable thermal damage processes. The concentration of each of the damage markers, $C(\tau)_i$, is accumulated; and a distributed field description of the predicted histologic endpoint is generated – for example, the boundary of thermally induced birefringence loss can be predicted. The model predictions are then suitable for comparison to histologic results.

Functional Behavior of the Damage Model

The characteristic behavior of the kinetic damage model is that below a threshold temperature the rate of damage accumulation is negligible, and it increases precipitously when this value is exceeded. This behavior is to be expected from the exponential nature of the function. For purposes of discussion, it is useful to define the critical temperature, T_{crit}, as the temperature at which the damage accumulation rate, $d\Omega/dt$, is 1:

$$\frac{d\Omega}{dt} = 1 = A\, e^{\left[\frac{E_a}{R\,T_{\text{crit}}}\right]} \tag{13.41}$$

so,

$$T_{\text{crit}} = \frac{E_a}{R\ \ln\{A\}}$$

For a process with (arbitrary) representative coefficients of $A = 1.0 \times 10^{75}$ and $E_a = 5 \times 10^5$, the critical temperature is 74.8°C. Figure 13.16 illustrates the damage accumulation rate dependence on temperature for the hypothetical example process. Note that the plot is converted from [K] to [°C] for the figure. In Fig. 13.17 the damage accumulation rate is plotted vs. 1/T [K^{-1}] for comparison.

Constant temperature exposures of the example process will result in a decrease in concentration of native state material depending on the time of exposure. A constant temperature exposure reduces the integral of Eq. (13.40) to a simple multiplication. Figure 13.18 shows the remaining concentration for this hypothetical example damage process for constant temperature exposures of time $\tau = 0.1$, 1.0, and 10 s. The concentration may be seen to gradually decrease with increasing temperature for fixed exposure times, as expected; the strong exponential nature of the process is evident as well. Applying this model framework relies heavily on identifying independent damage processes that can be quantitatively measured.

From plots like Fig. 13.18, estimates of A and E_a may be made from the $\Omega = 1$ line (i.e. $C(\tau) = 36.8\%$). Each curve will give one point on an Arrhenius plot ($\ln\{\tau\}$ vs. $1/T$) where $\Omega = 1$:

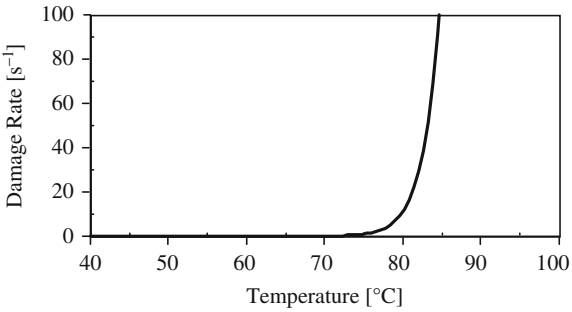


Fig. 13.16 Damage accumulation rate, $d\Omega/dt$, vs. temperature for the example process; $A = 1.0 \times 10^{75}$ and $E_a = 5 \times 10^5$

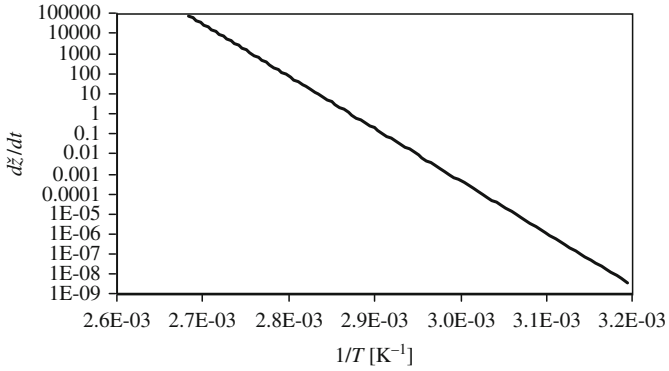


Fig. 13.17 Damage accumulation rate, $d\Omega/dt$, vs. $1/T$ [K^{-1}] for the hypothetical example process; $A = 1.0 \times 10^{75}$ and $E_a = 5 \times 10^5$

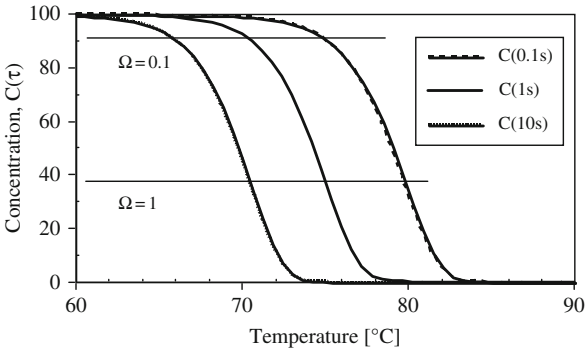


Fig. 13.18 Remaining undamaged tissue, $C(\tau)$, vs. T [°C] for the hypothetical example process; $A = 1.0 \times 10^{75}$ and $E_a = 5 \times 10^5$ for exposure times of 0.1, 1 and 10 s. Lines for $\Omega = 0.1$ and $\Omega = 1$ are also shown. At $\Omega = 10C(\tau) = 4.5 \times 10^{-3}$ and is not resolvable on the plot

$$\ln\{\tau\} = \left(\frac{E_a}{R}\right)\frac{1}{T} - \ln\{A\} \tag{13.42}$$

Estimates of A and E_a may then be obtained from a least-squares regression fit to the line of Eq. (13.42). Note the necessity of using exposure times that span several orders of magnitude in order to separate the curves sufficiently to give acceptable accuracy in the determination of the kinetic coefficients.

Though the damage parameter, Ω, cannot be measured directly in histologic section, often a clearer picture of the functional behavior can be obtained from it. The exponential dependence of damage on the inverse of absolute temperature means that the temperature required to obtain comparable damage levels is quite sensitive to time of exposure. For example, we may define a threshold temperature, T_{TH}, as the temperature at which the damage parameter, Ω, is 1 for a given duration, τ. Rearranging Eq. (13.42) gives the threshold temperature:

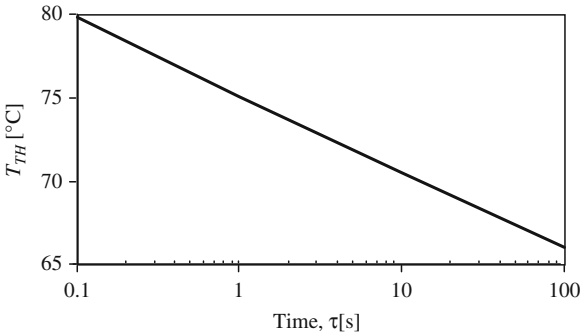


Fig. 13.19 Threshold temperature T_{TH} [°C] for the hypothetical example process; $A = 1.0 \times 10^{75}$ and $E_a = 5 \times 10^5$ vs. τ

$$T_{TH} = \frac{E_a}{R\ [\ln\{\tau\} + \ln\{A\}]} \tag{13.43}$$

Figure 13.19 is a plot of the threshold temperature as a function of duration for the example damage process of Figs. 13.16 and 13.18. Note that threshold temperature is approximately exponentially dependent on duration, as expected, with (in this case) a slope of about −4.65°C/decade. So, while 79.5°C is sufficient to result in $\Omega = 1$ (i.e. 63.2% damage) at 0.1 s, 102.3°C would be required for a 1 μs exposure for this process.

13.3.3.2 Experimental Determination of Rate Process Coefficients

Thermal damage kinetic coefficients are usually determined from constant temperature exposures of relatively long duration. Threshold damage results are selected out of a set of damaged tissue samples for analysis from which estimates of A and E_a are obtained.

Constant Temperature Exposures

Because of the sensitivity of the damage integral to small changes in temperature, the typical approach in obtaining A and E_a is to expose the tissue to a constant temperature, identify experiments in which the damage, $\Omega = \ln\{C(0)/C(\tau)\}$ is known and obtain A from the intercept and E_a from the slope of an Arrhenius plot of $\ln(\tau)$ vs. $1/T$ for the those experiments. If the temperature is held constant at T (K) Eq. (13.40) becomes:

$$\ln\{\tau\} = \left(\frac{E_a}{R}\right)\frac{1}{T} - \ln\{A\} + \ln\{\Omega\} \tag{13.44}$$

For experiments with $\Omega = 1$, $\ln\{\Omega\} = 0$ and a linear regression fit to the data plotted on $\ln\{\tau\}$ vs. $1/T$ axes gives the desired coefficients: intercept $b = -\ln\{A\}$ and slope $m = (E_a/R)$.

For constant temperature experiments in which Ω is known but not equal to 1, a simple re-arrangement of Eq. (13.44) yields the values required for an Arrhenius plot. In Eq. (13.45) the time of exposure for an equivalent experiment that would have $\Omega = 1$ at the experiment temperature T, i.e. τ_1, is simply related to the actual experiment conditions – Ω, T and τ – by:

$$\ln\{\tau\} - \ln\{\Omega\} = \ln\{\tau_1\} = \left(\frac{E_a}{R}\right)\frac{1}{T} - \ln\{A\} \tag{13.45}$$

In practice, one then plots $\ln\{\tau_1\}$ vs. $1/T$ (K^{-1}) for all useful data points and fits a linear regression line to determine the process coefficients from the slope and intercept.

The usual experimental method is to expose thin slices of tissue to constant temperature in a water bath, by surface application of heated water [72, 77] or on a heated metallic plate for desired time intervals. Table 13.4 lists rate coefficients obtained in various experiments. When using the kinetic models and measured coefficients, it is imperative that an adequate description of the particular end point be given. This is because different end points in the same tissue will have widely varying critical temperatures and damage accumulation rates. For example, in Table 13.4 in addition to the variation in tissues, the thermal damage histologic end points in the various studies differ.

Retina

The measurements of Welch and Polhamus [78] used the diameter of the edge of the visible lesion formed in rhesus monkey retina in vivo under argon laser irradiation as the end point. The temperatures were not measured directly, but were determined in separate experiments on retinas in which a correlation between temperature and radius was established using micro-thermocouples (about 5 μm in diameter) advanced from the posterior surface of the eye to a point just below the retina. The correlates were used to estimate the retinal temperature given laser beam power for durations between 0.1 and 10 s. The critical temperature for these coefficients is 56.0°C.

Takata et al. [79] have used a similar decision criterion for retinal damage for shorter exposure times. They fit the data with three sets of coefficients because a single first order model was not sufficient. It would appear that several parallel damage processes were at work in their study, thus a set of coefficients was required – the usual case is that there will be "breakpoints" in the Arrhenius plots, and the regions require different coefficients to fit the data. The critical temperature for the high temperature set of their coefficients is 59.9°C.

Birngruber et al. [80, 81] estimated a frequency factor and activation energy from consideration of the thermodynamics of protein and enzyme denaturation processes. Their estimates have a critical temperature of 74.5°C.

Skin

The end point for $\Omega = 1$ in the original studies of Henriques and Moritz corresponded to a continuum of cascaded effects [72, 74]. In their study the skin of pigs

Table 13.4 Experimentally determined Arrhenius rate coefficients

Source/process	Damage process coefficients			Notes
	A (s^{-1})	E_a (J/mole)	T_{crit} (°C)	
Heat Shock Proteins				
Beckham [86]	6.90×10^{282}	1.74×10^{6}	48.2	
Cell Death				
Sapareto [85]	2.84×10^{99}	6.18×10^{5}	51.4	CHO Cells, ≥ 43 °C
Skin				
Henriques [75]	3.1×10^{98}	6.28×10^{5}	59.9	Not recommended
Diller [82, 87]	1.3×10^{95}	6.04×10^{5}	58.5	Recommended (same data)
Weaver [83]	2.185×10^{124}	7.82×10^{5}	55.4	$T \leq 50$°C
	1.823×10^{51}	3.27×10^{5}	60.1	$T > 50$°C
Wu [88]	3.1×10^{98}	6.27×10^{5}	59.4	$T \leq 53$°C
	3.1×10^{98}	$6.27 \times 10^{5} - 5.1 \times 10^{5}$ $(T-53)$		$T > 53$°C
Fuggitt [89]	3.1×10^{98}	6.28×10^{5}	59.9	$T \leq 55$°C
	5.0×10^{45}	2.96×10^{5}	65.2	$T > 55$°C
Takata [90]	4.322×10^{98}	4.18×10^{5}	64.6	$T \leq 50$°C
	9.389×10^{104}	6.69×10^{5}	59.7	$T > 50$°C
Retinal Damage				
Welch [78]	3.1×10^{99}	6.28×10^{5}	57.6	Damage
Takata [79]	4.322×10^{64}	4.18×10^{5}	64.6	$T \leq 50$°C (Coagulation)
	9.389×10^{104}	6.69×10^{5}	59.7	$T > 50$°C
Vassiliadis [91]	9.95×10^{43}	2.90×10^{5}	71.1	Coagulation
Birngruber [80, 81]	1×10^{44}	2.93×10^{5}	74.7	
Collagen Changes				
Miles [92]	1.60×10^{137}	8.59×10^{5}	53.9	In lens capsule
Jacques [93]	7.35×10^{64}	4.251×10^{5}	69.2	Contraction of mouse dermis
Maitland [39]	1.77×10^{56}	3.676×10^{5}	68.2	Rat tail birefringence loss
Pearce [77]	1.606×10^{45}	3.06×10^{5}	80.4	Rat skin birefringence loss
Muscle				
Gaylor [94] See also: Toner [95] or Lee [96]	2.9×10^{37}	2.4×10^{5}	61.5	Cell membrane rupture
Jacques [97]	2.94×10^{39}	2.596×10^{5}	70.4	Myocardium whitening ~ birefringence loss

Table 13.4 (continued)

Source/process	Damage process coefficients			Notes
	A (s^{-1})	E_a (J/mole)	T $_{crit}$ (°C)	
Agah [98]	3.0×10^{23}	1.62×10^5	87.3	Optical properties changes
Erythrocytes				
Moussa [51]	6.8×10^{36}	2.49×10^5	80	Membrane denaturation
Lepock [52]	7.6×10^{66}	4.55×10^5	82.2	Hemoglobin coagulation
Flock [99]	1×10^{31}	2.12×10^5	84	Membrane denaturation
Egg				
Yang [84]	3.8×10^{57}	3.85×10^5	76	Egg albumin, whitening
Yang [84]	3.05×10^{56}	3.89×10^5	86.6	Egg yolk, whitening
Liver				
Jacques [100]	5.51×10^{41}	2.769×10^5	73.4	Whitening, pig liver
Mathewson [101]	2.09×10^{33}	2.219×10^5	74.7	Necrosis, rat liver
Prostate				
Jacques [102]	2.08×10^{27}	1.866×10^5	83.6	Whitening
Skinner [103]	3.8×10^{57}	3.85×10^5	76	Whitening, absorption coeff.
Kidney				
He [104]	1.48×10^{60}	3.996×10^5	73.7	Whitening, pig kidney
	3.3×10^{38}	2.569×10^5	75.2	Delayed necrosis
Pop [105]	5.73×10^{34}	2.404×10^5	88	Electrical conductivity
	5.85×10^{28}	2.023×10^5	94.2	Electrical permittivity

Coefficients are arranged by damage or tissue type and in the order of Increasing T_{crit}. References are listed by first author.

was exposed, in vivo, to flowing water at a controlled temperature for exposure times varying over several orders of magnitude. They calibrated their coefficients so that $\Omega = 0.5$ corresponded to the onset of erythema (characterized as "first degree" in their paper). Then $\Omega = 1$ corresponded to a "second degree," or a partial thickness, burn, and $\Omega = 10^4$ to a full thickness, or "third degree," burn. Their published coefficients, $A = 3.1 \times 10^{98}$ and $E_a = 6.28 \times 10^5$, have a critical temperature of 59.7°C. Interestingly, although these coefficients have been used for many years by numerous investigators, they do not fit the original data very well. In a later analysis, Diller and Klutke applied linear regression to their original data for temperatures less than 52°C – for which $A = 1.3 \times 10^{95}$ and $E_a = 6.04 \times 10^5$ – and if the data point at

52°C is included, $A = 0.865 \times 10^{95}$ and $E_a = 6.03 \times 10^5$, virtually the same result [82]. In any event, these new coefficients should be used in future work as they fit the original data much better than those originally published in 1947.

Weaver and Stoll [83] used similar criteria to the original 1947 studies and applied two sets of coefficients, as in Table 13.4 (the upper values are applicable above 50 °C) to match the experimental data. The critical temperature for their highest temperature coefficient set is 59.4°C.

Egg White and Egg Yolk

Yang et al. [84] exposed egg white and egg yolk to constant temperature (± 0.2°C) in a water bath for varying exposure times. Coagulation was defined as the onset of whiteness (coagulum formation due, apparently, to an increase in scattering in the clear liquid egg white) observed by the naked eye. Water bath temperatures ranged from 60 to 90°C in 4°C increments. At each temperature the time to threshold (onset of observable change) was measured and plotted in accordance with Eq. (13.44) and rate coefficients determined from linear regression, as described. Approximately 3–5 s were required to obtain whitening in the egg white at 70°C; while in the egg yolk 82°C was required at the same exposure time. The Table 13.4 values have a critical temperatures of 76.0°C for egg white and 86.6°C for egg yolk.

Birefringence Loss in Collagen

Birefringence loss in rat skin collagen was measured at temperatures between 40 and 90°C in 5°C increments for times ranging from 120 to 6000 s [77]. The gross extracted skin was wrapped in aluminum foil and immersed in a water bath. The heated samples were sectioned diagonally to increase their area, mounted on glass slides and stained (H&E) for analysis (see Fig. 13.9). The relative intensity of birefringence was determined using the exposure meter on the microscope camera in the "spot" mode: assuming that reciprocity applies, the intensity is the inverse of the indicated exposure time. This approach was necessary because the birefringence image is very low intensity: normal film exposure times required range up to 20 s in ASA 64 film. Intensity was calculated from data normalized by the background light intensity for each microscope slide, I_0. Relative birefringence intensity was then calculated by subtracting the fully denatured birefringence intensity, I_d, from the specimen intensity, I_s, and native state (undamaged) intensity, I_n, with all intensities individually normalized by I_0 before the calculation:

$$B = \frac{I_s - I_d}{I_n - I_d} \tag{13.45b}$$

A typical set of histologic sections from these data was presented in Fig. 13.5. The $\Omega = 1$ values were determined from the plot similar to Fig. 13.18 in order to calculate A and E_a. For the estimated $\Omega = 1$ points in this data set, $\ln\{\tau\} = 36,753\,(1/T) - 104$ with $r^2 = 0.900$, so the calculated values are: $A = 1.606 \times 10^{45}(s^{-1})$ and $E_a = 3.06 \times 10^5(\mathrm{J \cdot mole^{-1}})$, and $T_{\mathrm{crit}} = 80°\mathrm{C}$.

Deriving Damage Coefficients from Survival Curves

Another format typically used to present thermal damage data is the cell survival curve, derived from cell culture experiments. The simplest format in which to calculate A and E_a is a plot of surviving fraction *vs*. time at a fixed temperature. In that case the surviving fraction is simply $C(\tau)$, and a straight forward calculation can be done for $\Omega(\tau)$, from which A and E_a can quickly be extracted by curve fitting.

Alternately, the time required for the population to decrease by a factor of e^{-1} may also be used; see Fig. 13.20 [85]. D_0 is the time for the colony formation rate to decrease by e^{-1}:

$$\frac{S}{N_0} = e^{\frac{t_0 - t}{D_0}} \tag{13.46}$$

where $N_0 =$ the initial cell count measured at time t_0, and $S =$ surviving cell count at time t.

This parameter is plotted for asynchronously dividing Chinese Hamster Ovary (CHO) cells, and for CHO cells in the G1 phase in Fig. 13.20. The G1 phase is cell

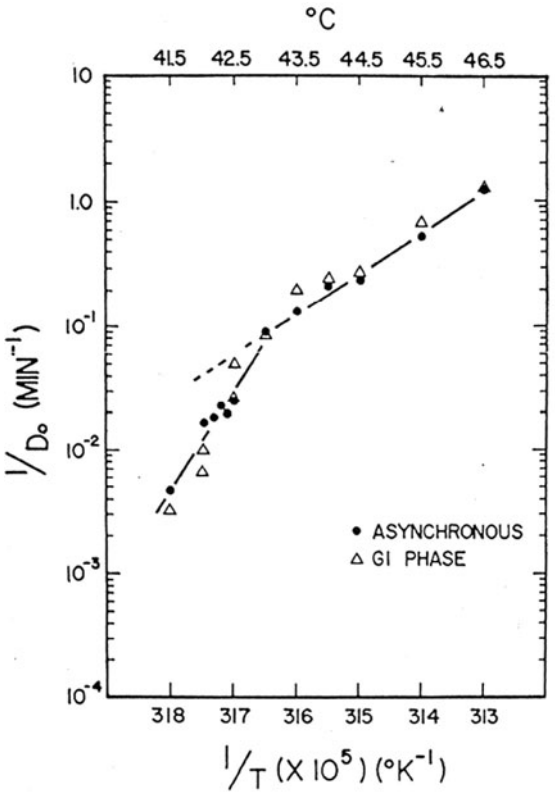


Fig. 13.20 Cell survival curve for Chinese Hamster Ovary (CHO) cells, a standard tumor cell line (reproduced from [85] and used by permission). The plot is $1/D_0$ [minutes^{-1}] vs. $1/T$ [K^{-1}]. Note the breakpoint at 43°C

growth immediately after a new cell is produced. The asynchronous cells have a random distribution of growth phases. CHO cells are common model cell lines used in tumor hyperthermia studies. By Eq. (13.46) the relationship for D_0 reduces to:

$$\ln\left\{\frac{C(0)}{C(\tau)}\right\} = 1 = \int_0^{D_0} A\, e^{\left[\frac{-E}{RT}\right]}\, dt \tag{13.47}$$

from which it may quickly be determined that:

$$\ln\{D_0\} = \frac{E}{RT} \; - \; \ln\{A\} \tag{13.48}$$

and the Arrhenius plot may be constructed as previously described. For the asynchronous CHO cells in this data set at temperatures above the break point (43°C) $A = 2.84 \times 10^{99}(\mathrm{s}^{-1})$ and $E_a = 6.18 \times 10^5[\mathrm{J \cdot mole}^{-1}]$, and $\mathrm{T_{crit}} = 51.4°\mathrm{C}$.

The usual damage prediction in hyperthermia studies is to express the exposure in terms of "Cumulative Equivalent Minutes" of exposure at the breakpoint temperature, typically very near 43°C in most tissues. The goal in those studies is to ensure cell death in the tumor, so the threshold for "success" is usually taken to be CEM = 30–60 min, depending on the application:

$$CEM\,43 = \sum_{i=1}^{N} R^{(43-T_i)}\, t_i \tag{13.49}$$

where t_i = the time of exposure [minutes] at temperature T_i [°C] and R is derived from a plot like Fig. 13.20. For most tissues above the breakpoint $R \simeq 0.5$, and can be calculated from:

$$\ln\{R\} = \frac{\ln\left\{\frac{D_0(T_{\mathrm{break}})}{D_0(T)}\right\}}{(T_{\mathrm{break}} - T)} \tag{13.50}$$

Example 13.7: This example derives Arrhenius coefficients from the CHO cell survival curve of Fig. 13.20 for temperatures below the break point, 43°C.

Solution:

(1) Reading lower temperature values for synchronous CHO cells from the plot:

T (°C)	T(K)	$D_0{}^{-1}$ (min^{-1})	D_0(s)	Ln{D_0}
43	316.2	0.09	667	6.503
42.5	315.7	2.6×10^{-2}	2308	7.744
42	315.2	1.6×10^{-2}	3750	8.230
41.7	314.9	4.8×10^{-3}	12,500	9.433

(2) Linear regression fit, $\ln\{D_0\}$ vs. T^{-1}

$$\ln\{D_0\} = 206,035\left[\frac{1}{T}\right] - 645.1$$

(3) $\ln\{A\} = 645.1$ means that $A = 1.46 \times 10^{280}$ (is this a frightening number, or what?)
(4) $E/R = 206,035$ means that $E = 1.71 \times 10^6[\text{J} \cdot \text{mole}^{-1}]$.

In practice, Eq. (13.49) is the discrete form of a continuous integral for time-varying temperatures. It should be noted that the *CEM43* damage calculation yields only an equivalent time of exposure and not a damage probability result. A *CEM43* value can be converted into a probability calculation if $D_0(43)$ is known for the tissue, however, by application of Eq. (13.46).

In summary, Fig. 13.21 compares several of the representative processes selected from Table 13.4 in terms of their relative rates of damage accumulation vs. temperature.

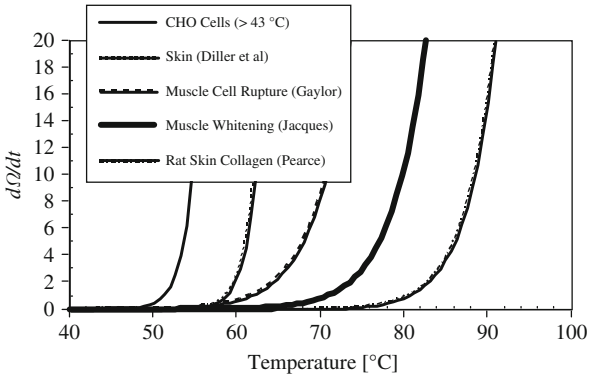


Fig. 13.21 Comparison of five representative rates of damage accumulation, $d\Omega/dt$, based on coefficients selected from Table 13.4, arranged in order of increasing critical temperature

Transient Thermal History Experiments

Problems with the standard constant temperature approach to determining A and E_a are that: (1) with only very few exceptions, the tissue must be excised to perform the experiments, disrupting its normal blood perfusion and activating autolytic (post mortem degeneration and necrosis) processes; (2) no exposure is truly constant temperature, and rise time segments must necessarily be short compared to the total exposure time in order to be negligible; and (3) the methods are essentially impossible to use in laser damage studies since constant temperature exposures are not possible to obtain and the typical small spot sizes mean that substantial thermal gradients confound reliable determination of the temperature history. Working with transient thermal data adds a high level of uncertainty to the damage coefficients

since the time of exposure is problematic. Nevertheless, estimates of rate coefficients can be made upon which treatment protocols can be evaluated.

Here we present a new method specifically applicable to transient temperature heating cases [106]. While it does not address the small-spot, steep thermal gradient problem, the new method does solve the transient temperature exposure problem so that shorter exposures may be used in a data set, thus reducing the uncertainty in the estimated damage coefficients. We can rearrange Eq. (13.44) for $\Omega = 1$ experiments to see the basis for the method:

$$\ln\{A\} = \left(\frac{1}{RT}\right) E_a - \ln\{\tau\} \tag{13.51}$$

So, for a single constant temperature experiment resulting in $\Omega = 1$, we can plot a straight line in the $\ln\{A\}$ vs. E_a plane, with slope $1/RT$ and intercept $\ln\{\tau\}$. An ensemble of such lines from several experiments would intersect at the value of $\ln\{A\}$ and E_a that describe the thermal damage process under study. This exercise provides no new information for constant temperature experiments, of course.

However, such a plot does provide a convenient basis for analyzing transient heating experiments. For any experiment in which $C(t)$ or $\Omega(t)$ can be directly measured, determining A and E_a is a simple matter of curve fitting. Given an ensemble of experiments for which only the histologic endpoint $\Omega(\tau)$ and the transient history, $T(t)$, are known, we can assume values for $\ln\{A\}$ and E_a and calculate a relative "Cost" that is minimized when the integral of Eq. (13.40) is evaluated to calculate $\Omega\,(A, E_a)$:

$$\text{Cost} = \ln\left\{\frac{\int_0^\tau A\, e^{\left[\frac{-E_a}{RT(t)}\right]}\, dt}{\Omega\,(\tau)}\right\} \tag{13.52}$$

This is a standard optimization technique (Cost has no units). Here the "Cost" has been defined such that its magnitude approaches zero when the assumed values of $\ln\{A\}$ and E_a yield an integral equal to the measured $\Omega(\tau)$ value.

A two-dimensional gray-scale plot of the magnitude of the Cost on the $\ln\{A\}$ vs. E_a plane using hundreds of assumed values of A and E_a shows that the locus of points for which the |Cost| approaches zero is a straight line (see Fig. 13.22). In Fig. 13.22a, a hypothetical low-energy damage process with $A = 26.5 \times 10^9 [\text{s}^{-1}]$ and $E_a = 7 \times 10^4\,[\text{J} \cdot \text{mole}^{-1}]$ has been hypothetically heated in transient manner – i.e. with discrete temperatures shown in the plot calculated from a $t^{0.5}$ formula – including 2 s of cooling. The resulting damage value is $\Omega = 0.9$. In Fig. 13.22b the $\ln\{A\}$ vs. E_a plane is scanned for minima in the Cost, as calculated by Eq. (13.52) from the transient curve. In this figure the clarity of the functional behavior is enhanced by plotting the magnitude of $(\log_{10}\{1/\text{Cost}\})$. Other transient heating experiments will yield an ensemble of plots such as Fig. (13.22b), each with differing slope and intercept.

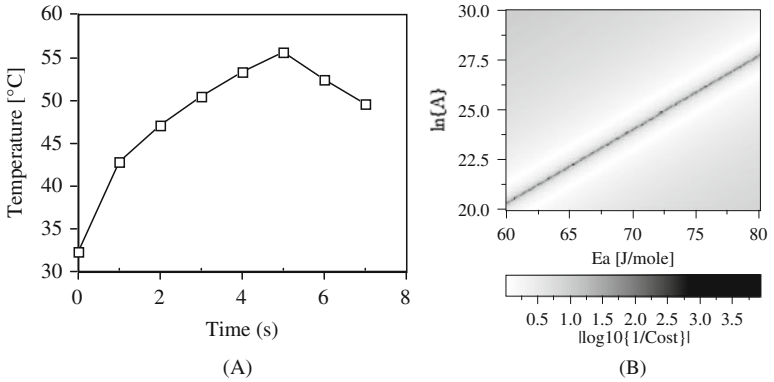


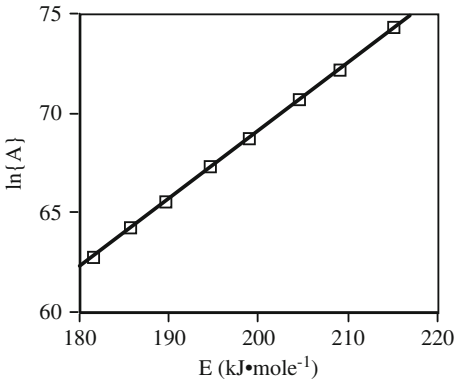
Fig. 13.22 Transient thermal history analyzed in the $\ln\{A\}$ vs. E_a plane. (**a**) Transient heating for a hypothetical process with $A = 26.5 \times 10^9$ and $E_a = 7 \times 10^4$ (linear in $t^{0.5}$) results in $\Omega = 0.9$. (**b**) Assuming values for A and E_a the transient temperature is integrated according to Eq. (13.52) and the inverse of Cost plotted in the plane. The locus of minima in Cost follow a straight line, the slope of which gives T_{eq} and the intercept $\ln\{\tau_{eq}\}$ for an equivalent constant-temperature experiment

The slope and intercept of the line of Cost minima in Fig. 13.22b contain extremely useful information. Referring to Eq. (13.51), the slope is $(1/RT_{eq})$ and the intercept is $\ln\{\tau_{eq}\}$, where T_{eq} and τ_{eq} are, respectively, the temperature and duration of an equivalent constant-temperature experiment for which $\Omega = 1$. It is then a simple matter to plot the ensemble of (T_{eq}, τ_{eq}) points on a standard Arrhenius plot, from which A and E_a are obtained in the ordinary fashion, i.e. from the slope and intercept of the linear regression line through them. In practice, the appropriate range to scan in the $\ln\{A\}$ vs. E_a plane can quickly determined by a coarse discrete scan of the values, followed by a fine scan to refine the estimate of T_{eq} and τ_{eq}.

This approach has the advantage over other curve fitting approaches that it preserves the underlying physics and reveals processes and relationships. It turns out to be quite critical that accurate values for the temperature be used to calculate the integral of Eq. (13.52). Consequently, an experiment design in which the temperature is uniform over a relatively large area is required so that spatial averaging can be used to reduce uncertainty in both the temperature and the quantitative damage measure. This excludes almost all laser spot approaches to damage process estimation. Note that it is neither necessary nor advantageous to have $\Omega = 1$ to make an Arrhenius plot when using this approach

Example 13.8 This example calculates the equivalent constant temperature experiment for the process described in Fig. 13.22. Here the hypothetical process has been integrated for the transient heating curve that resulted in $\Omega = 3.78$ for the hypothetical thermal damage process. The search plane was estimated using the strategy outlined in the text. The line of convergence in Fig. 13.22b has these points:

Cost	ln{A}	E (kJ/mole)
3.33×10^{-3}	62.75	181.4
4.05×10^{-3}	64.25	185.7
1.28×10^{-3}	65.58	189.6
5.23×10^{-3}	67.33	194.6
5.04×10^{-3}	68.75	198.8
3.70×10^{-3}	70.75	204.5
3.93×10^{-3}	72.25	208.9
2.73×10^{-3}	74.33	215.0



The fit for this line is: linear regression

$$\ln\{A\} = 0.345E\left(\text{kJ}\cdot\text{mole}^{-1}\right) + 0.201$$

So,

$$\frac{1}{R\,T_{eq}} = 0.345\times10^{-3}$$

$$\ln\left\{\tau_{eq}\right\} = -0.201$$

Or, $T_{eq} = 75.5°\text{C}$ and $\tau_{eq} = 0.818\,\text{s}$. For the original damage process ($A = 1 \times 10^{75}, E = 5 \times 10^5$) a constant temperature exposure of 0.818 s at 75.5°C results in $\Omega = 0.677$; close to $\Omega = 1$. Note that τ_{eq} is much shorter than the original experiment time of 5 s because the resulting damage in the simulated experiment gave Ω larger than 1. For experiments that have $\Omega < 1$, τ_{eq} will be longer than the experiment heating time. The difference between $\Omega = 0.677$ and the expected value of 1 is most likely due to an accumulation of round off errors: note how small 0.201 is compared to $\ln\{A\} = 172.7$ in Eq. (13.44), for example.

13.4 Summary

First order kinetic models for tissue damage are useful for predicting trends in laser damage experiments. Unfortunately, there are only a few damage processes for which the frequency factor and energy have been determined in situ – critical temperatures for known processes range from about 50 to 90°C. Nevertheless, these models can be used to make direct comparisons between numerical predictions of laser damage and histologic results, something that cannot, as yet, be achieved any other way. A very careful definition of the particular histologic end point is necessary. Also, to avoid the arbitrariness that characterizes many of the early damage studies, it is essential that a quantitatively measurable damage measure be identified. Birefringence intensity is one such variable; other excellent candidates include, for example, changes in optical properties such as the scattering or absorption coefficients on exposure to elevated temperatures.

In part because of the effects of biologic inhomogeneities and random fluctuations in tissue characteristics, and in part due to the difficulty of resolving small temperature differences, thermal damage data are inherently noisy. However, predicted boundaries in small spot size experiments compare favorably to those observed. This is probably due to the very steep thermal gradients typical of small spot experiments – that is, a rather large error in actual critical temperature may be swamped out by the very steep thermal gradient so that the location of the predicted damage contour may, in fact, agree fairly well with experimental histologic observation. Certainly, even though the results of a particular experiment may eventually prove impossible to duplicate in numerical models, a careful analysis of the trends which one would obtain from, say, changing beam power, spot size and duration can be studied in detail in the numerical model and on a spatial scale similar to that of microscopic observation. Also, the numerical model allows dissection of the transient development of thermal damage – something that cannot be achieved in any other way. So, while there are many uncertainties associated with kinetic models of thermal damage, they can be extremely illuminating and helpful in dosimetry planning.

References

1. Thomsen S. Pathologic analysis of photothermal and photomechanical effects of laser-tissue interactions. *Photochem. Photobiolog.*, 53(6):825–835 (1991).
2. Thomsen S. Identification of lethal thermal injury at the time of photothermal treatment. In: eds. G. Muller and A. Roggan (eds) *Intersitial thermotherapy*. SPIE, Bellingham, WA, pp. 459–467 (1995).
3. Thomsen S. Qualitative and quantitative pathology of clinically relevant thermal lesions. *SPIE Crit. Rev*., CR75:425–459 (2000).
4. Thomsen S. Non-thermal effects in thermal treatment applications of non-ionizing irradiation. *Proc. SPIE*., 5698:1–14 (2005).
5. Thomsen S and Coad JE. Developing clinically successful biomedical devices by understanding the pathophysiology of the target tissue: Insights from over 25 years at the microscope. *Proc. SPIE*., 6440:1–15 (2007).

6. Kumar,V, Abbas AK, Fausto N, and Mitchell RN. *Robbins basic pathology*. Saunders-Elsevier, Philadephia, 8th edition, pp. 1–224 (2007).
7. Lee H-C and Wei Y-H. Mitochondrial role in life and death of the cell. *J. Biomed. Sci*., 7:2–15 (2000).
8. Letal A and Scorrano L. Laying the foundations of programmed cell death. *Cell Death Differ.*, 13:1245–1247 (2006).
9. Chan D, Frank S, and Rojo M. Mitochondrial dynamics in cell life and death. *Cell Death and Differ.*, 13:680–684 (2006).
10. Coad JE. Thermal fixation: A central outcome of hyperthermic therapies. *Proc. SPIE*, 5698:15–22 (2005).
11. Metz CN. Fibrocytes: A unique cell population implicated in wound healing. *Cell. Mol. Life Sci.*, 60:1342–1350(2003).
12. Baum CL and Arpey CJ. Normal cutaneous wound healing: Clinical correlation with cellular and molecular events. *Dermatol. Surg*., 31:674–686 (2005).
13. Blakytny R and Jude E. The molecular biology of chronic wounds and delayed healing in diabetes. *Diabetic Med*., 23:594–608 (2006).
14. Williams P, Bannister LH, Berry MM et al. (eds). *Gray's anatomy. 38th international edition*. Churchill-Livingstone. Edinburg, pp. 1–2092 (1995).
15. Fawcett DW. *Textbook of histology*. Chapman and Hall, New York, 12th edition, pp. 1–964 (1994).
16. Ghadiallay FN. *Ultrastructural pathology of the cell and matrix*, Vols. 1 and 2. Butterworth-Heinemann, Boston, pp. 1–1450 (1997).
17. Alberts B, Johnson A, Lewis J et al. *Molecular biology of the cell.* Garland Science, New York, 4th edition, pp. 1–1463 (2002).
18. Junquiera LC and Carneiro J. *Basic histology: Text and atlas*. McGraw-Hill, Medical Publishing Division, New York, 11th edition, pp. 1–502 (2005).
19. Thomsen S. Practical pathology for engineers: How to do the job right the first time. *Proc. SPIE*, 4754:1–26 (2003).
20. Welch AJ and van Gemert MJC. (eds). *Optical-thermal response of laser-irradiated tissue.* Plenum, New York, pp. 1–925 (1995).
21. Niemz MH. *Laser-tissue interactions: Fundamentals and applications*. Springer, Berlin, 3rd edition (2004).
22. Thomsen S, Pearce JA, Randeri R, and Chan E. Determination of isotherms of thermal damage. *Proc. IEEE/LEOS*, 2295–2296 (1995).
23. Thomsen S. Mapping thermal injury in biologic tissues using quantitative pathologic techniques. *Proc. SPIE*., 3595:82–95 (1999).
24. LeCarpentier GL, Motamedi M, Rastegar S, and Welch AJ. Simultaneous analysis of thermal and mechanical events during cw laser ablation of biological media. *Proc. SPIE*, 1064: 107–113 (1989).
25. Verdaasdonk RM, Borst C, and van Gemert MJC. Explosive onset of continuous wave laser tissue ablation. *Phys. Med. Biol.*, 35:1129–1144 (1990).
26. Gijsbeers GHM, Selten FM, and van Gemert MJC. CW laser ablation velocities as a function of absorption in an experimental one dimentional model. *Lasers Surg Med.,* 11 :287–296 (1991).
27. Pearce JA and Thomsen S. Rate process analysis of thermal damage. In: AJ Welch and MJD van Gemert (eds) *Optical-thermal response of laser irradiated tissue*. Plenum, New York, pp. 561–606 (1995).
28. Protsenko DE and Pearce JA. Thermo-electrical numerical model of electrosurgical RF cutting. *Proc. SPIE*, 4247:203–209 (2001).
29. Nelson DL and M.M. Cox MM. *Lehninger principles of biochemistry*. W.H. Freeman and Co, New York, 4th edition, pp. 1–1119 (2005).
30. Guyton AG and Hall JE. *Textbook of medical physiology*. Saunders Elsevier, Philadelphia, 11th edition, pp. 1–1116 (2006).

31. Bray D. *Cell movements: From molecules to motility.* Garland, New York, 2nd edition, pp. 1–372 (2001).
32. Boal D. *Mechanics of the cell.* University Press. Cambridge, UK, pp. 1–406 (2002).
33. Ramachandran GN and Ramakrishan C. Molecular structure. In: Ramachandran GN and Reddi AH (eds) *Biochemistry of Collagen*. Plenum, New York. pp. 45–84 (1976).
34. Fleischmajer R, Olsen BR, and Kühn K. Biology, chemistry and pathology of collagen. *Annals NY Acad. Sci.*, 460:1–537 (1985) and 580:1–592 (1990).
35. Uitto J and Perejda AJ. *Connective tissue disease: molecular pathology of the extracellular matrix*. Marcel Decker, New York, pp. 1–563 (1987).
36. Chadwick DJ and Goode JA (eds). *The molecular biology and pathology of elastic tissues. Ciba symposium 192.* Wiley, New York, pp. 1–351 (1995).
37. Flory P and Garrett RR. Phase transitions in collagen and gelatin systems. *J. Am. Chem. Soc.*, 80:4836–4845 (1958).
38. Lim JJ. Transition temperature and enthalpy change dependence on stabilizing and destabilizing ion in the helix coil transition in native tendon collagen. *Biopolymers* 15:2371–2381 (1976).
39. Maitland DJ and Walsh JT Jr. Quantitative measurements of linear birefringence during heating of native collagen. *Lasers Surg. Med*., 20:310–318 (1997).
40. Yoshimura H, Viator JA, and Jacques SL. Relationship between damaged fraction and reflected spectra of denaturing tissues. *Lasers Surg. Med*., 37:308–313 (2005).
41. Schmidt SJ. Die Doppelbrechung von karyoplasma, zytoplasma und metaplasma. In: *Protoplasma-Monographien. Vol. II.* Verlag von Gebruder Borntraeger, Berlin, pp. 154–267 (1937).
42. Fisher E. The birefringence of striated and smooth muscles. *J. Cell. Comp. Physiol*., 23:11–130 (1944).
43. Peckham M and Irvin M. Myosin crossbridge orientation in demembranated muscle fibres studied by birefringence and X-ray diffraction measurements. *J. Mol. Biol*., 210:113–126 (1989).
44. Harris P and Heath D. Structure and function of vascular smooth muscle. In: *The human pulmonary circulation: Its form and function in health and disease.* Churchill Livingstone, New York, pp. 161–182 (1986).
45. Hulmes DJS, Miller A, Parry DAD et al. Crystalline regions in collagen fibrils. *J. Mol. Biol.*, 184:473–477 (1985).
46. Thomsen S, Pearce JA, and Cheong W-F. Changes in birefringence as markers of thermal damage in tissues. *IEEE Trans. Biomed. Eng.*, BME-36:1174–1179 (1989).
47. Thomsen S. Quantitative morphologic markers of laser thermal injury in cardiovascular muscle. *Proc. SPIE*, 1878:152–158 (1993).
48. Thomsen S, Jacques SL, and Flock S: Microscopic correlates of macroscopic optical property changes during thermal coagulation of myocardium. *Proc. SPIE*, 1202:2–11 (1990).
49. Thomsen S, Vijverberg H, Huang R, Schwartz J. Changes in optical properties of rat skin during thermal coagulation. *Proc. SPIE*, 1882:230–236 (1993).
50. Fielding CJ (ed). *Lipid Rafts and Caveolae: From membrane biophysics to cell biology*. Wiley-VCH, New York, pp. 1–294 (2006).
51. Moussa NA, Tell EN, and Cravalho EG. Time progression of hemolysis of erythrocyte populations exposed to supraphysiological temperatures. *Trans. ASME J. Biomech. Eng*., 101:176–184 (1979).
52. Lepock JR, Frey HE, Bayne H, Markus J. Relationship of hyperthermia-induced hemolysis of human erythrocytes to the thermal denaturation of membrane proteins. *Biochim. Biophys. Acta*, 980:191–201 (1989).
53. Bernardi, P, Scorrano L, Colonna R, Petronelli V, and Di Lisa F. Mitochondria and cell death: Mechanistic aspects and methodological issues. *Eur. J. Biochem*., 264:687–701 (1999).

54. Van Cruchten, S and Van den Brobeck W. Morphological and biochemical aspects of apoptosis, oncosis and necrosis. *Anat. Histol. Embryol*., 31:214–223 (2002).
55. Shoshan-Barmatz V, Israelson A, Brdiczda D, and Sheu SS. The voltage-dependent anion channel (VDAC): Function in intracellular signaling, cell life and cell death. *Curr. Pharm. Des*., 12:2349–2270 (2006).
56. Kim J-S, He L, Qian T, and Lemasters JJ. Role of the mitochondrial permeability transition in apoptotoc and necrotic death after ischemia/reperfusion injury o hepatocytes. *Curr. Mol. Med*., 3:527–235 (2003).
57. Anghieri LJ and Robert J. *Hyperthermia in cancer treatment.* CRC, Boca Raton (1986).
58. van der Zee J. Heating the patient: A promising approach? *Ann. Oncol*., 13:1173–1184 (2002).
59. Beynon, R and Bond JS. *Proteolytic enzymes.* University Press. Oxford, 2nd edition, pp. 1–340 (2001).
60. Thomsen S, Schwartz JA, Joseph R, Pearce JA, Rae B, and McMurray TJ. Temperatures associated with thermally induced red blood cell changes in tissues irradiated in vivo. *Proc. SPIE*, 2130:156–163 (1994).
61. Milleron RS and Bratton SB. "Heated" debates in apoptosis. *Cell. Mol. Life Sci*., 64:2329–2333 (2007).
62. Cohen IK, Diegelmann RF, and Lindblad WJ. *Wound healing: Biochemical and clinical aspects*. W.B. Saunders, Philadelphia, pp. 1–630 (1992).
63. Falanga V. *Cutaneous wound healing.* Martin Dinitz, London, pp. 1–484 (2001).
64. DiPietro LA and Burns AL. *Wound healing; Methods and protocols.* Humana, Totowa, NJ, pp. 1–453 (2003).
65. Grey JE and Harding KG. *ABC of wound healing.* Blackwell, Malden, MA, pp. 1–47 (2006).
66. Keenan JH and Keyes FG. *Thermodynamic properties of steam*, 1st edition, 19th printing. Wiley, New York (1948).
67. Van Wylen GJ. *Thermodynamics.* Wiley, New York (1964).
68. Reynolds WC. *Thermodynamics*. McGraw-Hill, New York (1970).
69. Torres JH, Motamedi M, Pearce JA, and Welch AJ. Experimental evaluation of mathematical models for predicting the thermal response of tissue to laser irradiation. *Appl. Opt.*, 32(4):597–606 (1993).
70. Brutsaert W. *Evaporation into the atmosphere: Theory, history and applications.* D. Reidel, Dordrecht, Holland (1982).
71. Yang D, Converse MC, Mahvi DM, and Webster JG. Expanding the bioheat equation to include internal water evaporation during heating. *IEEE Trans. Biomed. Eng*., 54(8):1382–11388 (2007).
72. Moritz AR and Henriques FC. Studies of thermal injury II. The relative importance of time and surface temperature in the causation of cutaneous burns. *Am. J. Pathol*., 23:695–720 (1947).
73. Moritz AR. Studies of thermal injury III. The pathology and pathogenesis of cutaneous burns: An experimental study. *Am. J. Pathol*., 23: 915–934 (1947).
74. Henriques FC and Moritz AR. Studies of thermal injury in the conduction of heat to and through skin and the temperatures attained therein: A theoretical and experimental investigation. *Am. J. Path*., 23:531–549 (1947).
75. Henriques FC. Studies of thermal injury V. The predictability and significance of thermally induced rate processes leading to irreversible epidermal injury. *Arch. Pathol*., 43:489–502 (1947).
76. Maron SH and Lando JB. *Fundamentals of physical chemistry*. McMillan, New York (1974).
77. Pearce JA, Thomsen S, Vijverberg H, and McMurray T. Kinetics for birefringence changes in thermally coagulated rat skin collagen. *Proc. SPIE*, 1876:180–186 (1993).
78. Welch AJ and Polhamus GD. Measurement and prediction of thermal injury in the retina of Rhesus monkey. *IEEE Trans. Biomed. Eng*., 31:633–644 (1984).

79. Takata AN et al. *Thermal model of laser induced eye damage*. Final Report, USAF School of Aerospace Medicine, Brooks AFB TX, Contract F41609-74-C-0005, IIT Research Institute, Chicago, IL (1974).
80. Birngruber R. Thermal modeling in biological tissue. In: F Hillenkamp, R Pratesi, and CA Sacchi (eds) *Lasers in biology and medicine,* Plenum, New York, pp. 77–97 (1980).
81. Birngruber R, Hillenkamp F, and Gabel VP. Theoretical investigations of laser thermal retinal injury. *Health Phys*., 48(6):781–796 (1985).
82. Diller KR, Pearce JA, and Valvano JW. Bioheat transfer, Chapter 4, Section 4. In: *CRC Handb. Therm. Eng.*, (cat. no. 9581), pp. 4–114, 4–187 (2000).
83. Weaver JA, and Stoll AM. Mathematical model of skin exposed to thermal radiation. *Aerospace Med*., 40(1):24 (1969).
84. Yang Y, Welch AJ, and Rylander HG. Rate process parameters of albumen. *Lasers Surg. Med*., 11:188–190 (1991).
85. Sapareto SA. *Physical aspects of hyperthermia*. G Nussbaum (ed). American. Institute of Physics, Chapter 1 (1982).
86. Beckham JT, Mackanos MA, Crooke C, Takahashi T, O'Connell-Rodwell C, Contag CH, and Jansen ED. Assessment of cellular response to thermal laser injury through bioluminescence imaging of heat shock protein 70. *Photochem. Photobiol.*, 79:76–85 (2004).
87. Diller KR and Klutke GA. *Accuracy analysis of the Henriques model for predicting burn injury*. Advances in Bioheat and Mass Transfer, AM. Soc. Mech. Eng., New York, 268 (1993).
88. Wu YC. *A modified criterion for predicting thermal injury*. Nat. Bur. Stand. Washington, District of Columbia (1982).
89. Fugitt CE. A rate process of thermal injury. Armed Forces Special Weapons Project No. AFSWP-606 (1955).
90. Takata AN. Development of criterion for skin burns. *Aerospace Med.*, 45:634–637 (1974).
91. VassiliadisA, Christian HC, and Dedrick KG. Ocular laser threshold investigations. *Aerospace Med. Rep*., F41609-70-0002 (1971).
92. Miles CA. Kinetics of collagen denaturation in mammalian lens capsules studied by differential scanning calorimetry. *Int. J. Biol. Macromol*., 15:265–271 (1993).
93. Jacques SL and Prahl SA. Modeling optical and thermal distributions in tissue during laser irradiation. *Lasers Surg. Med*., 6:494–503 (1987).
94. Gaylor DC. *Physical mechanism of cellular injury in electrical trauma*. Ph.D. Thesis, Department of Electrical Engineering, Massachusetts Institute of Technology (1989).
95. Toner M, Cravalho EG, Gaylor DC, and Lee RC. Cellular mechanisms of thermal injury in electrical trauma. *Proc. Ann. Conf. Eng. Med. Biol.*, 4:1505–1506 (1990).
96. Lee RC Gaylor DC, Prakah-Asante K, Bhatt D, and Israel DA. Skeletal muscle cell rupture by pulsed electric fields. *IEEE Eng. Med. Biol. Soc. Annu. Conf*., 712–714 (1987).
97. Jacques SL and Gaeeni MO. Thermally induced changes in optical properties of heart. *IEEE Eng. Med. Biol. Mag.* 11(Part 4/6):1199–1200 (1989).
98. Agah R, Gandjbakhche AH, Motamedi M, Nossal R, and Bonner RF. Dynamics of temperature dependent optical properties of tissue: dependence on thermally induced alteration. *IEEE Trans. Biomed. Eng.*, 43(8):839–846 (1996).
99. Flock ST, Smith L, and Waner MD. Quantifying the effects on blood of irradiation with four different vascular-lesion lasers. *Proc. SPIE*, 1882: 237–242 (1993).
100. Jacques SL, Newman C, and He XY. Thermal coagulation of tissues: Liver studies indicate a distribution of rate parameters not a single rate parameter describes the coagulation process. *Proc. Winter Annu. Meeting Am. Soc. Mech. Eng.* (1991).
101. Matthewson K, Coleridge-Smith P, O'Sullivan JP, Northfield TC, and Bown SG. Biological effects of intrahepatic neodymium:yttrium-aluminum-garnet laser photocoagulation in rats. *Gastroenterology*, 93:550–557 (1987).
102. Jacques SL, Motamedi M, and Rastegar S. Computer simulation of laser coagulation of prostate: A guide to dosimetry. *Am Soc. Lasers Med. Surg. Lasers Surg. Med.*, Suppl 5, abstract 311 (1993).

103. Skinner MG, Everts S, Reid AD, Vitkin IA, Lilge L, and Sherar MD. Changes in optical properties of *ex vivo* rat prostate due to heating. *Phys. Med. Biol*., 45:1375–1386 (2000).
104. He X, McGee S, Coad JE, Schmidlin F, Iaizzo PA, Swanlund DJ, Kluge S, Rudie E, and Bischof JC. Investigation of the thermal and tissue injury behaviour in microwave thermal therapy using a porcine kidney model. *Int. J. Hyperthermia*, 20:567–593 (2004).
105. Pop M, Molckovsky A, Chin L, Kolios MC, Jewett MA, and Sherar MD. Changes in dielectric properties at 460 kHz of kidney and fat during heating: Importance for radio-frequency thermal therapy. *Phys. Med. Biol*., 48:2509–2525 (2003).
106. Pearce JA and Thomsen S. Arrhenius model thermal damage coefficients for birefringence loss in rabbit myocardium. *Proc. ASME Int'l. Mech. Eng. Conf. Expos*., IMECE2003-41986 (2003).

Chapter 14
Pulsed Laser Ablation of Soft Biological Tissues

Alfred Vogel and Vasan Venugopalan

14.1 Introduction

In this chapter we focus on the key elements that form our current understanding of the mechanisms of pulsed laser ablation of soft biological tissues. We present a conceptual framework providing mechanistic links between various ablation applications and the underlying thermodynamic and phase change processes [1]. We define pulsed laser ablation as the use of laser pulses with duration of ~1 ms or less for the incision or removal of tissue regardless of the photophysical or photochemical processes involved. However, we will confine this presentation to pulsed ablation performed on a tissue level that does not involve laser-induced plasma formation. Ablation processes within transparent tissues or cells resulting from nonlinear absorption have been considered in reviews by Vogel and Venugopalan [1] and by Vogel and co-workers [2].

14.2 Tissue Properties Relevant for Ablation

Although early consideration of laser ablation treated tissue with properties equivalent to water, it soon became apparent that *tissue composition*, as well as its optical, thermal, and mechanical properties, play important roles on both the ablation dynamics and outcome. Soft biological tissues consist of cells that reside in and attach to an extracellular matrix (ECM). By mass, the composition of most soft tissues is dominated by water (55–99%) and collagen (0–35%). In "cell-continuous" tissue such as liver and epithelia, the ECM fraction is quite small and consists mostly of cell adhesion proteins. By contrast, "matrix-continuous tissues" that include the corneal stroma, dermis, cartilage, and tendon have a very small cellular fraction and are almost entirely ECM. In matrix-continuous tissues (see Fig. 14.1), the ECM consists largely of collagen, with the collagen content being as high as 35% [1, 3–5].

A. Vogel (✉)
Institute of Biomedical Optics, University of Lübeck, Lübeck, Germany
e-mail: vogel@bmo.uni-luebeck.de

A.J. Welch, M.J.C. van Gemert (eds.), *Optical-Thermal Response of Laser-Irradiated Tissue*, 2nd ed., DOI 10.1007/978-90-481-8831-4_14,

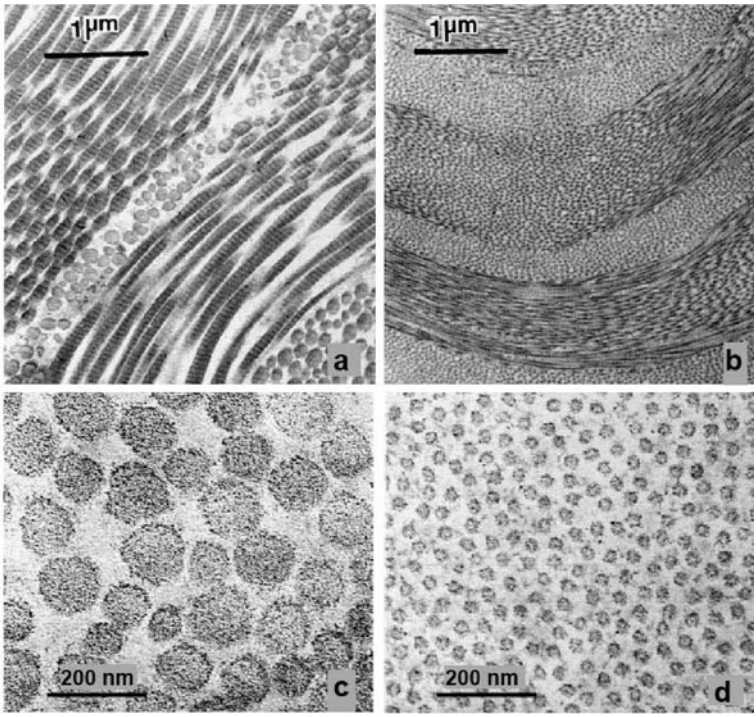


Fig. 14.1 Transmission electron micrographs of the sclera (**a**, **c**) and corneal stroma (**b**, **d**) showing an arrangement of collagen fibrils that are embedded in a ground substance with high water content. (Reprinted with permission from [6])

Nearly all the water content resides in the ground substance between the collagen fibrils [1, 7]. The primary ECM function is the maintenance of the tissue's structural integrity. As a result, the ECM inhibits both tissue vaporization and material removal that represent the desired outcomes of ablation processes.

Figure 14.2a summarizes the absorption spectra of proteins, melanin, hemoglobin, and water that govern the deposition of the incident laser energy in pulsed laser ablation. A more detailed discussion of the *optical absorption properties* of tissues is presented in Chapter 3 and Ref. [1]. In non-turbid samples, optical transmission is governed by Beer-Lambert's law, and the amount of deposited laser energy decays exponentially with depth. In the absence of scattering, the reciprocal of the absorption coefficient $\mu_a[\mathrm{cm}^{-1}]$ defines the optical penetration depth δ [cm], i.e. $\delta = (1/\mu_a)$, and thus the characteristic depth of laser energy deposition.

For most laser wavelengths used only a single tissue constituent (e.g., water or collagen) absorbs the radiation; thus macroscopic energy deposition is inhomogeneous. Heat diffusion tends to level out these inhomogeneities, and the extent to which this occurs during the ablation process depends both on the spatial scale that characterize the distribution of these constituents and the laser pulse duration

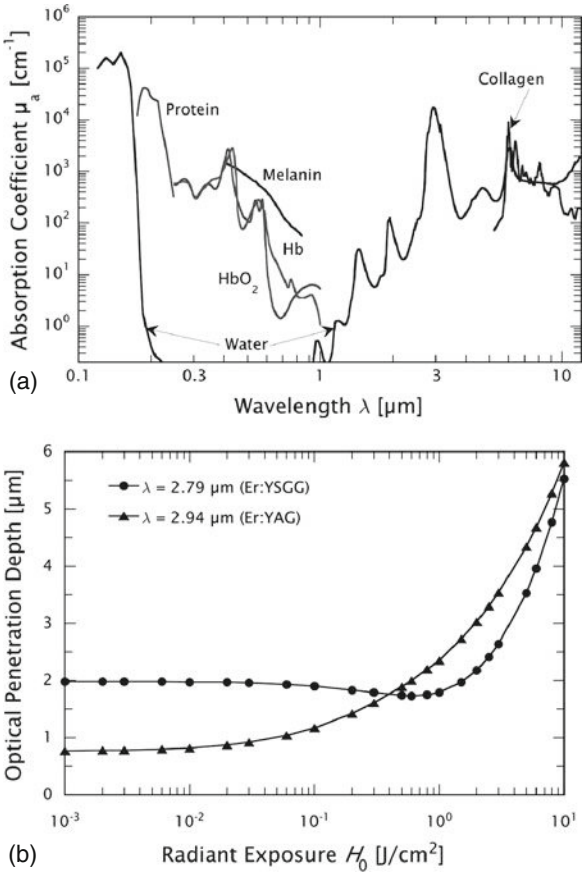


Fig. 14.2 Optical properties relevant for ablation. (**a**) Optical absorption coefficients of principal tissue chromophores in the 0.1–12 μm spectral region. (**b**) Variation of optical penetration depth of water with incident radiant exposure for Er:YAG ($\lambda = 2.94\,\mu$m) and Er:YSGG ($\lambda = 2.79\,\mu$m) laser irradiation. (Reprinted with permission from [1])

(see Section 14.3). An important spatial scale characterizing domains with different absorption properties is given by the diameter and spacing of the collagen fibrils because most tissue water resides in either cells or in the ground substance of the ECM between the collagen fibrils. The fibril diameter is ≈30 nm in cornea and varies between 20 nm and 120 nm in dermis. The center-to-center spacing is very regular (≈65 nm) in the transparent cornea and exhibits more variations in other tissues [1, 3, 7].

Optical scattering arises from spatial variations in refractive index within tissue that are particularly strong between collagen fibrils and ground substance. Typical reduced scattering coefficients for tissues in the visible spectral region are on the order of $\mu'_s = 10 - 40\,\mathrm{cm}^{-1}$ [1, 8]. Optical scattering will reduce the optical penetration depth, δ, of light relative to the absorption depth, $(1/\mu_a)$. In addition, when scattering is dominant over absorption, backscattering and total internal reflection lead to a fluence rate close to the tissue surface that can exceed by several times the incident irradiance [9]. Moreover, the peak fluence can be reached below the surface. However, ablation is typically performed at wavelengths where $\mu_a >> \mu'_s$.

Precise tissue ablation is achieved using laser wavelengths that exhibit very large absorption coefficients (Fig. 14.2a), such as the radiation of ArF excimer lasers ($\lambda = 193$ nm), Er:YSSG lasers ($\lambda = 2.79\,\mu$m), Er:YAG lasers ($\lambda = 2.94\,\mu$m) lasers ($\lambda = 10.6\,\mu$m). Since these wavelengths cannot be transmitted well through optical fibers, they are mainly used for ablation at tissue surfaces in air. For ablation inside the human body, XeCl excimer lasers ($\lambda = 308$ nm), thulium:YAG lasers ($\lambda = 2.0\,\mu$m) and holmium:YAG lasers ($\lambda = 2.1\,\mu$m) are frequently employed because their radiation is transmitted by low-OH quartz fibers.

The thermal and mechanical transients generated during the pulsed laser ablation process can result in *dynamic changes of the optical absorption properties*, as described in detail in Chapter 9. These changes are particularly important for wavelengths around the absorption peak of water at $\lambda = 2.94\,\mu$m that shifts towards shorter wavelengths for increasing temperature [10–12]. Figure 14.2b shows the variation in optical penetration depth with incident radiant exposure for $\lambda = 2.94$ and 2.79 μm [1]. For $H_0 > 0.5$ J/cm^2, Er:YSGG laser radiation ($\lambda = 2.79\,\mu$m) offers better spatial confinement of the laser energy compared to Er:YAG laser radiation ($\lambda = 2.94\,\mu$m), opposite to the behavior one would expect from the absorption coefficients measured at small radiant exposures (Chapters 3 and 9). Variations in optical absorption with temperature are also important for ultraviolet laser ablation. Absorption of UV radiation by peptide bonds is followed by heating of the surrounding water, resulting in a change in hydrogen bonding structure of water and thus to a red shift of the water absorption band, which at room temperature is located at 160 nm [13]. The absorption of water at $\lambda = 193$ nm may be raised to as much as $\mu_a \approx 10^4$ cm^{-1} at a volumetric energy density of $W = 2$ kJ/cm^3 [13].

Phase transitions of the tissue water are strongly influenced by the *mechanical tissue properties*. Figure 14.3 provides stress-strain curves for a variety of soft biological tissues. Although the mechanical characteristics vary considerably, these tissues all possess a nonlinear stress-strain characteristic with a "concave-up" shape. As a result, biological tissues are soft and elastic under physiologic conditions but become very stiff when loaded in an extreme fashion as in laser ablation.

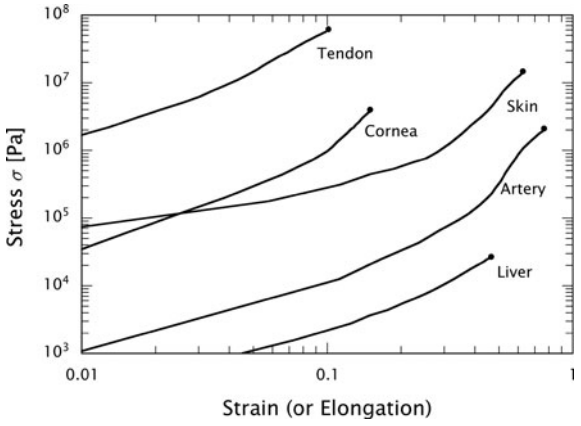


Fig. 14.3 Stress-strain curves characterizing the mechanical properties of various biological tissues under uniaxial tension. The · symbols represent the mechanical state at which tissue fracture occurs. Data are compiled from Ref. [14]

There is a positive correlation between tissue strength and collagen content. Tissues that represent extremes of mechanical strength are the liver and tendon. Liver is a cell-continuous tissue with little ECM and collagen content, which results in a very low ultimate tensile strength (UTS) of 23 kPa and moderate extensibility at fracture ~50% [3, 15]. Tendon is a matrix-continuous tissue that possesses high collagen content. This provides for high strength and stiffness with an UTS of $\gtrsim$100 MPa and fracture extensibility of ~10%. Skin has similarly high collagen content (25–33%) but lower UTS (~10 MPa) and much larger fracture extensibility of ~30–100% because the collagen fibrils in the dermis are "wavy" and form a loose three-dimensional network [3] (see Fig. 14.1).

Nearly all tissue mechanical data have been acquired under "quasi-static" loading conditions in which the tissue is deformed at very slow strain rates; typically on the order of 10^{-3} s^{-1}. However, the processes involved in pulsed laser ablation of tissue produce extremely high strain rates; on the order of 10^5–10^7 s^{-1}. Studies performed to examine the effect of strain rate in the range 0.3–170 s^{-1} have revealed that while the tissue strain at fracture remains roughly constant, the UTS increases in proportion to the logarithm of the strain rate [15–17]. The increase in UTS is due to the viscous dissipation between the collagen fibrils and the adjacent ground substance during the rapid deformation. It is not known whether the logarithmic dependence of tissue UTS vs. strain rate remains valid at the extreme rates produced by pulsed laser ablation. However, the available UTS data suggest that the tissue strength under ablative conditions can be considerably higher than that measured under "quasi-static" loading conditions.

Thermal denaturation of collagen fibrils affects the dynamics of the ablation process. Denaturation begins when the thermal energy of the constituent molecules overcomes the weak hydrogen bonds and van der Waals interactions stabilizing the helical configuration of the α-chains in the collagen molecule [18]. The "native" triple-helical structure of the molecule is thus transformed into a "denatured" random coil structure that is associated with a loss of the banding pattern of the native collagen fibrils in TEM (see Fig. 14.1) and with shrinkage of the fibrils along their longitudinal axis [4, 19]. However, when the fibrils are embedded in the surrounding tissue, shrinkage is impaired and tensile stress develops along the fibrils due to the covalent cross-links that connect the molecules and maintain the organization of the fibrils [20, 21]. Further heating denatures first the thermally-labile and then the thermally-stable covalent cross-links between the collagen molecules. This results in a stepwise disintegration of the collagen fibrils [22], a relaxation of the stresses developed during shrinkage [20, 21], and, finally, in total mechanical failure of the fibrillar tissue structure that now appears homogeneous in TEM [18, 19]. Older tissues possess a higher density of cross-links and thus require higher temperatures to undergo these transitions [4, 23].

To understand the degree to which collagen denaturation influences a particular ablation process, one must consider the kinetics of this rate process (see Chapter 13) that depends on both the magnitude and duration of thermal exposure [24]. For shorter heating times, considerably higher temperatures are required for denaturation. While the mechanical stability of collagen is destroyed at about 75°C when

heated for several minutes [20], temperatures far in excess of 100°C are required to affect mechanical stability for thermal exposures in the nanosecond to microsecond range characteristic of pulsed laser ablation [25]. Nevertheless, given that surface temperatures approaching 400–750°C have been measured during tissue ablation using laser pulses of 100 μs duration [26], the mechanical integrity of the tissue ECM will certainly be compromised. We shall see in Section 14.6 that the progressive denaturation of the tissue matrix with time is one reason for the continuation of the ablation process after the end of the laser pulse. In nanosecond ablation even moderate radiant exposures may produce temperatures that exceed 1000°C in the superficial tissue layer, at which point the mechanical integrity of the tissue ECM is completely lost due to thermal dissociation of the constituent molecules into volatile fragments [27] (Section 14.4.6).

14.3 Linear Thermo-Mechanical Response to Pulsed Irradiation

In the absence of photochemical processes, the laser energy absorbed by the tissue is completely converted to a *temperature rise* before a phase transition occurs. Under adiabatic conditions, the temperature rise T [K] at a location r is related to the local volumetric energy density $W(r)$ [J/m^3] by:

$$\Delta T(r) = \frac{W(r)}{\rho c_v} \tag{14.1}$$

where ρ[kg/m^3] is the tissue density and c_v[J·kg^{-1}·K^{-1}] the specific heat capacity at constant volume. The absorbed energy is redistributed by thermal diffusion [28] (see Chapter 10). In 1983, Anderson and Parrish introduced the concept that spatially-confined microsurgical effects can be achieved by using laser pulse durations t_p shorter than the characteristic thermal diffusion time of the heated volume [29]. For laser ablation, the heated volume is typically a layer with a thickness of the optical penetration depth ($1/\mu_a$), and the characteristic thermal diffusion time t_d is given as [30]

$$t_d = \frac{1}{\kappa \mu_a^2} \tag{14.2}$$

where κ [m^2/s] is the thermal diffusivity. By defining a dimensionless ratio $t_d^* = (t_p/t_d)$, the thermal confinement condition can be expressed as [31, 32]

$$t_d^* = \frac{t_p}{t_d} = \kappa \mu_a^2 t_p \lesssim 1 \tag{14.3}$$

Short-pulse laser irradiation of tissue not only leads to rapid heating but also to the *generation and propagation of thermoelastic stresses* [33]. The magnitude and

temporal structure of these stresses are governed by the longitudinal speed of sound in the medium c_a, the laser pulse duration t_p, the depth of the heated volume $(1/\mu_a)$ and the Grüneisen coefficient Γ [33, 34]. The dimensionless Grüneisen coefficient Γ is simply the internal stress per unit energy density generated when depositing energy into a target under constant volume (i.e., isochoric) conditions. This is given by the thermodynamic derivative

$$\Gamma = \left(\frac{\partial \sigma}{\partial W}\right)_v = \frac{\beta}{\rho c_v \kappa_T} \tag{14.4}$$

where σ [Pa] is the internal stress, W the volumetric energy density, v the specific volume, β [1/K] the coefficient of thermal expansion, ρ the mass density, c_v the specific heat capacity at constant volume and κ_T [1/Pa] the isothermal compressibility.

Thermoelastic stresses are most prominent when the laser pulse duration t_p is smaller than the characteristic time required for a stress wave to traverse the heated volume $t_m = (1/\mu_a c_a)$ [30]. This means that the stresses are confined within the heated region during the laser irradiation. By defining a dimensionless ratio $t_m^* = (t_p/t_m)$, the "stress confinement" condition can be expressed as [31, 32]

$$t_m^* = \frac{t_p}{t_m} = \mu_a c_a t_p \lesssim 1 \tag{14.5}$$

For $t_m^* << 1$, heating of the laser-affected volume occurs under isochoric conditions and the thermoelastic stress is maximal. The peak stress σ_p is given by [33]

$$\sigma_p = A\Gamma\varepsilon_0 = A\Gamma\mu_a\Phi_0 \tag{14.6}$$

where $A = 1$ and the duration of the thermoelastic stress transient t_a scales with the stress propagation time and $t_a \approx (4 - 6/\mu_a c_a)$. When the stress transient leaves the heated volume, the peak stress drops to $0.5\,\sigma_{\mathrm{p}}$.

In the limit $t_m^* \to \infty$, where there is no stress confinement, $A \to 0$ and the duration of the stress transient approaches that of the laser pulse. The variation of A with t_m^* for different temporal laser pulse shapes is shown in Fig. 14.4.

While thermal expansion of a heated volume generates compressive thermoelastic stresses, subsequent propagation of these stresses results in transients that contain both compressive and tensile components. Tensile stresses arise from the reflection of the compressive stress waves at an interface with a medium with lower acoustic impedance (tissue-air, tissue-water) or from the three-dimensional characteristics of acoustic wave propagation from a heated tissue volume of finite size [2, 34, 36–39]. Tensile stress wave generation originating from acoustic impedance mismatch at a tissue surface is shown in Fig. 14.5.

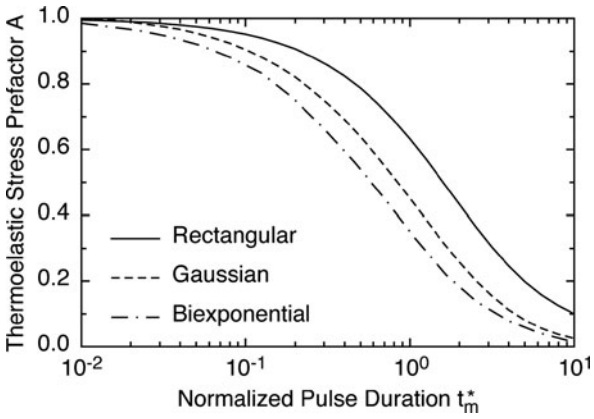


Fig. 14.4 Variation of the thermoelastic stress prefactor A (see Eq. (14.8)) with pulse duration relative to the stress propagation time across the heated volume t_m^* for various laser pulse shapes. Data are compiled from [33] and [35]

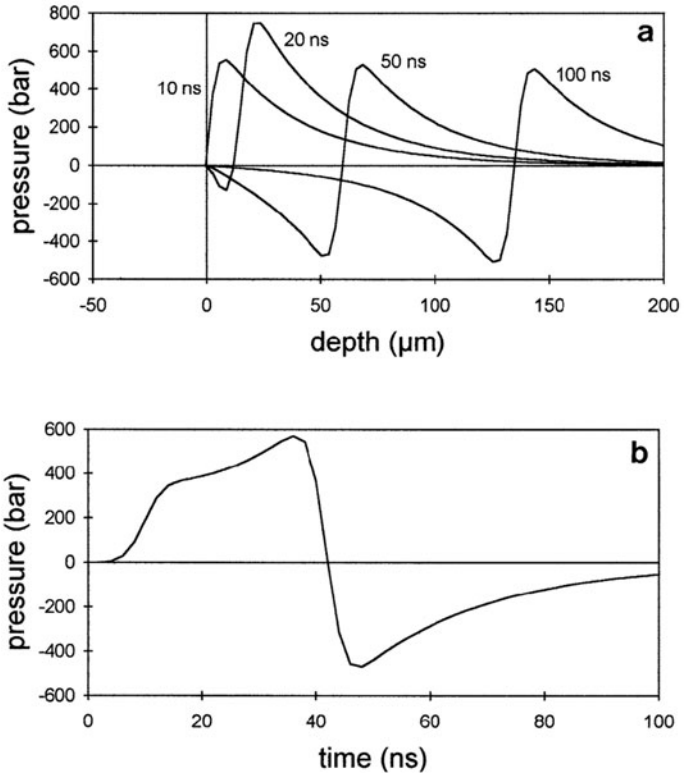


Fig. 14.5 (**a**) Development of a thermoelastic stress wave in water calculated for $H_0 = 2$ J/cm^2; $\mu_a = 200$ cm^{-1} and $t_p = 8$ ns. (**b**) Pressure as a function of time at a depth of 50 μm. (Reprinted with permission from [37])

14.4 Thermodynamics and Kinetics of Phase Transitions

All ablation processes involve the fracture of chemical bonds and lead to the removal of single molecules, molecular fragments, and molecular clusters. Bond fracture can also produce the formation of voids (i.e., bubbles or cracks) that facilitate the ejection of non-decomposed material fragments upon mechanical failure of the target. Vaporization, molecular fragmentation, and void formation can all be viewed as phase transitions that are accomplished via photothermal, photomechanical, and/or photochemical mechanisms. Given the central role of phase transitions in the ablation process, we devote this section to a systematic presentation and analysis of their thermodynamics and kinetics. We begin by considering the generation of phase transitions via photothermal and photomechanical phenomena and their modification in the presence of a tissue matrix. We then end the section with a treatment of photochemical mechanisms to achieve bond dissociation.

14.4.1 Phase Diagrams

We use the pressure vs. temperature projection of the phase diagram for liquid and gaseous water (Fig. 14.6) and the pressure vs. specific volume projection of the thermodynamic phase diagram (Fig. 14.7) to discuss the thermodynamics of phase transitions. The solid curve *A*–*C* on Fig. 14.6 represents those pressure/temperature pairs where liquid and vapor phases are in equilibrium with each other and is known as the "binodal." The curve *B*–*C*–*D*, the "spinodal," denotes a locus of states representing the intrinsic stability limit of the liquid or vapor phase [i.e., $(\partial T/\partial s)_p = 0$ and $(\partial p/\partial v)_T = 0$]. At the spinodal, the superheated liquid phase (*B*–*C*) or subcooled vapor (*C*–*D*) phase is no longer stable with respect to the random density

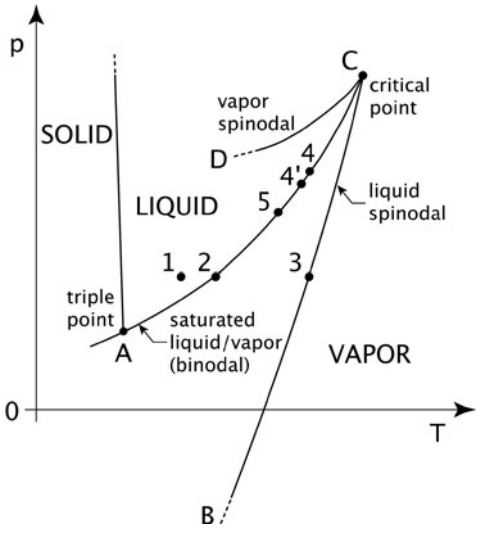


Fig. 14.6 Pressure vs. temperature projection of the thermodynamic phase diagram including the spinodal curve. Specific states of interest are (1) ambient temperature and pressure, (2) boiling temperature under ambient conditions, (3) spinodal temperature at ambient pressure, (4) saturated conditions corresponding to the ambient spinodal temperature. The importance of points 4′ and 5 are discussed in Section 14.4.4

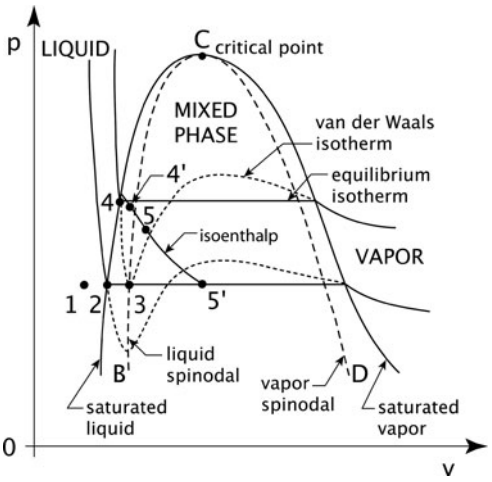


Fig. 14.7 Pressure vs. specific volume projection of the thermodynamic phase diagram including the spinodal curve and equilibrium and van der Waals isotherms. Points 1–4 correspond to those shown in Fig. 14.6. Importance of points 4′, 5, 5′ and the isoenthalp are discussed in Section 14.4.4

fluctuations that are present in all materials at nonzero temperatures. The region between segment *A*–*C* of the binodal and *B*–*C* of the spinodal represents metastable states of the superheated liquid for which the presence of a free surface or bubble nucleus is required for vaporization. The binodal and spinodal curves intersect at the critical point *C*, above which no thermodynamic distinction can be made between liquid and vapor phases. For water, the critical point is located at $T_c = 374.14°\text{C}$ and $p_c = 22.09$ MPa.

Liquid, vapor, and mixed phase regions are clearly demarcated in the *p*-*v* diagram of Fig. 14.7. The binodal encompasses the mixed phase region that specifies the range of specific volumes in which liquid and gaseous phases coexist for a given pressure and temperature. The dashed curve provides the spinodal where the segment *B*–*C* represents the stability limit of superheated liquid and segment *C*–*D* represents the stability limit of subcooled vapor.

14.4.2 Surface Vaporization

Equilibrium vaporization at a liquid-vapor interface is associated with an increase of specific volume at constant temperature that requires the latent heat of vaporization. This process can occur at any single location along the binodal (A–C) in the *p*-*T* diagram of Fig. 14.6, with the rate of vaporization increasing with *T*. In the *p*-*v* diagram in Fig. 14.7, surface vaporization is represented by a path following an equilibrium isotherm that connects a state of saturated liquid with a state of saturated vapor. Thus vaporization does not occur at a pre-determined temperature and theoretical models that adopt a fixed "vaporization temperature" violate the basic physics of the process [40]. The actual surface temperature is dictated by the rate of equilibrium vaporization that balances the irradiance supplied to the system. Thus,

surface vaporization becomes a "cooling" heat flux boundary condition in models describing the thermal distribution within the sample.

However, during laser ablation, rates of vapor formation in excess of that predicted by equilibrium vaporization are often achieved. This occurs because the increased equilibrium vapor pressure corresponding to the increased temperature of the liquid surface is not established instantaneously. This results in an increased mass flux of vapor transported into the surroundings is known as non-equilibrium interphase mass transfer and can be estimated using arguments from the kinetic theory of gases [41, 42]. Nevertheless, even this increased non-equilibrium interphase mass transfer provides a very small ablation rate. *Efficient* ablation is achieved at laser irradiances that deposit energy at rates that cannot be balanced by surface vaporization processes alone. This results in material removal via *volumetric* processes.

14.4.3 Normal Boiling

Normal boiling refers to a volumetric process that forms vapor at a thermodynamic state on the binodal as indicated by point 2 in Fig. 14.6. Thus, for a given pressure, the binodal defines the corresponding "boiling temperature." For water at atmospheric pressure this temperature is 100°C. However, for ablation processes with a high rate of mass removal, the boiling temperature is increased significantly because the recoil increases the pressure both at the target surface and within its bulk.

Vapor formation in normal boiling relies on the presence of pre-existing nuclei of vapor or dissolved gas within the liquid to catalyze the nucleation and growth of vapor bubbles. The transition from saturated liquid to saturated vapor occurs within a finite layer of mixed phase. The thickness of this "vapor-liquid" layer is comparable to the optical penetration depth of the incident radiation, and its composition varies from that of saturated liquid at its base to saturated vapor at the target surface [40, 43]. As a result, the surface temperature is fixed at the saturation conditions corresponding to the pressure at the target surface and there is no temperature gradient within the vapor-liquid layer.

Once a normal boiling process is established, temperatures slightly higher that the saturation temperature result in a growth of the vapor bubbles in the liquid. Therefore, normal boiling processes *always "involve" partial vaporization* of a liquid volume through the growth of vapor bubbles. Thus the concept sometimes found in biomedical ablation papers that vaporization only occurs once the entire latent heat of vaporization is deposited is not correct.

Nevertheless, normal boiling plays a negligible role for pulsed laser ablation for two reasons. First, the density of heterogeneous bubble nucleation sites is likely insufficient to provide a boiling process sufficiently vigorous to balance the high rates of energy deposition achieved in most pulsed laser ablation processes [32, 40]. Second, the high rates of energy deposition can be balanced only if the bubbles move to the target surface on a time scale set by the propagation velocity of the ablation front. Miotello and Kelly [44] showed that this is not possible when irradiating pure

water with nanosecond laser pulses and is possible for microsecond pulses only for radiant exposures proximal to the ablation threshold. In tissue this is even less likely because the mobility of vapor bubbles is further inhibited by the presence of the extracellular matrix.

14.4.4 Phase Explosion and Explosive Boiling

When the rate of volumetric energy deposition provided by laser radiation is more rapid than the rate of energy consumed by vaporization and normal boiling, the tissue water is driven to a metastable superheated state. The superheated liquid is metastable until the spinodal temperature is reached. The spinodal limit is defined by line *B*–*C* in Fig. 14.6 that represent the locus of states with infinite compressibility $[(\partial p/\partial v)_T = 0]$. At the spinodal limit, the superheated liquid undergoes "spinodal decomposition:" a spontaneous process by which a thermodynamically unstable liquid relaxes towards equilibrium by "phase separation" into a mixture of saturated vapor and saturated liquid [45–47]. The spinodal temperature of water at atmospheric pressure is ≈305°C with the corresponding equilibrium saturation vapor pressure of 9.2 MPa. Thus spinodal decomposition under atmospheric conditions involves an impressive pressure rise resulting in the violent emission of saturated liquid droplets by the expanding vapor.

For the phase diagram shown in Fig. 14.6, the heating phase corresponds to the path 1→3, and the spinodal decomposition will initially result in a nearly isochoric transition from point 3 on the spinodal to point 4′ in the mixed phase region possessing the same enthalpy. For pure water, the subsequent explosive expansion of this mixture will transition through a series of thermodynamic states that follows the curve of constant enthalpy (isoenthalp), as shown in Fig 14.7, until the mixture reaches atmospheric pessue at point 5′. During the expansion 4′→5′, the temperature of the mixture drops to 100°C, and about half of the liquid is transformed into vapor. The vapor fraction (∼49.6%) is given by energy density necessary to heat water from room temperature to the spinodal limit (1.27 kJ/g) as compared to the sum of the sensible and latent enthalpy of vaporization. The remaining saturated liquid is ejected in the form of droplets.

To provide a complete description of the phase transformation process as the liquid is heated to the spinodal limit, one must also consider the potential contribution of homogenous nucleation [32, 44, 48–50]. Homogenous nucleation refers to the spontaneous formation of vapor inclusions within the bulk liquid that arise solely from thermodynamic fluctuations and is not catalyzed by the presence of impurities or dissolved gas. While the formation of such vapor "nuclei" is spontaneous, their growth is not ensured and depends strongly on superheat temperature.

In classical nucleation theory, the driving force for growth of vapor nuclei is supplied by the difference in chemical potential between the superheated liquid outside the bubble and the vapor inside the bubble. This driving force is necessary to overcome the free energy barrier posed by the surface tension separating the vapor from the liquid [47]. The chemical potential difference between the superheated

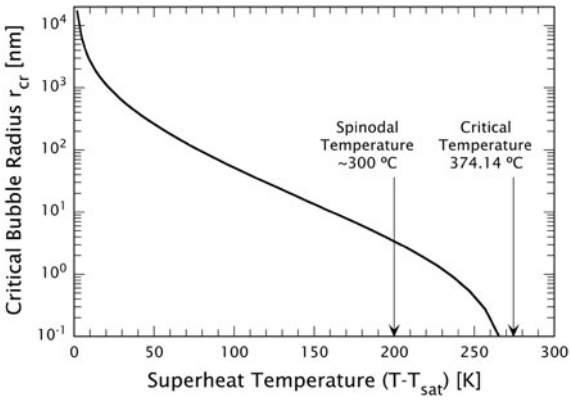


Fig. 14.8 Variation of the critical bubble radius r_{cr} required for spontaneous vapor bubble growth with superheat temperature. Note that the critical bubble radius goes to zero at the critical temperature

liquid and vapor scales with the bubble volume (i.e., r^3), while the contribution from surface tension scales with the bubble surface area (i.e., r^2). As a result, small vapor nuclei that form due to thermodynamic fluctuations spontaneously collapse while larger vapor nuclei will grow. The Gibbs free energy ΔG that describes the thermodynamics of bubble formation is given by:

$$\Delta G = \frac{4\pi r^3}{3}(\mu_v - \mu_l) + 4\pi r^2 \sigma \tag{14.7}$$

where μ_v and μ_l [J/m^3] are the chemical potentials of the vapor and liquid state, respectively, r is the radius of the vapor nuclei, and σ [N/m] is the surface tension of the surrounding liquid [47, 51]. Nuclei grow if they are larger than a critical radius r_{cr}. Figure 14.8 shows the dependence of r_{cr} on superheat temperature for water. Note that while r_{cr} strongly decreases as the superheat temperature increases, it remains finite even at the spinodal temperature. Thus nucleation remains an activated process with a finite free energy barrier [45]. The strong reduction of Δr_{cr} results in a dramatic rise in the nucleation rate J [s^{-1}] with the superheat temperature that attains a large, but finite, value at the spinodal temperature as shown in Fig. 14.9. Here we define nucleation as formation of a bubble with radius $r > r_{cr}$. The energy barrier that must be overcome for the conversion from the liquid to vapor phase disappears only when surface tension disappears and this occurs at the critical point. To account for the influence of statistical fluctuations on the stability limit, Kiselev has introduced the concept of a *kinetic spinodal*. The kinetic spinodal is defined as the locus of thermodynamic states where the time for spontaneous formation (driven by superheat temperature) of vapor nuclei becomes smaller than the characteristic time for their decay to local equilibrium (driven by surface tension) [52]. The superheat temperatures defined by the kinetic spinodal are much lower than the critical temperature and slightly lower than the classical spinodal. The kinetic spinodal represents the physical limit of metastable liquid states that can be achieved prior to spinodal decomposition.

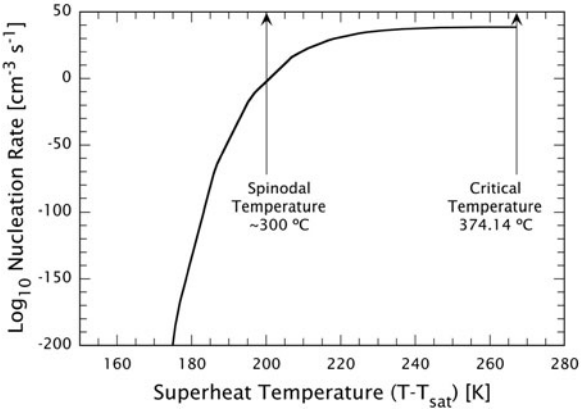


Fig. 14.9 Variation of vapor bubble nucleation rate with superheat temperature. Nucleation is here defined as formation of a bubble with radius $r > r_{cr}$ (see Fig. 14.8)

Thus, in general, the transformation of superheated (metastable) liquid to an equilibrium state of mixed phase may involve both bubble nucleation (large density fluctuations extending over a small spatial extent) and spinodal decomposition (small density fluctuations extending over a large spatial extent). We refer to the collective phase transition process as a *phase explosion*. A more detailed consideration of nucleation theory and spinodal decomposition as it relates phase transitions and tissue ablation can be found in our earlier review [1].

Thus far we have focused on processes tracing a path indicated by $1 \rightarrow 3 \rightarrow 4' \rightarrow 5 \rightarrow 5'$ in Figs. 14.6 and 14.7. This path corresponds to the extreme case in which no vapor nuclei are present in the liquid. When the heating occurs very rapidly at high radiant exposures, the liquid experiences a recoil pressure from surface vaporization/interphase mass transfer that can be substantial due to the nonequilibrium conditions produced during the beginning of the laser pulse. Thus spinodal conditions are reached at an elevated pressure somewhere between point 3 and the critical point *C*. Because the resulting phase explosion occurs at elevated temperature and pressure, the pressure jump associated with the phase separation is less severe. The elevated temperature corresponds to a higher volumetric energy density of the superheated liquid, and therefore more than half of the liquid will be transformed into vapor during the phase separation process.

When vapor nuclei are present in the liquid and the heating occurs on a time scale such that a significant fraction of the incident laser energy (but not the entire energy) contributes to the growth of heterogeneous and homogeneous vapor nuclei, the resulting phase transition process again follows a path that is intermediate between normal boiling and the path $1 \rightarrow 3 \rightarrow 4$. Spinodal conditions in the superheated liquid are again reached at a location between point 3 and the critical point *C* followed by phase separation. In this case however, the pressure rise is due to vapor formation at the nucleating centers rather than recoil from vapor leaving the target surface. Such intermediate processes are termed *explosive boiling*. In general, the energy necessary to reach spinodal conditions is higher for explosive boiling than for a phase explosion with surface vaporization. The amount of vapor formation is greater due to contributions from both the growing nucleation centers and the phase separation.

Both phase explosion and explosive boiling are volumetric processes in which a portion of the target material is ejected in the liquid phase and the latent heat of vaporization is not supplied to the entire ablated mass. As a result, the ablation efficiency (mass removed by a given amount of laser energy) is higher for these processes as compared to surface vaporization and normal boiling where all material is removed in the vapor phase.

14.4.5 Effects of the Tissue Matrix on the Phase Transitions

In pulsed laser ablation of tissues, the phase transition processes are affected by the presence of the extracellular tissue matrix (ECM). For boiling processes within tissue, the vapor pressure necessary to drive bubble growth must not only overcome surface tension but also the elasticity of the tissue matrix [53]. Therefore, bubble growth in tissue requires a higher internal pressure than in pure liquids, and the elevated pressure is coupled to an increase in the boiling (saturation) temperature. The pressure increase that develops during the boiling process continues until it exceeds the ultimate tensile strength of the ECM and results in explosive tissue ablation [53]. We term this process *confined boiling*. On the *p*-*T* phase diagram in Fig. 14.10, the confined boiling process corresponds to a path 1→2→6, where the 2→6 transition is coincident with the binodal and terminates where the saturated vapor pressure equals the ultimate tensile strength of the tissue. In the presence of a tissue matrix, explosive material ejection will thus occur regardless of the rate of energy deposition. Thus, it is not surprising that explosive material ejection due to confined boiling has also been reported to occur in ablation using continuous irradiation at relatively low irradiances [54, 55].

In the above scenario, little vaporization occurs prior to the onset of ablation because bubble growth is impeded by the necessity to deform the tissue matrix. For

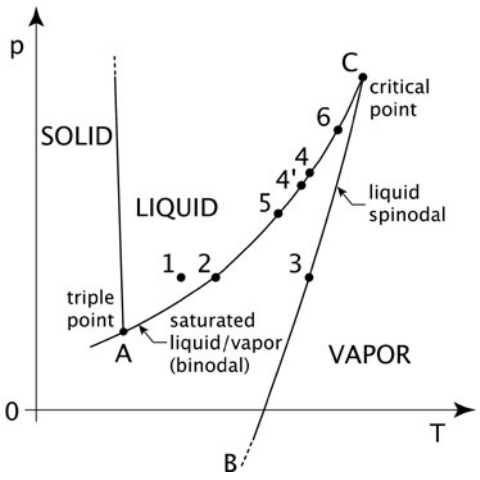


Fig. 14.10 Path taken through the *p* vs. *T* projection of the thermodynamic phase diagram for confined boiling (1 → 2 → 4 → 6) and for tissue ablation involving a phase explosion (1 → 2 → 3 → 4′ → 5 → 6). The actual path followed depends on the rate of energy deposition, number density of heterogeneous nuclei, and the mechanical strength of the tissue matrix relative to the saturation vapor pressure corresponding to the ambient spinodal temperature

mechanically weak tissues, much of the ejected mass consists of tissue that is fragmented and accelerated by the phase explosion. As a result, the ablation enthalpy can thus be considerably smaller than the vaporization enthalpy of water. However, for tissues that possess a strong matrix (e.g., skin), temperatures of 400–700°C are required to produce a saturation vapor pressure exceeding the ultimate tensile strength to initiate ablation [26]. Under these conditions, the ablation enthalpy often exceeds the vaporization enthalpy of water.

The tissue matrix retains its mechanical integrity during nanosecond or microsecond laser exposures even for temperature rises of several hundred degrees. There are several factors that are responsible for this. First, as discussed in Section 14.2, the temperatures required for disintegration of the matrix increase strongly as the duration of heat exposure decreases. Second, the application of tensile stresses to collagen fibrils stabilizes the helical architecture and results in a significant increase of the denaturation temperature. Thus the generation of tensile stresses resulting from pulsed laser heating is expected to further stabilize a collagen matrix with respect to possible collagen denaturation. Third, the extreme strain rates produced by pulsed laser ablation processes ($\sim 10^5 - 10^7$ s^{-1}) likely increase the ultimate tensile strength of the tissue matrix (see Section 14.2). Thus ablation does not involve a "liquefaction" of the tissue as assumed in earlier models [56, 57] but proceeds via the ejection of tissue fragments driven by the vaporization of tissue water.

For low rates of volumetric energy deposition $S = \mu_a \phi$ [W/cm^2] and high number densities of heterogeneous nuclei, the nature of the ablation process is largely independent of tissue mechanical properties. The laser irradiation will initially heat the tissue under equilibrium conditions at constant pressure (1→2) and then continue on the binodal until the ultimate tensile strength of the tissue is reached; resulting in explosive material removal. However, if the tissue is heated rapidly and/or a small number density of heterogeneous nuclei are present, the tissue water will be driven into a metastable state and a phase explosion will be initiated when the spinodal limit is reached ($1 \rightarrow 2 \rightarrow 3 \rightarrow 4' \rightarrow 5$). The subsequent evolution of the process now depends on the mechanical properties of the tissue. Immediate material ejection will result for tissues that are unable to withstand the stresses and deformations associated with the phase explosion. However, tissues possessing high collagen content, and thus high UTS, will not fail mechanically due to the phase explosion. The laser irradiation will then drive a confined boiling process as indicated by path (5→6) in Fig. 14.10 until the tissue ruptures at higher vapor pressures resulting in material removal. For pulsed ablation of skin, surface temperatures of 400–750°C have been measured [26] indicating that the dynamic tensile strength of the tissue matrix is higher than the pressure at the critical point ($p_c = 22.09$ MPa).

14.4.6 Vapor Explosion and Photothermal Dissociation of the Tissue Matrix

For ablation using nanosecond pulses, the volumetric energy densities achieved in the tissue water usually exceeds the vaporization enthalpy of water [27, 32]. Under

these conditions, the liquid water is completely transformed into vapor in a process termed *vapor explosion*. Moreover, at temperatures exceeding ~1000°C the constituent molecules of the extracellular matrix are thermally dissociated into volatile fragments. Energetically, these processes result in an ablation enthalpy higher than the vaporization of water and are "less efficient" than phase explosion or confined boiling because they do not involve the ejection of condensed material.

An important factor contributing to the high volumetric energy densities achieved in nanosecond laser ablation is the recoil pressure produced by the ablation of superficial layers of the target. At the beginning of laser exposure, ablation is governed by non-equilibrium surface vaporization. During this phase the recoil pressure is relatively small and a phase explosion occurs as soon as the temperature reaches the spinodal limit. With the onset of the phase explosion, the ablation has transformed from a surface-mediated to a volumetric process resulting in a rapid increase of the recoil stress. This large compressive stress inhibits the ablation of deeper tissue layers until the volumetric energy density is sufficiently high to cause a phase transition that can overcome these higher pressures. Therefore the phase transition of subsurface tissue layers will be more vigorous than the initial surface vaporization because a larger volumetric energy density is required to initiate the phase change process. This, in turn, produces a higher recoil pressure that impedes ablation in deeper tissue layers until even higher volumetric energy densities are reached. At any given depth, ablation starts as soon as the vapor pressure exceeds the recoil pressure resulting from the explosive removal of more superficial layers. This results in a "positive-feedback" process in which the volumetric energy density and pressure values required for the onset of ablation at deeper tissue layers will continue to increase as long as the laser irradiance is increasing within the laser pulse. After the peak irradiance of the laser pulse has passed, the volumetric energy density and pressure at the target surface will decrease while the ablation front continues to propagate into the target. The ablation process becomes most vigorous shortly after the peak of the laser pulse, as that is when the volumetric energy density in the target reaches a maximum value [27]. Since the evolution of thermodynamic states within the target is determined both by the incident laser irradiation and by the recoil produced by the ablation plume, ablation will likely continue well beyond the end of the laser pulse [1, 27, 58]. When the volumetric energy density in the target drops below the value required for thermal dissociation of the tissue matrix, the ejection of particulate tissue fragments will commence [27, 58]. Ablation ceases when the vapor pressure within the tissue falls below the ultimate tensile strength of the tissue matrix which itself is influenced by the local denaturation kinetics.

14.4.7 Effect of Stress Confinement on the Ablation Process

When performing ablation under conditions of stress confinement, the thermoelastic stresses modify significantly the phase transition processes that drive material removal. As discussed in Section 14.3, the thermoelastic stress wave propagation results in both compressive and tensile components (Fig. 14.5). The tensile stress

component can produce material ejection at temperatures less than 100°C in liquids with heterogeneous vapor/cavitation nuclei and in mechanically weak tissues such as liver. This phenomenon has been compared to back surface spallation produced by high-pressure impact [59, 60] and investigated in several studies [61, 62]. Nevertheless, temperatures above 100°C are usually necessary to initiate efficient tissue removal [61]. The important influence of tensile stress waves at temperatures above 100°C has only recently attracted attention [1, 33]. Figure 14.11 shows the path taken through the p–T phase diagram for a temperature rise above 100°C produced by laser irradiation with stress confinement. A heating phase (1→2), coupled with the generation of compressive stress, is followed by the passage of a tensile stress wave (2→3) that leads to a crossing of the spinodal limit, resulting in a phase explosion. The passage of the stress wave is followed by explosive boiling into the large number of bubbles produced shortly before (point 4), resulting in vigorous material ejection. This ejection occurs at temperatures lower than the spinodal temperature at atmospheric pressure (T=305°C) because, as shown in Fig. 14.11, the spinodal temperature is reduced with the decrease in pressure provided by the tensile thermoelastic stresses. It is important to note that the sequence of events described above may occur not only during surface ablation in a gaseous environment but also when laser pulses are focused into a transparent material. In fact, this process provides the basis of the high precision in femtosecond laser nanosurgery of cell [2].

At volumetric energy densities in excess of the spinodal limit at ambient pressure, i.e., for T>300°C, the superheated liquid is unstable and the onset of explosive ablation need not be initiated by the tensile component of the thermoelastic stress. Nevertheless, the thermoelastic stress transient can still contribute to material removal. The magnitude of thermoelastic transients produced by a given

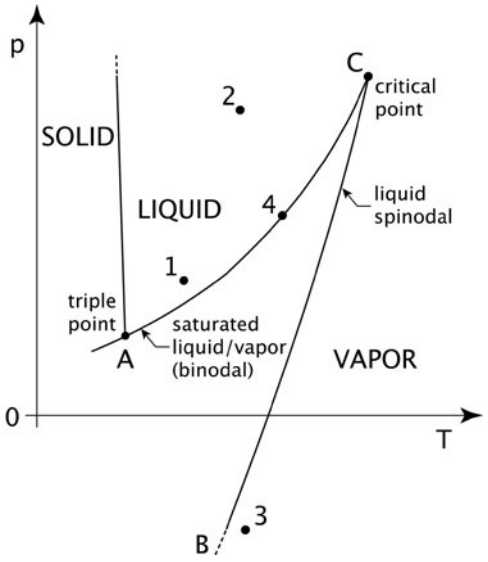


Fig. 14.11 Path taken through the p vs. T projection of the thermodynamic phase diagram for a temperature rise above 100°C under stress confinement conditions. The transition 1→2 corresponds to the heating phase that is coupled with the generation of compressive stress. The transition 2→3 corresponds to the passage of the tensile stress wave that leads to a crossing of the spinodal limit resulting in phase separation. After the passage of the stress wave, the system reaches point 4 that corresponds to explosive boiling into the large number of bubbles produced shortly before

temperature rise under conditions of stress-confinement is much larger than the saturation vapor pressure resulting from the same temperature rise and for $T > 1000°C$ may well exceed 1 GPa. The compressive component of the thermoelastic stress wave upon propagation will develop into a shock wave. The propagation of this shock wave into the depth of the target along with energy dissipation at the shock front [63, 64] result in tissue heating at locations beyond those heated directly by the laser irradiation and subsequent heat diffusion. Shock wave propagation thus serves as a form of convective heat transfer that extends the ablation depth and increases ablation efficiency [65]. Experimental evidence for shock wave induced phase changes of water after laser-induced breakdown was provided by Vogel and Noack [66]. For pulsed laser surface ablation, temperatures in the shock wave region will mostly be below the spinodal limit since a pressure jump in the neighborhood of 5 GPa is required to heat water from room temperature to 300°C [63]. Convective heat transfer will become important for ablation only for sufficient large volumetric energy densities and for very high degrees of stress confinement, i.e. very short pulse durations. We conclude that regardless of the volumetric energy density, stress confinement invariably serves to lower the ablation threshold and increase ablation efficiency [1, 33, 61, 65, 67].

14.5 Photochemical Decomposition

In 1982 it was discovered that intense ultraviolet laser pulses can etch synthetic polymer surfaces with sub-micrometer precision and without evidence of thermal or mechanical damage [68, 69]. Within one year, argon fluoride (ArF) excimer laser radiation ($\lambda = 193$ nm) was being explored to cut or reshape the surface of the cornea for the correction of visual defects [70]. To explain the unprecedented ablation characteristics, investigators invoked mechanisms unique to ultraviolet laser radiation. Because the dissociation energies of many organic molecular bonds are smaller than the photon energies (4–6.5 eV) at UV laser wavelengths ranging from $\lambda = 193$–355 nm, Srinivasan and Leigh hypothesized that the UV photons cause direct "bond-breaking" of the absorbing molecular chains in the organic target leading to ablative "photochemical decomposition" [71, 72]. Their theory assumes that material is ablated above a certain threshold irradiance which provides a bond dissociation rate that exceeds the rate of bond recombination [73, 74]. The molecular fragments have a lower density than the native macromolecules, and that creates an internal pressure rise and the ejection of decomposed material [75, 76].

Initially, it was believed that only a photochemical mechanism would provide the observed etching precision [75]. However, photothermal ablation can yield similarly small ablation depths provided that the optical penetration depth is sufficiently small and the conditions for thermal confinement are met [72, 76–78]. It was also believed that photochemical decomposition using UV wavelengths involved characteristics that would easily distinguish it from thermal ablation processes. However, this distinction proved difficult because following electronic excitation via UV absorption, other processes always compete with photochemical dissociation [79, 80]. Principal

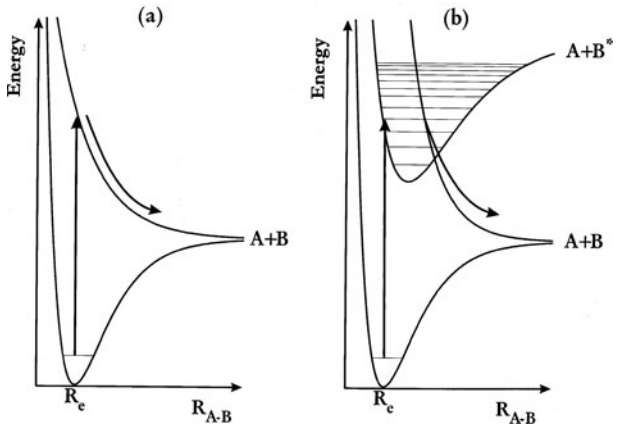


Fig. 14.12 Energy level diagram illustrating various pathways for photochemical bond-breaking. See text for further details. (Reprinted with permission from [80])

amongst these competing pathways is internal conversion of the absorbed photon energy to the vibrational modes of the molecule that is the basis for ablation via photothermal mechanisms.

Two generic types of *photochemical dissociation pathways* of a diatomic molecule *A-B* are illustrated in Fig. 14.12 [80]. The first type, shown in Fig. 14.12a, is one where electronic excitation promotes the molecule directly into an electronic state with no net bonding and results in a direct dissociation of the molecule into its constituent atoms. The second possibility, shown in Fig. 14.12b, represents a case where the electronic excitation promotes the molecule into a bound excited state matching that of a second, dissociative electronic state. Thus while absorption of the UV photon promotes the molecule into a bound vibrational level within the excited electronic state $A + B^*$, the electronic configuration can acquire the repulsive character of a dissociative state as the bond extends. The bond and repulsive energy potentials shown in Fig. 14.12 imply that for both of these pathways, photon energies significantly larger than the bond energy are necessary to achieve photochemical dissociation. The electronic energy exceeding the binding energy contributes to the kinetic or internal energy of the dissociation products. Therefore, upon thermalization of this excess energy, the ablation products will always be hot even if bond breaking is achieved solely through photochemistry.

Experimental investigations revealed that: (a) regardless of ablation mechanism the ablation fragments are thermalized on time scales comparable to the laser pulse [81]; (b) surface temperatures ranging from 800 to 2200°C are generated during the UV ablation of polymer substrates [82–84]; and (c) the fraction of laser pulse energy that does not manifest as heat in the remaining substrate is completely consumed for the radiation of a blast wave into the surrounding air [85]. While such studies provide strong evidence that a large fraction of the energy provided by UV radiation is thermalized in the substrate and in the ablation plume, they do not exclude the possibility that photochemical processes drive the molecular dissociation process [86].

Further attempts to identify the involvement of photochemical processes in UV ablation were made by examining the reaction products and possible chemical reaction pathways leading to their generation [74, 76, 87, 88]. Investigators expected that if photochemical processes were substantial, the prevalence of smaller molecular fragments would increase with shorter wavelengths because high single photon energy should be capable of breaking bonds of very high energy [76, 86, 89]. Unfortunately, these studies proved inconclusive for many polymers and for tissues as a similar composition of photoproducts is generated by both UV and IR ablation [78, 90]. However, Küper and co-workers' [91] analysis of the reaction pathways and products resulting from UV laser ablation of polymethyl methacrylate (PMMA) provided evidence of a significant photochemical contribution for ArF ($\lambda = 193$ nm) and KrF ($\lambda = 248$ nm) excimer laser ablation but not for XeCl ($\lambda = 308$ nm) excimer laser ablation, and Lippert and Dickinson presented several other examples for a photochemical contribution to polymer ablation [88]. The primary photochemical process in PMMA ablation is the scission of side chains that forms gaseous and volatile fragments. Degradation of the main polymer backbone was found to be a minor contribution relative to the total amount of photoproducts generated and considered to be a photothermal process. Kitai and co-workers [92] introduced the concept that the splitting of the bonds along the backbone of polymer molecules arises from the production of volatile molecules by photochemical scission of side chains. The volatile products possess a larger specific volume than PMMA and fracture the bonds in the main polymer by placing them under stress. The authors argued that the quantum yield of a purely photochemical process is much too low to explain PMMA ablation at the observed threshold radiant exposure [92–94], and that this mechanical decomposition pathway will increase the efficiency of the ablation process. They applied this concept to the physics of UV laser ablation of cornea by assuming that boiling of the water present in the tissue will provide the necessary stresses for the mechanical splitting of the bonds [92].

As a result of these studies and model development [88, 95], it is now generally accepted that UV laser ablation is never driven exclusively by photochemical dissociation and that thermal processes are also present [96]. Photochemical decomposition plays a significant, but not exclusive, role in the ablation of tissues at wavelengths around $\lambda = 200$ nm and, for some tissues, is also important at $\lambda = 248$ nm. However, its contribution drops rapidly when longer wavelengths are used [77, 97, 98]. Oraevsky and co-workers [98] estimated that the fraction of absorbed XeCl laser energy ($\lambda = 308$ nm) contributing to the photolysis of protein is only 2%.

14.6 Ablation Plume Dynamics

The phase transitions described in the previous section drive the formation of a plume consisting of material removed from the ablation site. Usually, the ablation dynamics and plume formation are not governed by just a single type of phase transition but result from an interplay of different transitions occurring at the target surface and in its bulk. The type and strength of the phase transition may change during

the laser pulse depending on the volumetric energy densities reached at each target location when the phase change occurs. The characteristics of the ablation plume reflect the underlying ablation dynamics and its analysis provides the insight necessary to draw conclusions about the phase transitions involved in a given ablation event. Furthermore, the plume dynamics influence the ablation process in various ways. The primary ejection of ablation products perpendicular to the tissue surface induces a recoil pressure that may produce additional, secondary material removal processes and cause collateral effects in the bulk tissue. Flow components parallel to the tissue surface that develop at later times may result in a redeposition of ablated material. Scattering and absorption of the incident light by the ablation plume reduce the amount of energy deposited in the target and limit the ablation efficiency at high radiant exposures.

To date, most investigations of the plume dynamics and acoustic phenomena associated with pulsed laser ablation of biological tissues have been performed experimentally by time-resolved photography, probe beam deflectometry, and spectroscopic techniques as reviewed in Refs. [1] and [99]. Here, we focus on the description of the plume dynamics itself rather than on the techniques of investigation. We first discuss the dynamics for water ablation and then progress to the more complicated case of tissue ablation where the primary ablation process and recoil-induced material expulsion are modified by the tissue matrix.

14.6.1 Primary Material Ejection in Nanosecond Ablation

For Q-switched laser pulses of 50–100 ns duration, the rate of energy deposition is extremely large. Close to threshold, the ablation process for liquids such as water is typically characterized by non-equilibrium mass transfer at the target surface [42] followed by a phase explosion of the superficial liquid layer [27]. However, when pulse energies well above the ablation threshold are used, large volumetric energy densities are produced in the target material that result in an ablation process characterized by more vigorous types of phase transitions. To illustrate this, Fig. 14.13 shows the sequence of events in the early phase of Q-switched Er:YAG laser ($\lambda = 2.94\,\mu\text{m}$, $t_p = 70\,\text{ns}$) ablation of water for a radiant exposure of $2.8\,\text{J/cm}^2; \approx 25\times$ the ablation threshold. The ablation dynamics is characterized by explosive vaporization followed by shock wave emission and the ejection of very fine droplets. The plume is initially fairly small but rapidly expands shortly after the peak intensity of the laser pulse is reached [47]. As a result, the majority of the ablated material is ejected towards the end and after the laser pulse. The layered structure of the plume reveals that different types of phase transition follow each other as the ablation front propagates into the target. The fact that the top part of the plume is completely transparent indicates that the volumetric energy density in the superficial target layers is larger than the vaporization enthalpy of water at room temperature under atmospheric pressure ($W = 2.59\,\text{kJ/cm}^3$). Therefore, this entire liquid volume is transformed into vapor in a "vapor explosion." When the ablation front has reached a depth where the energy density becomes smaller than

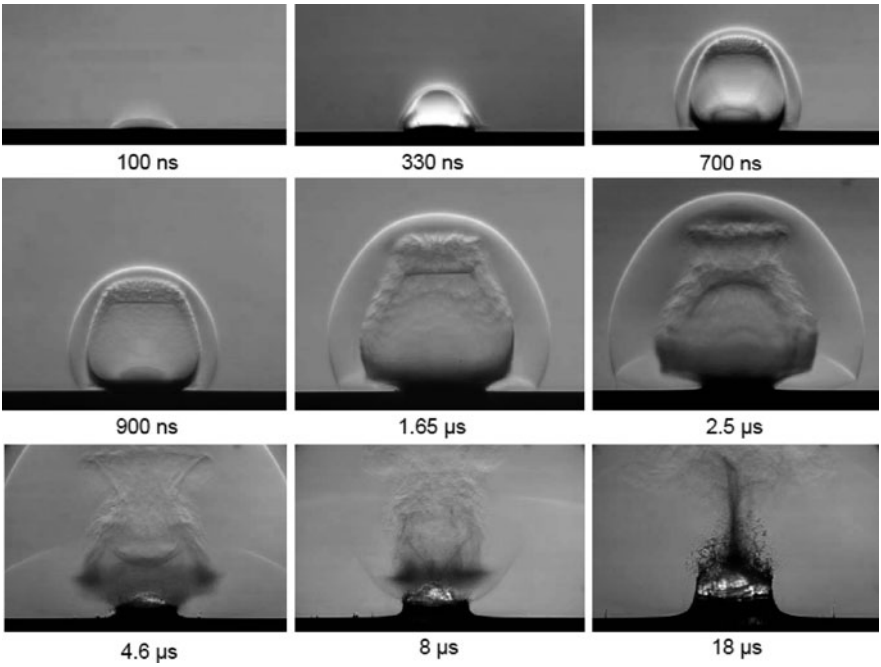


Fig. 14.13 Early phase of water ablation by a Q-switched Er:YAG laser pulse of 70 ns duration, photographed using a novel white light Schlieren technique [99]. The irradiated spot size was 700 µm, the radiant exposure 2.8 J/cm^2 (25x ablation threshold). All times refer to the beginning of the laser pulse. The dynamics is characterized by vapor plume formation, the emission of external and internal shock waves, droplet ejection, and the onset of recoil-induced material expulsion

the vaporization enthalpy of water, the superheated tissue water starts to decompose into vapor and liquid in a phase explosion, and droplet ejection commences. Droplet ejection is first visible after ≈700 ns and lasts for a few microseconds. The droplets cannot be resolved on the photographs and appear as a reddish haze. The reddish color indicates that the droplet size is sufficiently small to cause Rayleigh scattering by which blue light is scattered much stronger than red light [100]. As a consequence, the red spectral components of the illumination dominates the light that passes through the imaging optics. While the droplet ejection still continues, an indentation of the water surface forms and a "splash" region develops at the periphery of the ablation spot due to the recoil pressure produced by the phase transitions (see Section 14.6.3 below).

When soft tissues are ablated at moderate radiant exposures, the entire ablation plume consists of tissue fragments, as illustrated in Fig. 14.14b for Er:YAG laser ablation of liver at $t_p = 70$ ns and a radiant exposure of 1.4 J/cm^2. At the same radiant exposure, the top layer of a water target is already completely vaporized and thus transparent as shown in Fig. 14.14a. At a larger radiant exposure of 5.4 J/cm^2 (Fig. 14.14c), the top part of the plume becomes transparent for both water and liver ablation, and particulate fragments are ejected only after about 200 ns. The sequence

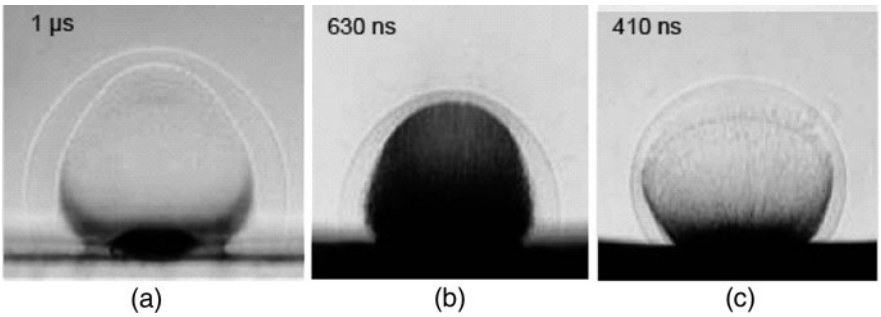


Fig. 14.14 Q-switched Er:YAG laser ablation of (**a**) water at $H = 1.4$ J/cm^2, (**b**) liver at $H = 1.4$ J/cm^2, and (**c**) liver at $H = 5.4$ J/cm^2. The plume consist of water vapor (*top*) and a droplet/vapor mixture in (**a**), tissue fragments in (**b**), and dissociated biomolecules (*top*) and tissue fragments (*bottom*) in (**c**). The volumetric energy densities averaged over the optical penetration depth are ≈ 5.2 kJ/cm^3 in (**a**), ≈ 4 kJ/cm^3 in (**b**), and ≈ 9 kJ/cm^3 in (**c**)

of gaseous ablation products followed by particulates could be visualized only by means of a photographic setup suited for detecting phase objects. In a previous study only the particulate fragments were observed and it was concluded mistakenly that the ablation process commences ≈150 ns after the end of the laser pulse [101]. In reality, the transparency of the top part of the plume indicates that during the initial ablation phase tissue water is completely vaporized and biomolecules are thermally dissociated into volatile fragments, which occurs at temperatures above 1000°C. For the liver target, the subsequent ejection of larger, non-transparent tissue fragments is driven by a phase explosion of the tissue water. The pressure developed during the phase separation (see Section 14.4.5) suffices to rupture the weak tissue matrix in liver parenchyma. The ejection ceases when the ablation front reaches a depth where the temperature drops below the stability limit of the superheated tissue water. The different optical appearance of the transparent and opaque parts of the ablation plume is due to differences in molecular composition and particle size distribution but not necessarily indicative for disparities in the average mass density.

For the ablation of skin at large radiant exposures, a similar sequence of biomolecule dissociation followed by ejection of tissue fragments was observed [27]. However, in this case the ejection of tissue fragments occurred over a shorter time interval than for liver. Ablation ceased when the ablation front reached a depth where the vapor pressure dropped below the tensile strength of the extracellular tissue matrix. Nevertheless, fragment ejection was found to continue for several microseconds after the laser pulse while the tissue matrix is increasingly weakened by thermal denaturation. Generally, the size of the ejected tissue particles is small at early times after the laser pulse and increases with time [27, 101]. The entire sequence of phase transitions occurring during water and tissue ablation is summarized in Fig. 14.15.

Since ablation becomes a volumetric process once the spinodal limit is exceeded and a phase explosion sets in, it is not obvious why large volumetric energy densities

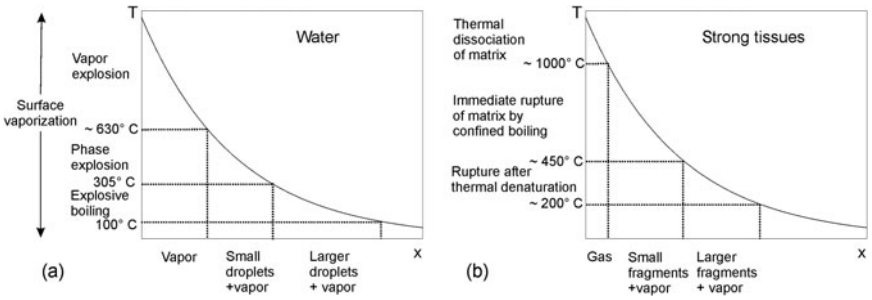


Fig. 14.15 Sequence of phase transitions and corresponding plume constitutents in ablation at radiant exposures well above threshold for (**a**) water ablation, and (**b**) ablation of mechanically strong tissues such as, for example, skin

sufficient for a vapor explosion and dissociation of biomolcules should be reached in pulsed laser tissue ablation. This occurs because the recoil stress produced by the phase transitions of the uppermost tissue layers delays the phase transitions in underlying layers through an increase of the spinodal temperature (see Fig. 14.6). The ongoing absorption of laser energy into the underlying layers can thus drive the thermodynamic state into the supercritical regime. Even larger recoil stresses are produced when these layers are ablated, and the phase transitions in deeper layers are delayed even more. This "positive-feedback" process continues at least until the intensity peak of the laser pulse is reached after which a relaxation process resulting in explosive ablation commences and continues for several microseconds after the end of the laser pulse. The energy densities generated during the runaway process are in the order of 10 kJ cm^{-3} and give rise to recoil pressures of several hundred MPa [27] (see Section 14.6.3 below).

The high volumetric energy density in the target material produced in Q-switched Er:YAG laser ablation results in a very large initial expansion velocity of the ablation plume that drives the emission of an equally fast shock wave. Shock front velocities are usually on the order of 2000–4000 m/s for both IR and UV wavelengths [27, 102–104], i.e., they reach values up to Mach 12. Measured shock wave and plume velocities correlate with the water content of the samples because lower water content results in smaller volumetric energy densities and less vigorous ablation. By contrast, the velocity of particulate fragments is larger for mechanically strong tissues (up to 1700 m/s for skin) than the velocity of droplets ejected in water ablation (up to 280 m/s) [27]. This is because the temperature required for thermal dissociation of the tissue matrix into volatile products is higher than the temperature required for complete vaporization of water. Therefore, tissue fragments become visible early in the ablation process when the ablation front has reached a depth at which the temperature is below the level required for thermolysis. At this time, the pressure driving the ejection is still very high. By contrast, droplet ejection starts only once the temperature at the ablation front has reached a lower level corresponding to the onset of a phase explosion. This results in smaller velocities for the droplet ejection.

The ablation plume exhibits complex dynamics. The plume expansion is nearly spherical during the initial phases of expansion but begins to propagate preferentially in the forward direction after 1–2 μs. For small radiant exposures, the interaction of the piston-like forward movement with the ambient air at rest results in ring vortex formation [27, 105]. For larger radiant exposures, a region of high density and pressure is created at the contact front between plume and surrounding air. The molecules and molecular clusters propagating with the plume possess a nonzero average velocity. When they collide with air molecules that are, on average, at rest, they are partially reflected back into the plume. As visible in Fig. 14.13, this reflection leads to the formation of an internal shock wave that begins to propagate toward the target surface when the rarefaction from the plume expansion has reduced the pressure in the plume considerably below its initial value [27, 106, 107]. The internal shock interacts with the particles and droplets of the plume and deforms the shape of the particle cloud during a time interval lasting about 10 μs. Due to the heating at the shock front, the passage of the internal shock wave through the reddish droplet cloud results in their vaporization.

The propagation of the shock front after a strong explosion in a homogeneous atmosphere was first theoretically described by Taylor [108] and Sedov [109, 110] and, using a higher order approximation, by Sakurai [111, 112]. These theories neglect the mass of the gas and debris driving the shock wave and are thus valid only once the shock wave has swept over a mass of atmospheric gas much greater than the mass in which the energy was initially concentrated. Various authors have later obtained solutions for the mass-dependent flow regime [113–115], and simple analytic solutions are available for some limiting cases. When the mass of the gas encompassed by the shock wave is much greater than the initial ablated mass and the pressure driving the shock is much greater than the atmospheric pressure ahead of the shock front, the position $R\,(t)$ of a spherical shock wave is governed by [116]

$$R(t) = \xi_1 \left(E_0/\rho_0\right)^{1/5} t^{2/5} \tag{14.8}$$

and that of a planar shock wave such as emitted from a large irradiated spot size by [85]

$$R(t) = \xi_2 \left(E_0/\rho_0\right)^{1/3} t^{2/3} \tag{14.9}$$

Here E_0 [J] is the energy driving the explosion, ρ_0 the density of the undisturbed gas, and $\xi_1[\mathrm{m}^{2/5}.\mathrm{kg}^{1/5} \cdot \mathrm{J}^{-1/5} \cdot \mathrm{s}^{-2/5}]$ and $\xi_2[\mathrm{kg}^{1/3} \cdot \mathrm{J}^{-1/3} \cdot \mathrm{s}^{-2/3}]$ are constants that depend on the specific heat capacity ratio γ of the gas. The peak pressure in the spherical case scales proportional to $E_0^{2/5}$ [110]. Once the shock wave pressure becomes comparable to the ambient pressure, its propagation is better described by the Jones' approximation [117, 118]. When the mass of the material removed is very large or the background pressure very low (including vacuum), the motion of a planar shock wave can be described by [119]

$$R(t) = \xi_3 \left(E_0/M_0\right)^{1/2} t \tag{14.10}$$

where M_0 is the mass of the explosive debris. ξ_3 is a constant with dimensions $[m^2 \cdot kg \cdot J^{-1} \cdot s^{-2}]$. A comparison of experimental $R(t)$ data with Eqs. (14.8)–(14.10) allows an assessment of the transduction of laser pulse energy into blast wave energy E_0 [85, 118].

More refined numerical simulations by Brode [107] and the analytical treatment by Arnold and coworkers [106] include the spherical movements of the external shock front, the contact front between plume and ambient gas, and the internal shock front within the plume. Recently, Chen, Bogaerts and Vertes [120] presented a model for the propagation of the external shock wave propagation in atmospheric pressure laser ablation of water-rich targets that incorporates the nonlinear absorption of water and the phase explosion due to superheating. The model predicts a succession of an initially slow plume emission followed by a vigorously accelerated expansion, in good agreement with the experimental results of Apitz and Vogel [27] and the views presented above.

14.6.2 Primary Material Ejection in Microsecond Ablation

Free-running lasers typically provide pulse durations longer than 100 μs. Thus, unlike nanosecond ablation, plume formation and expansion occurs largely during the laser irradiation. As a result, the ablation plume influences the energy deposition of the laser radiation into the tissue target, and the plume dynamics is also influenced by the interaction of the laser beam with the ejected material. Nevertheless, the succession of a sub-ablative phase, development of a vapor plume, and material ejection is similar as with nanosecond pulses even though it occurs on a much longer time scale [121]. However, the heating rates available from microsecond laser pulses are generally much smaller than those available from nanosecond laser pulses of moderate to high radiant exposures. These lower heating rates are not sufficient to generate the temperatures necessary to dissociate molecules of the extracellular matrix and are only able to produce supercritical water at very large radiant exposures.

Free-running laser emission is characterized by intensity fluctuations during the laser pulse ("spiking" behavior). These intensity peaks modulate the vaporization and material ejection rates [101, 122] as well as the emission of acoustic transients generated during the ablation process [123, 124]. The intensity spikes of the laser pulse are coupled with the generation of individual transients as shown in Fig. 14.16.

The mechanisms leading to material ejection are the same as for nanosecond pulses: a phase explosion for mechanically weak materials and a succession of phase explosion and confined boiling for mechanically stronger tissues. Previously it was believed that the generation of a phase explosion requires very fast heating rates achieved only when using nanosecond pulses [32]. However, using time-resolved photography, Nahen and Vogel [121] demonstrated that a phase explosion can also be produced with laser pulses with duration on the order of 200 μs. This is shown in Fig. 14.17 that compares the ablation dynamics for Er:YAG laser irradiation of water, gelatin and skin using identical radiant exposures. The rapid droplet ejection during Er:YAG laser ablation of water that is observed after 40 μs can only

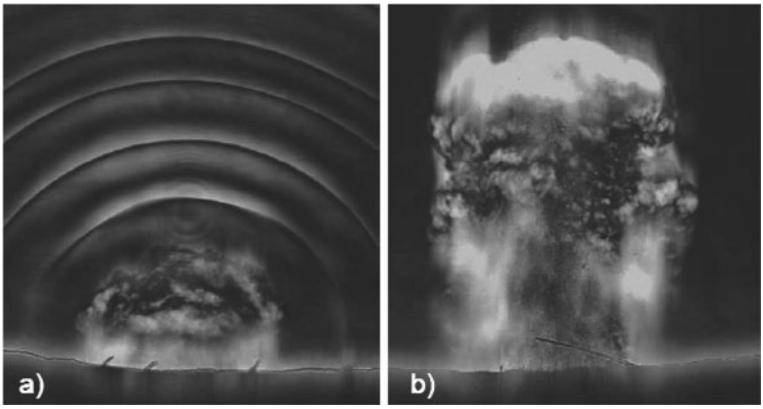


Fig. 14.16 Dark-field Schlieren images of the acoustic transients and ablation plume during skin ablation with a 200-μs Er:YAG laser pulse ($H = 20$ J/cm^2, spot size 2.3 mm) photographed (**a**) 22.4 μs and (**b**) 40 μs after the onset of the laser pulse. The images show acoustic transients arising from individual spikes in the free-running laser irradiation, and the plume containing vapor and tissue fragments. (Reprinted with permission from [99] with permission.)

be produced by a phase explosion because in the absence of stress confinement, no other mechanism gives rise to a material ejection perpendicular to the water surface. In gelatin, a phase explosion occurs at the same time as in water. However, the phase explosion only deforms the gelatin surface without rupturing it, and fracture of the gelatin surface and rapid particle ejection are observed only after a further pressure build-up through confined boiling (Section 14.4.5). The material ejection during skin ablation is also characterized by a phase explosion followed by confined boiling. However, the higher mechanical strength of skin causes a further delay of material ejection compared to gelatin. It is important to note that both for skin and gelatin targets fragments are ejected in the form of solid particles. The absence of droplet-like ejecta indicates that gelatin exposed to temperatures near the spinodal limit does not melt within of 200 μs, even though it melts at 60°C for sufficiently long heat exposures. This finding is consistent with the strong increase in denaturation temperature for very short exposures discussed in Section 14.2.

Initial material ejection velocities observed for microsecond laser ablation are roughly one order of magnitude lower than those reported for nanosecond ablation [125]. For free-running pulses, an increase of the radiant exposure results in an earlier onset of the material ejection but does not change the ejection velocity significantly. It is only for very large radiant exposures in which the first intensity spike of the free-running pulse provides a dose in excess of the ablation threshold that an increase of the ejection velocity is observed. By contrast, for nanosecond exposures an increase of the radiant exposure is always coupled with an increase of the volumetric energy density that translates directly into a higher temperature, pressure, and ejection velocity.

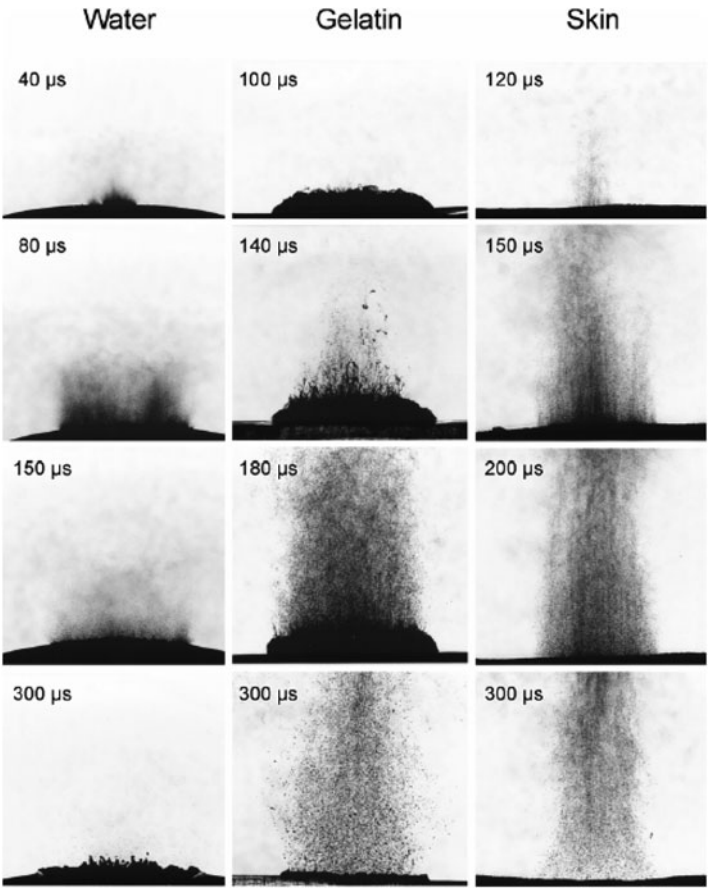


Fig. 14.17 Dynamics of Er:YAG laser ablation of water, gelatin with 70% water content, and skin using a radiant exposure of 4.6 J/cm^2, 5 mm spot size, and 200 µs pulse duration. The times after the beginning of the laser pulse when the photographs were taken are indicated on the individual frames. Note the increasing delay in the ejection of particulate matter with increasing mechanical strength of the target. (Reprinted with permission from [121] and [1].)

In both Q-switched and free-running laser ablation of soft tissues, material ejection continues for a considerable time following laser irradiation that can last up to several milliseconds [101–103, 105, 121, 126]. In general, post-pulse ablation lasts longer for mechanically weaker tissues, larger radiant exposures, and larger laser beam diameters. One possible driving force for the continuation of the ablation process after the end of the laser pulse is the heat retained in the tissue. A progressive weakening of the tissue matrix through thermal denaturation enables a propagation of the ablation front until the vapor pressure in the residual tissue drops below the ultimate tensile strength of the weakened tissue matrix. Another very important source of post-pulse ablation are hydrodynamic phenomena such as recoil stress-induced material expulsion.

14.6.3 Recoil Stress and Secondary Material Ejection

Both the rapidly expanding vapor plume and the ejected particles generate recoil stresses that impart momentum to the tissue. The linear momentum per unit area of the ablated material I [Pa·s] is the time integral of the recoil stress σ_{rec} [Pa] at the target surface

$$I = \int_0^\infty \sigma_{\text{rec}}(t)dt \tag{14.11}$$

A derivation of the peak recoil stress requires assumptions on the nature and duration of the ablation process. Various authors have presented solutions for the peak stress amplitude produced by a continuous vaporization process [31, 57, 127, 128], and by explosive ablation where the entire laser pulse is deposited prior to the onset of material removal [32, 129].

Experimental values for the recoil stress produced by nanosecond laser ablation have been obtained through direct pressure measurements using piezoelecteric transducers [31, 32, 130, 131], and, for water ablation, through analysis of the speed of the recoil-induced shock wave [27]. Peak pressures range from a few MPa at the ablation threshold up to several hundred MPa for radiant exposures well above threshold. For free-running microsecond laser pulses, average stress values during the laser irradiation have been determined through measurement of the recoil momentum using the ballistic pendulum method [56, 128], while the peak stress amplitudes produced by the intensity maxima of the free-running pulses, was obtained via transducer measurements [131]. Peak values of recoil stress produced during cornea ablation using free-running Er:YSSG laser irradiation at a radiant exposure of 50 J/cm^2 amounted to 2 MPa [131], while the average pressure value for skin ablation at the same radiant exposure was only 0.3 MPa [128].

In stress-confined tissue ablation, the compressive recoil stress transient is superimposed on a bipolar thermoelastic transient [132]. Figure 14.18 demonstrates the transition from a bipolar stress transient for radiant exposures below the ablation threshold to a monopolar compressive transient when the ablation threshold is exceeded. This transition and the corresponding increase in peak pressure is a sensitive method for the determination of the ablation threshold [31, 32, 130].

The recoil stress produced by both vaporization and material ejection in the primary ablation phase can induce a secondary material expulsion process that leads to a strong increase of the ablation efficiency [27, 127]. Recoil-induced material expulsion is most pronounced during ablation of liquids and mechanically weak tissues. Material will be ejected whenever the recoil stress component in the radial direction exceeds the mechanical strength of the tissue, as illustrated in Fig. 14.19.

The sequence of primary material ejection and recoil-induced material expulsion is shown in Fig. 14.20 for free-running and Q-switched Er:YAG laser ablation of liver. While the primary material ejection visible at short delay times takes place across the entire ablation area, recoil-induced expulsion occurs preferentially at

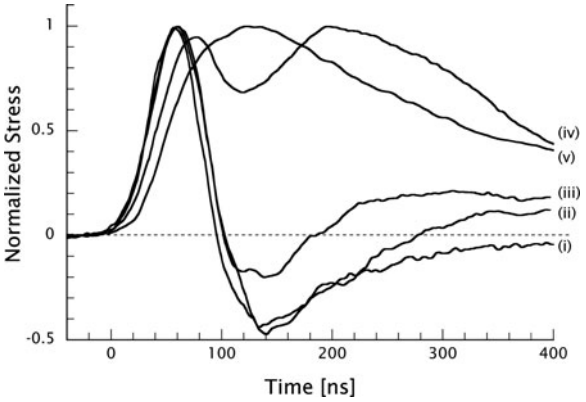


Fig. 14.18 Stress transients resulting from TEA CO_2 laser ($t_p = 30$ ns) irradiation of porcine dermis for radiant exposures below (i), at (ii), and above (iii)–(v) the ablation threshold. Radiant exposures below threshold produce bipolar thermoelastic stress transients. For radiant exposures equal to and above threshold, a compressive pulse is produced by the ablative recoil. (Adapted from [32] with permission. Copyright 1996 Biophysical Society)

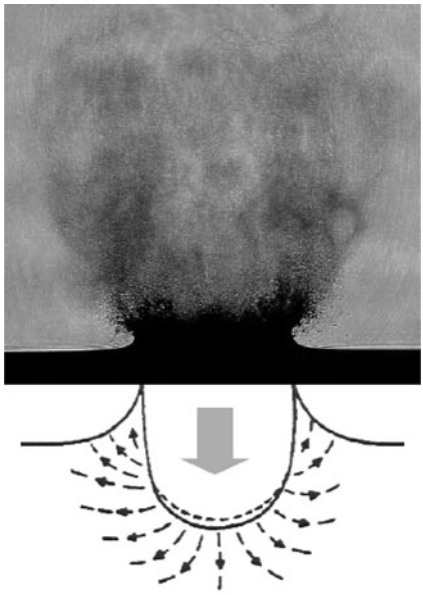


Fig. 14.19 Recoil-induced material expulsion in water ablation by 200-μs Er:YAG laser pulses, together with a schematic illustration showing the pathlines of the ejected material fragments. The lateral component of the recoil-induced flow collides with the surrounding fluid that is at rest, thus producing an upward directed splash

the ablation crater rim and includes the ejection of tissue fragments much larger than those ejected during the initial phase explosion. The recoil-induced ejection dynamics resembles the surface indentation and subsequent "splash" produced by the impact of liquid droplets on bulk liquids that has already been investigated in considerable detail [133, 134]. The mass expelled at later times far exceeds the

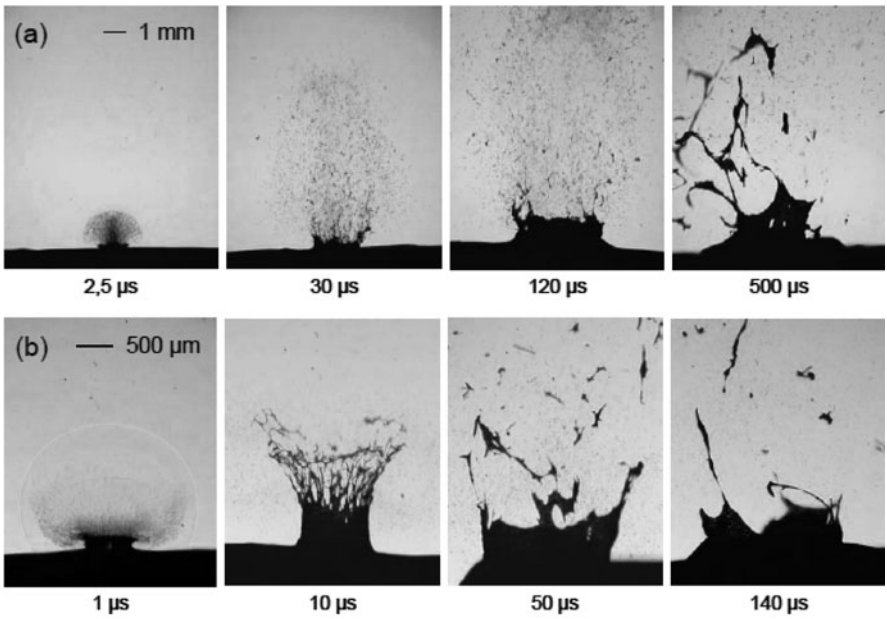


Fig. 14.20 Recoil-induced material expulsion for liver ablation by (**a**) 200-µs Er:YAG laser pulses at 100 J/cm^2 radiant exposure and 1.1 mm spot size, (**b**) 70-ns Er:YAG laser pulses at 5.4 J/cm^2 radiant exposure and 0.5 mm spot size. The primary material ejection produced by the phase changes in the target is also visible all images of (**a**) and in the first image of (**b**)

mass ejected during the primary ablation phase. However, the velocity of the ejecta is considerably slower.

Recoil-induced material expulsion begins after the primary ejection process, requires a radiant exposure well above the ablation threshold, and provides an increase of the ablation efficiency. A marked increase of the ablation efficiency at a certain radiant exposure has been observed for weak tissues as liver and myocardium as well as for gelatin with high water content but not for tissues with greater mechanical strength such as skin [27, 135]. Remarkably, no recoil-induced ejection was observed in skin ablation using Q-switched Er:YAG laser pulses even when the recoil stress was about 50 times larger than the quasi-static ultimate tensile strength of skin [2]. To understand this discrepancy, we must first consider that the recoil-induced tensile and shear stresses that contribute to tissue fracture may be considerably smaller than the measured compressive recoil stress. Moreover, the dynamic tensile strength of tissue at the extreme strain rates produced in pulsed laser ablation is much higher than the quasi-static values for the ultimate tensile strength found in the literature [15–17] (see Section 14.2). Finally, tissue fracture will only occur at sufficiently large strain that may not be achieved by stress transients of very short duration [1, 136, 137].

Figure 14.21a demonstrates that for mechanically weak tissues, the recoil-stress-induced material expulsion produces craters with a depth much larger then

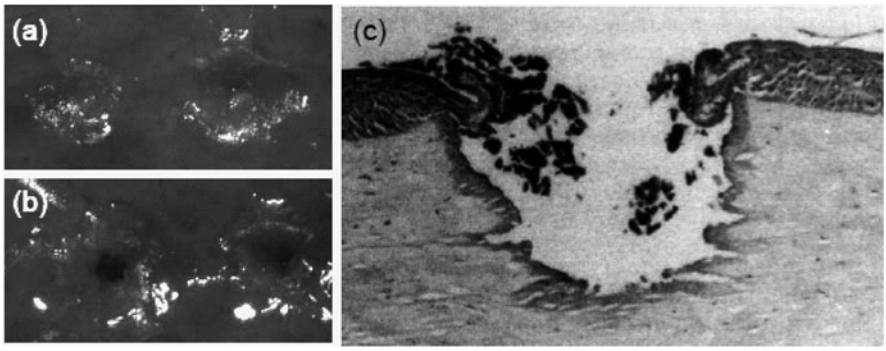


Fig. 14.21 (**a**) Crater produced during liver ablation by a 200-μs Er:YAG laser pulses of 32 and 50 J/cm^2 radiant exposure (2.5 and 1.6 mm spot size, respectively). (**b**) Histologic slide showing an ablation crater in bovine cornea produced by an Erbium:YSGG laser pulse with $t_p = 250$ μs and $H = 100$ J/cm^2. The tissue around the ablation crater exhibits a 25–50 μm zone of thermal damage (dark) and mechanical tearing between the corneal lamellae. ((**b**) Reprinted with permission from [128])

the optical penetration depth and a diameter much larger than the irradiated spot size. For mechanically strong tissues, the recoil stress does not lead to material expulsion. However, it can produce tissue tearing at the sides of the ablation crater as seen in Fig. 14.21b. The cracks and tearing patterns arise preferentially along morphological structures with reduced mechanical strength such as the transitions between corneal lamellae, sinusoid spaces holding blood between plates of cells in liver tissue, and their orientation is also influenced by the weakness of the longitudinal strength of blood vessels compared to their circumferential strength [128]. Tissue tearing at the rim of ablation craters was not observed for skin due to its 3-dimensional collagen network that can bear mechanical loads in an isotropic fashion.

Recoil-induced stress transients can also produce more subtle forms of collateral tissue damage further from the irradiation site. ArF-excimer ($\lambda = 193$ nm) laser ablation of skin may create epidermal and dermal photoacoustic damage [138, 139]. The recoil stress wave produced during ArF-excimer laser ablation of the cornea is focused into the eye ball and, due to diffraction, transformed into a bipolar shape with a tensile stress amplitude of up to 3.5 MPa [140]. Könz and coworkers [131] demonstrated recoil-induced damage of the corneal endothelium after mid-IR laser ablation of the corneal stroma that was due to the tensile stress generated upon partial reflection of the compressive recoil stress transient at the cornea-aqueous interface.

Thus, to achieve precise and gentle tissue ablation it is not sufficient to simply select a laser wavelength with small optical penetration depth and a pulse duration providing thermal confinement. In addition, one must avoid the production of extensive recoil stresses that may degrade the quality of the ablated surface and/or induce collateral mechanical damage. This restriction imposes an upper limit for the incident radiant exposure.

14.6.4 Shielding and Flow-Induced Material Redeposition

Absorption, scattering, and diffuse reflection of incident laser light by the ablation plume leads to a reduction of the energy delivered to the target tissue and a reduction of the ablation efficiency. Direct measurements of the diffuse reflectance of the plume [141] and of the entire reduction of optical transmission through the plume [121] provided values of the extinction coefficient within the plume produced by soft tissue ablation using Er:YAG laser irradiation ($\lambda = 2.94\,\mu$m, $t_p = 200\ \mu$s) on the order of 1 cm^{-1} [121].

Shielding is strongly enhanced when a series of pulses is applied instead of single pulses [141]. For skin ablation by means of Er:YAG laser pulses ($t_p \approx 300\,\mu$s) irradiating a spot of 2 mm diameter, Kaufmann and Hibst [142] observed a decrease of the etch depth per pulse from 40 to 10 μm when the pulse repetition rate was increased from 1 to 10 Hz. The reduction of the ablation efficiency was attributed to increased shielding by the ablation plume. When a considerably smaller spot size is used, the lateral spread of the plume removes a larger fraction of the ablation products out of the beam path, and the etch depth does not decrease with increasing repetition rate [143].

At very high radiant exposures, plasma formation in front of the target may lead to a further decrease of the optical transmission to the target [27, 105, 126] as illustrated in Fig. 14.22. Plasma formation starts at the target surface but the plasma grows rapidly after ignition into the space in front of the target surface, which results in very efficient shielding.

When radiant exposures close to the ablation threshold are used, the plume acquires a mushroom-like shape that exhibits a ring vortex at its top, a thin stem

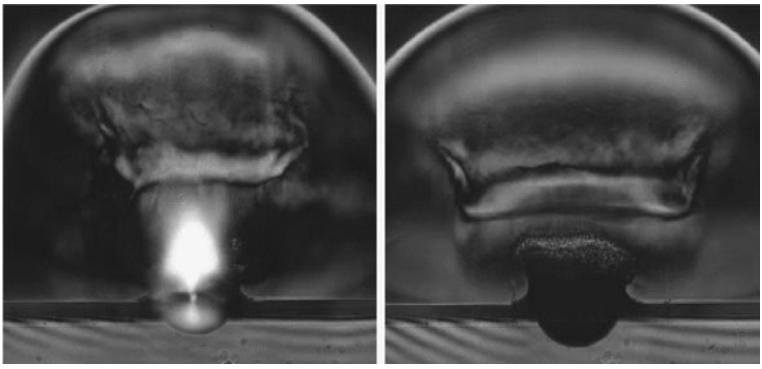


Fig. 14.22 Dark field images of water ablation at 5.4 J/cm^2 radiant exposure with and without plasma formation (0.5 mm spot size, image taken 2 μs after the laser exposure). The plasma originates in hot spots at the water surface and grows into the incoming laser beam. The recoil-induced cavity is considerably smaller in the case with plasma formation due to the "shielding" of the target by light absorption in the plasma. (Reprinted with permission from [27])

with a diameter smaller than the ablation spot and a radial flow component parallel to the surface at the foot of the plume [27, 105, 126]. The radial flow parallel to the tissue surface can result in a redistribution of ablation products across the ablation spot. For example, when performing corneal refractive surgery using large laser spot sizes, ablation rates were found to be smaller in the center of the ablation zone than in its periphery even though the irradiance was spatially homogeneous. As a result, "central islands" remained that distorted the intended refractive correction [145]. Photographic investigations of the plume dynamics revealed that the "central islands" are the result of a redeposition of ablated material after the end of the laser pulse that preferentially occurs near the stagnation point of the flow at the center of the ablated area [105]. A second factor contributing to the nonuniformity of the ablation rate is the attenuation of subsequent laser pulses in the center region of the ablated area by remnants of the plume from previous pulses that preferentially stay in the vicinity of the stagnation point [105].

14.7 Ablation Models and Metrics

In the previous sections we established that pulsed laser ablation always consists of a sequence of different phase transition processes that occur during and after laser irradiation. Moreover, we have shown that the initial primary material ejection is often followed by a secondary, recoil-induced ejection. The type and vigor of the phase transition and ejection processes during an individual ablation event depend on both the laser irradiance and radiant exposure as well as on the optical and mechanical tissue properties. Thus, it is impossible to formulate a simple comprehensive ablation model that describes these different aspects of the ablation process. Nevertheless, it is useful to formulate simplified models that elucidate basic features of the ablation behavior and parameter dependencies for specific ablation regimes. Particularly useful are models that illustrate the ablation behavior in extreme cases such as the "steady state" model for long laser pulses and the "blow-off" model for very short pulses. As long as one keeps the limitations of these models in mind, they can provide useful guidelines to interpret experimental data and to select laser parameters for an intended ablation outcome.

In this section, we first introduce common metrics that characterize the ablation process such as the threshold, enthalpy, and efficiency of ablation. We then describe the heuristic blow-off and steady-state models that use these metrics to predict ablation rates without reference to mechanistic aspects of the ablation process. In the next step, we discuss models that link the ablation outcome to underlying mechanisms. Models have not yet developed to the stage where they can address the full complexity of phase transitions and ejection mechanisms. Nevertheless, the models have provided insight into specific aspects of the ablation dynamics and the resulting collateral damage. Finally, we briefly describe the molecular dynamics approach that provide a microscopic view of the inception and dynamics of the ablation process.

14.7.1 Ablation Metrics

The *ablation threshold* H_{th} [J/cm^2] represents the minimum incident radiant exposure required to achieve ablative material removal. Molecular dynamics simulations and experiments show that radiant exposures below the ablation threshold results in thermal desorption of individual molecules while exposures above the ablation threshold produce volumetric ejection of large molecular clusters [93, 146, 147]. Thus, the ablation threshold is defined as the incident radiant exposure necessary to achieve *volumetric* ejection of tissue.

Perhaps the simplest means to determine the ablation threshold is to determine the radiant exposure at which the extrapolation of data for material removal or etch depth vs. incident radiant exposure intersects the abscissa, as indicated in Fig. 14.23 [144]. However, this approach may be inaccurate due to a lack of measurement sensitivity at radiant exposures close to the ablation threshold. A more sensitive, albeit indirect, measurement of ablation threshold has been obtained using piezoelectric transducers to measure the appearance of recoil stresses connected with material removal [31, 32, 130, 148]. High-speed photography has provided a sensitive and direct determination for the onset of material removal during pulsed laser irradiation [27, 121, 149], as shown in Figs. 14.13, 14.16 and 14.17. These measurements, when combined with the time-resolved recording of the laser irradiance, can provide an accurate measurement of the threshold radiant exposure.

The *ablation enthalpy* h_{abl} [J/kg] or heat of ablation is often used to describe the energetics of ablation [1, 135]. As depicted in Fig. 14.23, ablation enthalpy is a differential quantity that varies with the incident radiant exposure and represents the incremental amount of laser energy required to ablate an additional mass of tissue. We shall see that steady-state models invoke a constant ablation enthalpy that can be determined by fitting a line to etch depth vs. incident radiant exposure data and taking the reciprocal of the slope of this line. In the blow-off model the ablation enthalpy increases monotonically with radiant exposure.

Ablation efficiency η_{abl} [kg/J] is a metric for the total energy necessary to remove a given mass of tissue (Fig. 14.23). Ablation efficiency decreases with

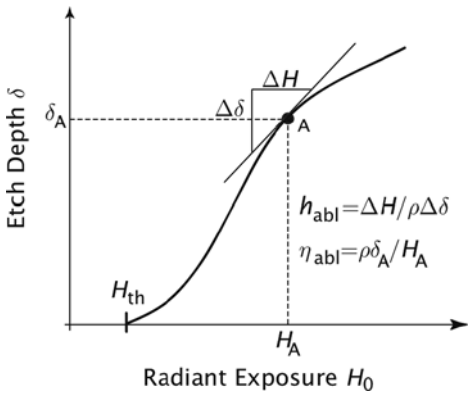


Fig. 14.23 Graph depicting the relationship between etch depth and radiant exposure for a hypothetical tissue ablation process to illustrate the concepts of ablation threshold H_{th}, ablation enthalpy h_{abl}, and ablation efficiency h_{abl}

increasing ultimate tensile strength (UTS) of tissue and increases with radiant exposure [27, 150]. For tissues with high UTS, such as skin and aorta, the ablation efficiency reaches a maximum at radiant exposures several times the ablation threshold and remains roughly constant while, for liver, a very weak tissue, the ablation efficiency continues to increase even for radiant exposures 10 times larger than the ablation threshold [27, 150]. The continued increase in ablation efficiency is due to the susceptibility of mechanically weak targets to recoil-induced material removal (Section 14.6.3).

14.7.2 Heuristic Models

Blow-off Model. The blow-off model was first developed to predict the etch depth resulting from nanosecond UV laser ablation of polymers [69, 151] and has also been applied to tissue ablation. This model is based on four assumptions. First, the Lambert-Beer law accurately describes the spatial distribution of absorbed laser energy in the tissue. Second, a threshold radiant exposure H_{th} is required for the initiation of ablation and lower radiant exposures result only in heating of the target. Third, material removal commences only after the end of the laser irradiation. Fourth, the conditions for thermal confinement are satisfied. Thermal confinement is nearly always achieved when performing laser ablation using pulse durations of 100 ns or less. The distribution of the volumetric energy density within the target immediately after the end of the laser pulse is then described by $W(z) = \mu_a H_0 \exp(-\mu_a z)$. For incident radiant exposures H_0 larger than the threshold radiant exposure H_{th}, all material receiving a radiant exposure in excess of H_{th} is removed. This results in a semi-logarithmic relationship between the etch depth δ [m] and the incident radiant exposure:

$$\delta = \frac{1}{\mu_a} \ln\left(\frac{H_0}{H_{th}}\right) \tag{14.12}$$

The use of laser wavelengths at which the tissue absorption coefficient is large results in a low threshold radiant exposure. As illustrated in Fig. 14.24a, the superficial penetration of laser radiation results in small etch depths that increase slowly with radiant exposure. By contrast, small tissue absorption coefficients result in much larger threshold radiant exposures, and the deeper optical penetration produces a much steeper slope of the etch depth vs. radiant exposure curve. These predictions of the blow-off model (Fig. 14.24a) are corroborated by a comparison with experimental data for 193 and 248 nm pulsed excimer laser ablation of cornea shown in Fig. 14.24b [152, 153].

Steady-State Model. For microsecond laser pulses, material removal typically occurs during irradiation of the target. Such processes can be modeled by assuming that a fixed amount of energy is required for removal of a unit mass of tissue, i.e. h_{abl} [J/kg] is constant. This is adequate to describe a continuous ablation process

in which the resulting rate of material removal balances the irradiance delivered to the tissue. It is also assumed that material removal begins soon after the beginning of laser irradiation and continues during the entire laser pulse. For doses in excess of the ablation threshold, the steady-state model predicts a linear relationship between the etch depth and incident radiant exposure. The slope of the etch depth and incident radiant exposure curve depends solely on the ablation enthalpy h_{abl} as:

$$\delta = \frac{H_0 - H_{\mathrm{th}}}{\rho\ h_{\mathrm{abl}}} \tag{14.13}$$

where h_{abl} is assumed to be constant. Note that there is no explicit dependence of the etch depth on absorption coefficient and there is a direct proportionality between the threshold radiant exposure and ablation enthalpy, namely, $H_{\mathrm{th}} = \rho h_{\mathrm{abl}}/\mu_{\mathrm{a}}$.

Comparison of Blow-off and Steady-State Models. To predict material removal, heuristic models require as input the optical absorption coefficient and either the

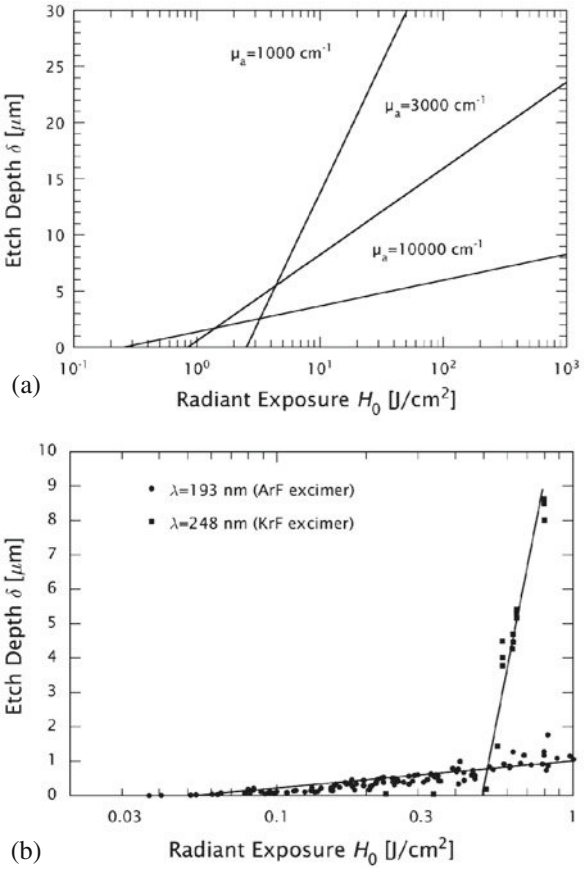


Fig. 14.24 (**a**) Predictions for etch depth vs. radiant exposure given by the blow-off ablation model for three different absorption coefficients and constant volumetric threshold energy density $W_{\mathrm{th}} = 2530$ J/cm^3 which is the sum of the sensible and latent heats for the vaporization of water. (**b**) Etch depth vs. radiant exposure data for ArF- and KrF-excimer laser ablation of cornea where $\mu_{\mathrm{a}} = 29000$ and 290 cm^{-1}, respectively. Data are compiled from [152] and [153]

threshold radiant exposure (blow-off model) or the ablation enthalpy (steady-state model). It is instructive to compare predictions of both models with respect to ablation efficiency. The ablation efficiency η_{abl} [kg/J] is defined as the amount of mass removed per unit energy delivered to the tissue and given by:

$$\eta_{\text{abl}} = \frac{\rho\delta}{H_0} \tag{14.14}$$

Substituting Eqs. (12) or (13) into Eq. (14) provides expressions for the ablation efficiency predicted by the blow-off and steady-state models, respectively:

$$\eta_{\text{abl}} = \frac{\rho}{\mu_a H_0} \ln\left(\frac{H_0}{H_{\text{th}}}\right) \text{(blow-off)} \tag{14.15}$$

$$\eta_{\text{abl}} = \frac{H_0 - H_{\text{th}}}{h_{\text{abl}} H_0} \text{(steady-state)} \tag{14.16}$$

For both blow-off and steady-state models, the ablation efficiency is zero at the threshold radiant exposure because a finite amount of energy is expended without any material removal. Also, in both models, a constant threshold energy density, $\mu_a H_{\text{th}}$, results in a maximum ablation efficiency that is independent of the optical absorption coefficient. In the blow-off model, the ablation efficiency reaches a maximum at an incident radiant exposure of approximately $2.7H_{\text{th}}$ and decreases monotonically at larger doses (see Fig. 33 in Ref. [1]). Thus for large radiant exposures much of the laser energy is poorly utilized by "overheating" superficial tissue layers far in excess of that required for their removal. In the steady state model, the ablation efficiency increases monotonically and asymptotically approaches a value $\eta_{\text{abl}} = 1/h_{\text{abl}}$ for $H_0 >> H_{\text{th}}$.

Hibst unified blow-off and steady state ablation models within a single framework to accommodate the effects of plume absorption on the ablation process [125]. This framework reduces to the steady-state model for cases where the ablation plume is transparent to the incident radiation and to the blow-off model for a steady-state process where the ablation plume possesses an absorption coefficient identical to the tissue. Defining $\gamma = (\mu_{a,p}/\mu_p)$ where $\mu_{a,p}$ and μ_a are the optical absorption coefficients of the plume and tissue, respectively, Hibst showed that the etch depth δ is given by:

$$\delta = \frac{1}{\mu_a \gamma} \ln\left(\gamma \frac{H_0}{H_{th}} - \gamma + 1\right) \tag{14.17}$$

Predictions equivalent to that of the steady-state and blow-off models are recovered for $\gamma \to 0$ and $\gamma \to 1$, respectively, and are shown for different values of γ in Fig. 14.25.

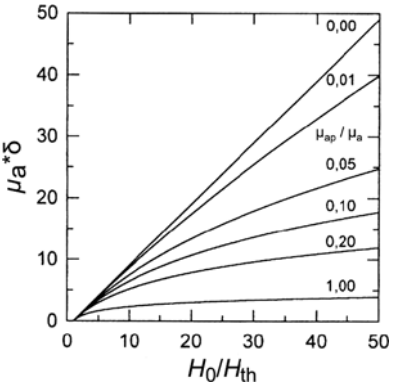


Fig. 14.25 Predictions for the variation of normalized etch depth with normalized radiant exposure predicted an ablation model that considers optical absorption associated by the ablation plume. Results are shown for various ratios for the optical absorption coefficient of the plume relative to the tissue. (Reprinted with permission from [125])

14.7.3 Mechanistic Models

The high degree of abstraction from physical ablation mechanisms that characterize heuristic models is a strength because it enables the description of different types of ablation processes as long as the model assumptions are fulfilled on a phenomenologic level. However, it is only when ablation mechanisms are explicitly considered that the origins of collateral effects can be identified and their spatial extent predicted.

Steady-state Vaporization Models. The first generation of mechanistic models developed in the 1980s and 1990s treated ablation as a rapid surface vaporization or boiling, considering thermal diffusion [43, 154–159]. The predictions of these models for threshold radiant exposure and material removal did not agree very well with experimental data because the models neglect explosive types of primary phase transitions, ignore the modification of all phase transitions by the tissue matrix, and they do not consider recoil-induced secondary material ejection. However, they provided estimates of the extent of thermal injury resulting from ablation and thus led to the discovery that thermal injury can be considerably reduced by using high laser irradiances that remove tissue at a speed faster than that associated with thermal diffusion [43, 156].

Thermo-mechanical Models. The first comprehensive attempt at integrating thermomechanical processes into a steady-state model of tissue ablation was made by Zweig, Frenz and co-workers starting in the late 1980s [56, 57, 160]. In this model, it is postulated that tissue can undergo two phase change processes: (a) a transition from a solid to liquid phase within the tissue bulk and (b) a transition from a liquid to vapor phase at the tissue surface. The solid–liquid phase transition is assumed to occur upon thermal denaturation, at which point the tissue loses mechanical integrity and "flows" under the action of shear stresses or pressure gradients. This allows for lateral pressure gradients in the vapor plume to eject denatured tissue "liquid" and thus provides material removal at ablation enthalpies lower than the vaporization enthalpy for water. Unfortunately, this model has three major deficiencies. First, experimental evidence shows that tissue is not liquified during ablation but ejected

in the form of fragments (Figs. 14.14 and 14.17). Second, the model does not consider phase explosions or confined boiling as driving the ablation process and thus neglects material ejection during the primary material removal process. Third, the mechanical component of ablation arises only as a by-product of pressure gradients in the ablation plume that are oriented *parallel* to the target surface, and the impact of recoil forces on material removal was neglected. The pressure gradients are assumed to originate from the radial variations of irradiance within a Gaussian beam that would result in radial variations of the vapor pressure in the plume. Zweig conceded that mechanical effects by the lateral vapor pressure gradients in the plume are relevant only for ablation using microsecond pulses and very small spot sizes because they "become insignificant if the liquid cannot transverse the irradiated spot during the laser pulse, as is typically the case for Q-switched pulses" [57]. However, he did not proceed to analyze the effects of the recoil pressure that provide a significant contribution to ablation, especially for nanosecond laser pulses [27].

A more realistic approach to describe the interplay between tissue matrix and phase transitions was taken by Majaron and co-workers who introduced the concept of confined boiling (see Section 14.4.5) [53]. Confined boiling was modeled by considering the thermodynamic behavior of tissue water when heated within an elastic tissue matrix. This was the first attempt at modeling explosive material removal as the primary process of tissue ablation. The model accurately predicts that explosive material removal can be initiated without supplying the entire vaporization enthalpy. Moreover it predicts, in agreement with experiment, that the threshold radiant exposure is weakly dependent on tissue stiffness but varies significantly with the ultimate tensile strength [150]. However, the model only examines the initiation of explosive material removal and does not consider the ablation process at radiant exposures larger than the ablation threshold. Furthermore, it does not consider phase transitions other than boiling.

Existing mechanistic models highlight only specific aspects of the ablation process but do not integrate the various phase transitions and secondary material removal processes involved to provide a comprehensive picture. The dynamics and interplay of both the phase transition processes and recoil induced effects depend on the spatio-temporal evolution of the volumetric energy density in the target as well as on the tissue properties. As a result of this complexity, the dependence between ablation depth and radiant exposure over a larger range of radiant exposures must be nonlinear. Thus any simple relationship between radiant exposure and ablation depth as predicted by heuristic and simple mechanistic models can only provide an approximate prediction over a limited range of radiant exposures. Within these limits, an assessment of the ablation rate and efficiency within the context of the heuristic blow-off or steady state models (Section 14.7.2), and an estimation of ablation thresholds and thermal damage zones based on mechanistic models is still valuable and remains in prevalent use. However, modeling of the sequence of continuous and explosive phase transitions in the primary ablation process and of recoil-induced material ejection will be required to obtain a mechanistically sound and reliable prediction of ablation depth and collateral damage. The development of such models thus represents an important challenge for future work.

14.7.4 Molecular Dynamics Simulations

With conceptual advances in modeling, as well as advances in computational power, molecular dynamics (MD) simulations are becoming a valuable approach for developing an understanding of ablation on the microscale [75, 97, 161–164]. Computational molecular dynamics offers predictions of the motion of molecular units produced by the deposition of laser energy. In these computations, macroscopic thermodynamic constraints are not imposed explicitly. Rather, ablation arises as a natural consequence of molecular interactions governed by the implementation of the equations of motion, intermolecular interactions, energy transfer, and molecular excitation within the model system.

The MD approach yields a wealth of information regarding the inception of phase transitions and the time-evolution of the size and velocity distribution of the ablation products that are difficult to obtain by other means. As an example, Fig. 14.26 provides the time evolution of a 2-dimensional system following the delivery of a 15 ps laser pulse with an incident radiant exposure two times the ablation threshold. These simulations reveal the volumetric nature of the ablative material removal process where small and large clusters leave the material bulk from locations not limited to those on the material surface. This is quite distinct from evaporative and desorption processes where material removal tends to occur strictly from the material surface

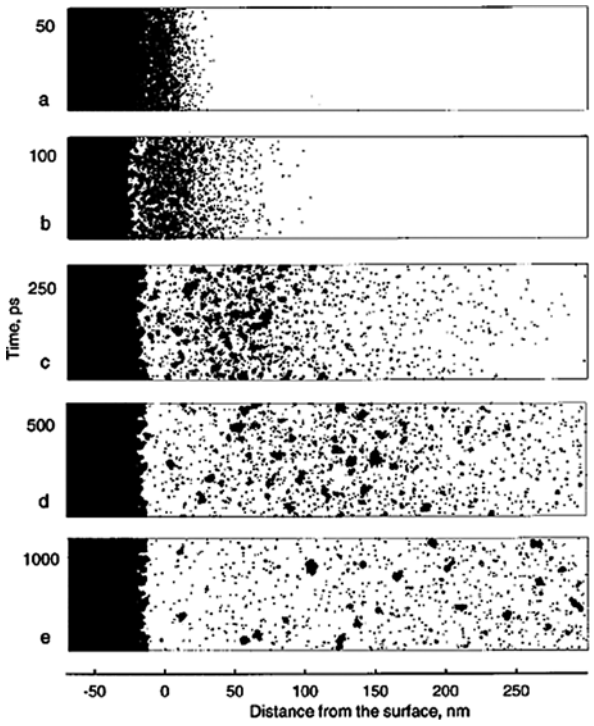


Fig. 14.26 Time development of an ablation plume for a 2D molecular dynamics simulation. The radiant exposure is roughly 2× the ablation threshold. The average energy deposited is 0.30 eV per particle in the irradiated region whereas the cohesive energy is 0.31 eV. The laser pulse width is 15 ps. (Reprinted with permission from [162])

and consists predominantly of single molecules. The MD approach has already provided valuable insights into the microscopic dynamics of desorption/ablation of both homogeneous and heterogeneous materials [165, 166] and is thus well suited to provide insights into the ablation dynamics of biological tissues. MD simulations are especially valuable for the analysis of stress-confined ablation processes because they are able to model the void formation and coalescence in heated and stretched materials [167, 168]. The main limitation of the MD approach are the very limited spatial and temporal scales that can be addressed.

14.8 UV and IR Ablation

The most common clinical application of pulsed tissue ablation is corneal refractive surgery using ArF excimer laser ($\lambda \approx 193$ nm) pulses of 20–30 ns duration. When ArF excimer laser pulses were first applied for the ablation of biological tissue, the smooth, precise, and nearly damage free character of the ablation [70, 152, 169–171] as shown in Fig. 14.27a was remarkable. This result seemed enigmatic in comparison to the much inferior ablation quality produced by hydrogen fluoride lasers ($\lambda = 3.0$ μm) [172, 173] or by Q-switched pulses from Er:YAG ($\lambda = 2.94$ μm) or Er:YSGG ($\lambda = 2.79$ μm) lasers [174], as shown in Fig. 14.27b because it appeared that the optical absorption coefficients at the two wavelength regions (190 nm vs. 2.8–3.0 μm) were comparable [170, 175]. Later this issue became even more perplexing when it was shown that a reduction in laser pulse duration from ≈ 50 ns to

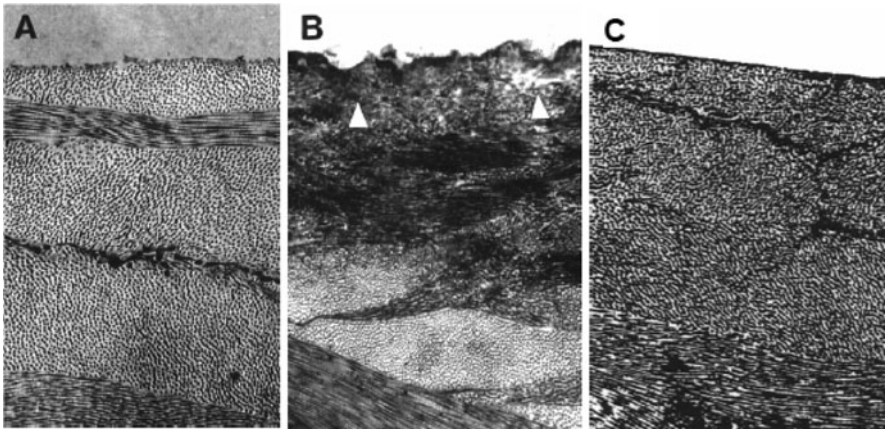


Fig. 14.27 Transmission electron micrographs of corneal ablation with (**a**) 193 nm ArF-excimer laser irradiation ($\lambda = 193$ nm), (**b**) Q-switched Er:YAG laser irradiation ($\lambda = 2940$ nm), and (**c**) 6-ns pulses from an optical parametric oscillator (OPO) emitting at 2.94 μm. Magnification 10000×. The damaged surface zone is visible as thin dark line in (**a**) and (**c**), and demarcated by arrowheads in (**b**). The damage zone has a thicknesses of ≤0.2 μm and in (**a**) and ≤0.3 μm (**c**), respectively, while it is 1–2 μm thick in (**b**) [176]

6 ns enabled 2.94 μm radiation to provide ablation quality similar to that of the ArF excimer laser (Fig. 14.27c) [176].

The common argument advanced to explain the differences between UV and IR ablation is that UV radiation induces photochemical processes which directly break chemical bonds, while IR laser ablation is governed solely by thermal processes. However, extensive investigations in the late 1980s and 1990s revealed that the actual picture is much more complex. As explained in Section 14.4.8, UV laser ablation is never an exclusively photochemical process, but involves an interplay of photochemical and photothermal processes. Moreover, for both UV and IR wavelengths, photomechanical mechanisms must be considered to fully understand the ablation dynamics and outcome. A detailed description of this picture and the underlying data are presented in the recent review article by Vogel and Venugopalan [1]; here, we will simply summarize the main findings.

Several studies using different measurement techniques [13, 153, 177] showed that the optical penetration depth of ArF excimer laser light is in the order of 0.25–0.5 μm rather than 3.7 μm as reported in the initial study [170]. In turn, mid-infrared spectroscopic studies of the temperature dependence of the optical absorption of water [10–12] revealed that the optical absorption of water drops significantly with temperature at both $\lambda = 2.94$ and 2.79 μm. While the optical absorption depth of water is 0.7 μm at room temperature, it exceeds 5 μm at radiant exposures $H_0 > 10\ \mathrm{J/cm^2}$ as shown in Fig. 14.2. Thus under ablative conditions the optical absorption coefficient of tissue at $\lambda = 193$ nm is actually about an order of magnitude larger than that in the $\lambda = 3$ μm region.

While the extremely small optical penetration depth explains the small etch depths and zones of thermal injury achieved by ArF excimer lasers, one must consider the role of photochemical effects, stress confinement, and heat transfer between tissue constituents to elucidate the differences in ablation dynamics and surface quality produced by UV and IR laser pulses of different durations. Table 14.1 provides a list of laser and tissue parameters relevant for this discussion. We use the ArF excimer laser ($\lambda = 193$ nm) and Q-switched Er:YSSG laser ($\lambda = 2.79$ μm) as examples for ablation without stress confinement, and the KrF excimer laser ($\lambda = 248$ nm) and an optical parametric oscillator (OPO) emitting radiation at $\lambda = 2.94$ μm as examples for ablation processes that are influenced by thermoelastic stresses. The table provides the pulse duration and tissue optical penetration depth for each laser source as well as the dimensionless pulse durations t_d^* and t_m^* that are normalized by the characteristic thermal diffusion and stress propagation times across the heated volume, respectively (see Section 14.3). Note that thermally confined conditions ($t_d^* \lesssim 1$) are achieved in all cases.

The volumetric energy density required for the onset of ablation is much smaller in cases with stress confinement, and increases as the degree of confinement decreases. When only thermal confinement is achieved, the threshold volumetric energy density for ablation resembles the energy density necessary to heat water from room temperature to the spinodal limit, which is $1270\ \mathrm{J/cm^3}$ for pure water and $890\ \mathrm{J/cm^3}$ for tissue with 70% water content. Thus it is considerably smaller than the vaporization enthalpy of water ($2255\ \mathrm{J/cm^3}$), which indicates the important role of material ejection for laser ablation.

Table 14.1 Wavelength λ, photon energy E_λ, pulse duration t_p, optical absorption depth μ_a^{-1}, dimensionless pulse durations t_m^* and t_d^* normalized by the stress propagation time and thermal diffusion time across the heated volume, respectively, volumetric energy density at ablation threshold W_{th}, and width of thermal damage zone d for IR and UV laser ablation of tissue. For comparison, the volumetric energy densities necessary to heat water from room temperature to the spinodal limit and to produce complete vaporization are 1270 and 25855 J/cm^3, respectively

Laser	λ [nm]	E_λ[eV]	t_p[ns]	μ_a^{-1} [μm]	$t_m^*(=\mu_a c_a t_p)$	$t_d^*\ (=\kappa\mu_a^2 t_p)$	H_{th}[J/cm^2]	S_{th}[J/cm^3]	d [μm]
OPO	2940	0.42	6	1.5	6	3.5×10^{-3}	0.035	230	≤ 0.3
Q-sw Er:YSGG	2790	0.44	40	3	20	5.8×10^{-4}	0.24	800	1−2 (Er:YAG)
KrF-excimer	248	5.0	34	30	1.2	3.5×10^{-6}	0.21	70	≈ 2
ArF-excimer	193	6.4	22	0.35	94	3.2×10^{-2}	0.050	1400	≤ 0.2

Data are compiled from [31] and [32], with the exception of the data for the thermal damage zones that are taken from [152] and [176].

With only thermal confinement, the thermal damage zone is similar to the optical penetration depth resulting in damage zones ≤ 0.2 μm for ArF laser irradiation and 1–2 μm for Q-switched Er:YSSG laser irradiation. However, when stress confinement is provided in addition to thermal confinement, the thermal damage zone is much smaller than the optical penetration depth. Specifically, the width of the damage zone is only 1/20 of the penetration depth of KrF laser irradiation ($t_m^* = 1.2$, relatively strong stress confinement) and 1/5 of the penetration depth of the IR-OPO irradiation ($t_m^* = 6$, weak stress confinement). The influence of thermoelastic stresses is the key factor that explains how ablation with IR-OPO radiation at $\lambda = 2.94$ μm achieves a damage zone that is almost as small as that produced by the ArF excimer laser even though its optical penetration depth is significantly larger.

We saw above that the precision of ablation achieved by different UV and IR lasers can be explained in terms of the optical penetration depth and thermal and stress confinement. However, to understand experimental observations that UV ablation starts earlier and progresses more smoothly than IR ablation, one must consider in addition the influence of the tissue matrix on phase transitions in the tissue water and the role of photochemical decomposition of this matrix [1, 31, 32]. The time necessary for the onset of material removal represents the timescale required for tissue decomposition and can be inferred from measurements of the acoustic transients generated by recoil of the ablation products. For Q-sw Er:YSGG laser ablation, these measurements indicate that radiant exposures close to the ablation threshold result in an explosive onset of material removal that occurs towards the end of the laser irradiation [32]. By contrast, in ArF-excimer laser ablation stress transients due to ablative recoil begin promptly after the start of the laser pulse [31, 81, 82, 130, 178], and material removal proceeds continuously during the irradiation [31]. This result is consistent with the finding that the ejection velocity of plume material produced by ArF excimer laser ablation ($v = 600$ m/s for cornea ablation using $H_0 = 0.3 = 1$ J/cm^2) is much smaller than that for Q-sw Er:YAG ablation at similar volumetric energy density within the target tissue ($v \approx 1400$ m/s for skin ablation using $H_0 \approx 25$ J/cm^2) [27, 101, 102, 126].

Both the different timing of the onset of material removal and the different ejection velocities can be traced back to different decomposition kinetics of the tissue matrix. For IR laser irradiation at radiant exposures similar to the ablation threshold, the mechanical integrity of the tissue matrix is essentially maintained during the laser pulse (Fig. 14.17b, c) and ablation occurs only after the tissue matrix has been ruptured via phase explosion or confined boiling. This scenario explains the explosive character of the ablation dynamics and the rough surface of the ablation crater produced by Q-switched Er:YSSG irradiation (Fig. 14.27b). By contrast, the more rapid onset of the recoil stress in UV ablation suggests that the tissue matrix is destroyed more rapidly than in IR ablation. Thermal denaturation cannot be used to explain the early degradation of the tissue matrix in UV ablation because a similarly rapid onset would then also be produced by IR exposures as the heat transfer between tissue water and collagen fibrils occurs on a nanosecond time scale (see Section 14.2 and Eq. (14.2)). Photochemical decomposition is much more likely to be responsible for rapid degradation of the tissue matrix during UV laser

irradiation. However, this does not imply that photochemical decomposition is the only mechanism relevant for UV ablation. In ArF-excimer laser ablation, the peptide bond linking the amino acids serves as the primary absorber. Several authors have shown that although a considerable fraction of the chemical bonds present in the tissue matrix is affected by photochemical decomposition [89, 94, 178] this fraction is insufficient to drive ablation by purely photochemical dissociation [92, 93]. To obtain a complete picture of all driving forces, one needs to include the phase transition of the tissue water. Thermal diffusion from the collagen heats the water to temperatures >> 100°C on a nanosecond time scale, which in turn promotes direct absorption of 193-nm laser light by water at elevated temperatures [1, 13, 179]. Thus ArF-excimer laser ablation is likely driven by a combination of photochemical and photothermal mechanisms. Photochemical effects are primarily responsible for tissue matrix decomposition and weakening that greatly reduces the explosive character of the ablation process while the photothermal processses are involved in the ejection of the fragments.

ArF-excimer laser ablation of tissue is less violent than IR laser ablation of either tissue or pure water which does not have a structural matrix. In IR laser ablation of water, the material ejection is driven by a phase explosion of superheated water or a vapor explosion of supercritical water that occurs mainly towards the end of the laser pulse [27]. By contrast, in ArF-excimer laser ablation, it is likely that the volatile gaseous products of the photodecomposition process produced during the heating of the tissue water serve to avert a phase explosion. A variety of organic gases including NH_2, C_2H_2, CH_2NH_2, and CO_2 in addition to water vapor have been found in the plume produced by 193-nm and 248-nm ablation of corneal tissue. [180] The volatile photodecomposition products provide a large number of boiling nuclei that promote a vaporization process and lower the degree of superheating in a way that prevents the spinodal limit from being reached. Analogous to a confined boiling process (Section 14.4.5), ablation will probably occur when the pressure from the volatile dissociation products, together with the vapor pressure from the heated tissue water, exceeds the tensile strength of the tissue matrix. However, since the strength of the matrix is strongly reduced by photochemical processes, the ablation process is less explosive than in IR ablation driven by a phase explosion.

It is interesting to note that the volumetric energy density at the ablation threshold is higher when using the ArF-excimer laser as compared to the Er:YSGG laser (Table 14.1). While the photochemical contribution to UV ablation changes the dynamics of tissue ablation, it does not reduce the ablation threshold such as predicted for photochemically enhanced ablation of organic solids [97]. The high ablation threshold correlates with findings that a large fraction of the tissue matrix is transformed into gaseous products [181] and that particulate ejecta of ArF laser ablation plumes [102, 105, 126] are much smaller than those produced by Q-switched IR laser tissue ablation close to the ablation threshold [27]. The photochemical decomposition of tissue into gaseous products and small particles requires more energy than the partial vaporization of tissue water and ejection of large tissue fragments produced by a phase explosion. By contrast, when photochemical weakening of

the tissue matrix arises in conjunction with the generation of thermoelastic tensile stresses (for example in KrF excimer laser ablation), ablation is associated with the ejection of large particles [126], and the volumetric energy density required for ablation is dramatically lowered (Table 14.1).

14.9 Ablation in a Liquid Environment

Ablation in a liquid environment is most often encountered when delivering laser radiation through optical fibers for medical applications inside the human body. Current technology places significant limits on the use of optical fibers to transmit the laser wavelengths most suitable for precise tissue ablation in air: $\lambda = 193$ nm (ArF-excimer), 248 nm (KrF-excimer), 2.79 μm (Er:YSSG), 2.94 μm (Er:YAG), and 10.6 μm (CO_2). Therefore, wavelengths that are well transmitted through low-OH quartz fibers e.g., $\lambda = 308$ nm (XeF-excimer), 2.01 μm (Cr:Tm:YAG), and 2.12 μm (Cr:Tm:Ho:YAG), are most often used for ablation in liquid environments. While these wavelengths are not as highly absorbed by protein or water, (see Fig. 14.2a), their optical penetration depth is still fairly small: ≈50 μm for $\lambda = 308$ nm [182, 183], ≈170 μm for $\lambda = 2.01$ μm (thulium laser), and ≈350 μm for $\lambda = 2.12$ μm (holmium laser), respectively [184].

In a liquid environment, a layer of aqueous fluid is usually present between the tip of the optical fiber and the tissue surface. This fluid absorbs all the IR laser wavelengths listed above and thus needs to be ablated or displaced before the radiation can reach the tissue surface and tissue ablation can start. A similar problem is encountered during UV ablation if the fluid is blood, due to UV absorption by hemoglobin (see Fig. 14.2a). However, once the radiation is delivered to the tissue, the kinetics of the phase transitions in the liquid environment are similar to those in a gaseous environment, with the exception that surface vaporization plays no role. As a result, ablation thresholds, as measured by the deposited volumetric energy density are almost the same in liquid environments as those observed in air [67, 130].

The most characteristic feature of ablation in a liquid environment is the confinement of the ablation products by the liquid. Therefore, ablation in a liquid environment is accompanied by bubble formation and much stronger mechanical effects than observed in a gaseous environment [1]. The expansion of gaseous products produced during tissue ablation creates a bubble in the liquid surrounding the ablation site. When the optical fiber is not in contact with tissue, a bubble is also formed by absorption of laser radiation in the liquid separating the fiber tip and the tissue surface, as shown in Fig. 14.28a. This bubble is essential for the transmission of optical energy to the target, and the mechanisms governing its formation and subsequent dynamics have received much attention by researchers [182, 185–192].

Experiments have demonstrated bubble formation at the tip of an optical fiber at volumetric energy densities smaller than the vaporization enthalpy of water at constant pressure [186, 187, 193]. This observation of "partial vaporization" has puzzled some researchers but can be easily explained by examining the kinetics

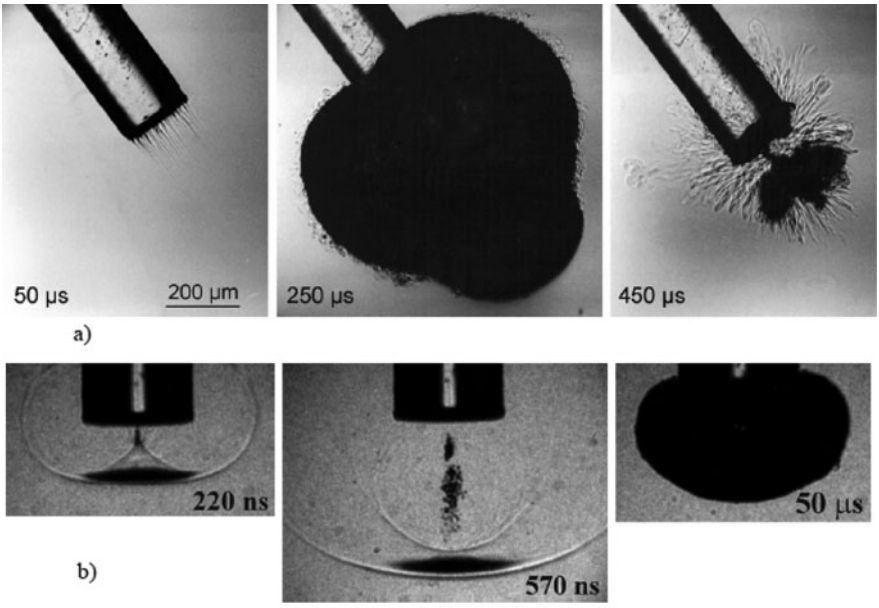


Fig. 14.28 Bubble formation in water without and with stress confinement (**a**) Bubble produced by a thulium laser pulse ($\lambda = 2.0$ µm) of 500 mJ pulse energy and 300 µs duration (FWHM) applied through a 200-µm diameter optical fiber. Thermal Schlieren at the fiber tip mark the optical penetration depth. Schlieren around the expanded and collapsed bubble indicate that the liquid next to the bubble remains hot for some time. (**b**) Bubble formation under stress confinement conditions [38]. 6-ns pulses of 490 nm wavelength were transmitted through a 200-µm optical fiber submerged in an aqueous solution of Orange G dye with an optical penetration depth of 11 µm. Tensile stress waves originating from the stress release ("diffraction") at the edges of the fiber produce cavitation bubbles at the fiber axis where they overlap. After ≈ 5 µs, a vapor bubble appears directly at the fiber surface that reaches its maximum size at 50 µs. (Reprinted with permission from [38])

of phase transitions. When the rate of energy deposition is low and heterogeneous vaporization nuclei are present, the phase transition is governed by normal boiling and starts once a temperature in excess of the equilibrium saturation temperature (100°C at atmospheric pressure) has been reached that is sufficient to overcome the surface tension of the vaporization nuclei (Fig. 14.8). Due to the large specific volume of water vapor, bubbles of considerable size are formed even when only a small fraction of the water within the optical penetration depth is vaporized. When the rate of energy deposition is high and the stress confinement condition is not fulfilled, a phase explosion occurs once the spinodal limit is reached [185, 187]. The separation of the metastable liquid into saturated vapor and saturated liquid corresponds again to a partial vaporization. The fraction vaporized is below 50% and depends on the volumetric energy density as described by the theory of metastable liquids [45, 194, 195] (Section 14.4.4). For intermediate rates of energy deposition in the presence of heterogeneous nuclei, the phase transition occurs as explosive boiling. At very large

radiant exposures and nanosecond pulse durations, supercritical conditions are produced and the heated volume is completely transformed to vapor as it expands to pressures below the critical point (Section 14.4.6).

With stress confinement, a thermoelastic stress wave is launched within a liquid volume directly adjacent to the fiber tip, and tensile stress waves are created due to the finite spatial extent of the absorbing liquid volume (Section 14.3). The tensile stress distribution can be described as a consequence of diffraction of the thermoelastic stress wave at the circular boundary of the acoustic source in front of the fiber tip. The ratio of the optical penetration depth ($1/\mu_a$) of the laser radiation to the diameter d of the fiber tip determines whether the cavitation-induced bubble formation occurs inside or outside the volume of energy deposition, which is the case for $\mu_a d \approx 1$, and $\mu_a d << 1$, respectively [38, 196]. An example of cavitation-induced bubble formation outside the volume of energy deposition is shown in Fig. 14.28b. Using nanosecond pulses delivered though a 400 μm fiber into a medium with an optical penetration depth of 380 μm, cavitation could be created by a temperature rise as small as 33°C [195]. Similar absorption and stress confinement conditions are achieved with a Q-switched holmium laser in aqueous (water or blood) environments. A certain degree of stress confinement is also present during the individual intensity spikes of free-running holmium laser pulses. The tensile portion of the thermoelastic stress transients produced by these spikes leads to a transient lowering of the boiling temperature and to bubble formation from heterogeneous nuclei at temperatures well below 100°C [187].

The dynamics of bubble formation and growth depends on the number density and size of heterogeneous nuclei as well as on the specific path taken in the phase diagram. Therefore, no general temperature-threshold for bubble formation exists. The conditions leading to bubble formation must be determined in each individual case by establishing the dynamics of temperature and pressure produced by the laser irradiation.

The expansion of the hot ablation products is always inhibited in a liquid environment, and is confined further in cases where the fiber tip is placed in direct contact with the tissue surface. Thus for a given radiant exposure, higher temperatures and pressures can build up within the target as compared to ablation in a gaseous environment. Due to the higher pressures, liquid environments offer a more effective transduction of the laser energy into mechanical energy [197, 198]. As a result, the potential for mechanical collateral damage is much larger for ablation in liquid environments as compared to air [197, 199–206].

The expansion of the heated liquid volume produces a compressive stress transient followed by bubble formation. When the laser-induced stress transient possesses a sufficiently short rise time, its propagation results in the formation of a shock wave [206]. The large pressure in the laser-induced vapor bubble leads to a very rapid expansion that overshoots the equilibrium state where the internal bubble pressure equals the hydrostatic pressure. The increasing difference between the hydrostatic pressure and the falling internal bubble pressure then decelerates the expansion and brings it to a halt. At this point, the kinetic energy of the liquid during bubble expansion has been transformed into the potential energy of the expanded

bubble. The bubble energy E_B [J] is related to the radius of the bubble at its maximum expansion, $R_{\max}$, and the difference between the hydrostatic pressure p_0 [Pa] and the vapor pressure p_v [Pa] inside the bubble by [207]

$$E_B = \frac{4\pi}{3}(p_0 - p_v)R_{\max}^3 \tag{14.18}$$

The expanded bubble collapses again due to the static background fluid pressure. The collapse compresses the bubble content into a very small volume, thus generating a very high pressure that can exceed 1 GPa for an approximately spherical bubble collapse [208]. The rebound of the compressed bubble interior leads again to the emission of a strong pressure transient into the surrounding liquid that can evolve into a shock wave [203, 204].

While the events during bubble generation are influenced strongly by the laser parameters, the subsequent bubble dynamics are influenced primarily by the properties of the fluid and the boundary conditions in the neighborhood of the laser focus. A spherical bubble produced in an unconfined liquid retains its spherical shape while oscillating, and the bubble collapse takes place at the site of bubble formation. However, when the bubble is formed near a material boundary, the collapse is asymmetric and associated with the formation of high-speed water jets that concentrate the bubble energy at some distance from the locus of bubble generation [201, 208]. When the bubble collapses in the vicinity of a rigid boundary, the jet is directed towards this boundary. The bubble collapse near an elastic, tissue-like boundary is characterized by the formation of two liquid jets that are directed both away from and towards the boundary and reach velocities as high as 960 m/s [208]. Jet formation is also induced by the fiber tip itself whereby the jet is usually directed away from the tip in the direction of the fiber axis [188, 209]. The jets have been shown to be responsible for collateral damage [209] and to increase the material removal [201, 202, 210].

The fraction of laser energy converted into the mechanical energy of the cavitation bubble depends on the laser pulse duration [191, 211]. Bubbles produced by free-running laser pulses with durations of several hundred microseconds expand during the laser pulse, and vaporization and condensation occur simultaneously in different bubble regions. Material is ablated at the bubble wall opposite to the fiber tip while vapor condenses in those regions of the bubble that are not exposed to the laser radiation [191, 211]. The bubble formation during the laser pulse limits the maximum energy density in the ablated target. The energy flow and condensation within the bubble further limits the maximum bubble size. By contrast, when nanosecond pulses are used for ablation, the bubble formation starts only towards the end of the laser irradiation [186, 187]. This corresponds to extremely high volumetric energy densities and temperatures adjacent to the fiber tip at radiant exposures well above the ablation threshold. As a result, the cavitation bubble expands to a much larger size than that produced by a free-running pulse with equal energy. Thus, cavitation can lead to structural deformation of the tissue adjacent to the ablation site that is much more pronounced than the direct ablative effect. These

mechanical effects are sometimes desired because they enhance the cutting speed in liquid environments. However, the mechanical damage often compromises the high precision of the original ablation effect.

Cavitation-induced tissue deformation has had dramatic consequences in laser angioplasty [182]. The bubble dynamics produced ruptures within the vessel wall that often initiated an overactive healing response leading to restenosis [199]. Placement of an optical fiber tip in contact with the tissue further enhances the level of inertial confinement and can result in violent material ejection and tissue dissection [212–214]. Furthermore, the confinement of the ablation products by the ablation channel leads to an increase of collateral thermal damage because the heat within the ablation products is conducted back into the residual tissue [213, 215, 216].

Thermomechanical tissue injury can be significantly reduced when a tailored series of short pulses is delivered to the target rather than single pulses [1, 189, 190, 192]. When the pulse duration is sufficiently short to provide both stress and thermal confinement, the lowering of the ablation enthalpy in the stress confinement regime reduces the residual heat in the tissue [192, 211]. A reduction of cavitation effects is achieved by applying a pre-pulse with small energy followed by one or several ablation pulses with larger energy that are separated by time intervals of 50–100 μs. The pre-pulse produces a small cavitation bubble that is then filled by the ablation

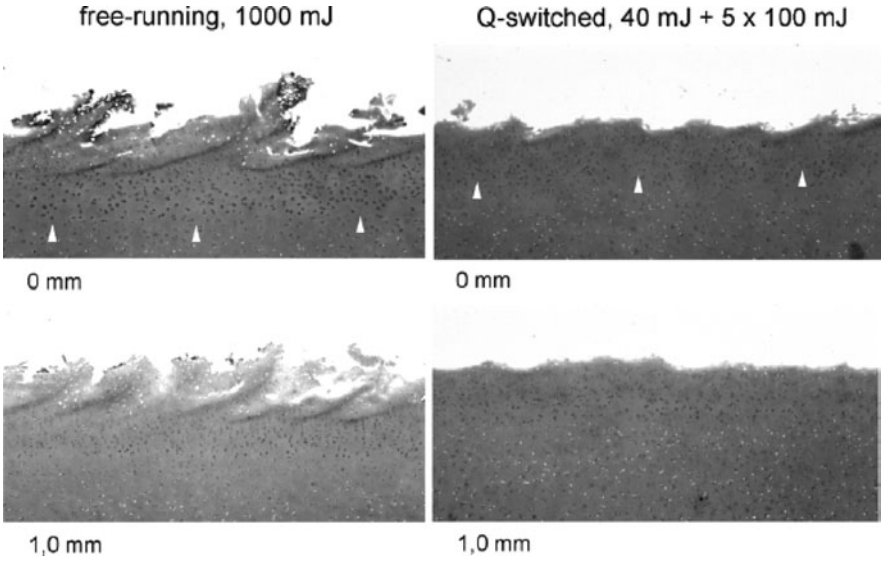


Fig. 14.29 Histological sections through ablated areas of cartilage that was irradiated with free-running thulium laser pulses (1000 mJ), and bursts of multiply Q-switched thulium-pulses (40 mJ + 5×100 mJ). The fiber was displaced by 0.4 mm between subsequent pulses. The distance between fiber tip and tissue surface was 0 mm (*top row*), and 1.0 mm (*bottom row*). The thermal damage zone (*arrows*) is much larger after free-running pulses, and the damage is more severe. Moreover, bursts of Q-switched pulses create smoother ablation crater walls than free-running pulses. Original magnification 20×

products produced by the subsequent pulses of higher energy. In this manner, no additional cavitation effects are produced, and tissue tearing and other mechanical side effects are minimized [189–192]. Moreover, the transient cavity created by the pre-pulse between fiber tip and tissue surface improves the optical transmission to the target which results in an increase of the ablation efficiency. The refinement of the tissue effects achieved by this technique is shown in Fig.14.29.

14.10 Control of Ablated Mass and Thermal and Mechanical Side Effects

In this section, we utilize predictions of ablation models as well as the qualitative insights formed in the previous sections to draw conclusions with regard to the control of ablated mass and thermal and mechanical side effects.

The heuristic ablation models described in Section 14.6.2 predict that, if the *removal of large amounts of material* is desired, the use of long laser pulses that achieve a steady-state-like ablation process will be more suitable than ablation based on a blow-off process. This arises because the ablated mass scales linearly with radiant exposure in a steady state ablation processes but logarithmically in a blow-off process (Eqs. (14.12) and (14.13)). However, the difference between the two types of ablation process becomes less pronounced if optical shielding by the ablation products is significant. A steady state process is most advantageous under conditions where the absorption of the incident laser beam by the ablation plume is markedly smaller than the absorption in the target tissue present in the blow-off situation. Very high ablation rates have been achieved using IR laser pulses with millisecond duration and used successfully for transmyocardial laser revascularization (TMR) where a deep hole is drilled across the full thickness of the myocardium during one fraction of the systolic cycle [214, 217].

When use of short laser pulses is required to minimize thermal side effects, the largest ablation efficiencies are achieved when the pulse duration is sufficiently short to guarantee not only thermal confinement but also stress confinement [67]. Stress confinement or inertial confinement conditions generally serve to lower the ablation threshold and increase the ablation efficiency (Section 14.4.7).

The most direct strategy to *control thermal side effects* involves the selection of a pulse duration that is sufficiently short to minimize heat diffusion during the laser pulse from the volume of energy deposition into the nonablated tissue (Eq. (14.3)). However, similar results may also be obtained using longer pulses if the velocity of the ablation front during the laser pulse is comparable or faster than the heat diffusion into the residual tissue. A theoretical analysis of this strategy was presented by Venugopalan and co-workers [43], and experimental evidence for its validity was presented by various authors [43, 142, 143, 218–220].

Thermal side effects can be diminished further by selecting laser pulse durations sufficiently short to provide both stress and thermal confinement because the lowering of the ablation enthalpy in the stress confinement regime reduces the residual

heat in the tissue [192, 211]. This is shown in Fig. 14.29 where the thermal damage produced by ablation using free-running thulium laser pulses is compared to the damage zone produced by ablation using Q-switched thulium laser pulses. The first irradiation modality provides merely thermal confinement while the second modality provides both stress and thermal confinement.

When delivering multiple pulses to a single location, both the spatial extent and degree of the thermal injury are influenced by heat accumulation. The pulse repetition rate must be sufficiently low to avoid progressive accumulation of the residual heat within the tissue [143, 221]. An alternative strategy to avoid heat accumulation is to scan the laser beam in order to lengthen the time interval between subsequent exposures at each location within the ablated area.

In other instances, the heat accumulation occurring when series of laser pulses are applied can be used intentionally to create a coagulation effect in conjunction with precise laser ablation. Since precise tissue ablation requires both a small optical penetration depth and a short pulse duration, hemostasis is often not produced at the tissue surface. Hemostasis can be achieved by applying a series of pulses at radiant exposures just below the ablation threshold at sufficiently high repetition rate [222].

Mechanical side effects: In principle, high precision can be achieved by selecting a laser wavelength featuring a very small optical penetration depth combined with a short pulse duration sufficient to provide thermal confinement, and, if possible, also stress confinement. However, the production of extensive recoil stresses should be avoided to maintain the precision provided by a small optical penetration depth combined with thermal confinement. Extensive recoil stresses may degrade the smoothness of the ablated surface and/or induce collateral mechanical damage; especially in friable tissues. This restriction imposes an upper limit for the incident radiant exposure and implies that the finest tissue effects can be achieved by working close to the ablation threshold.

The confinement provided by a liquid environment increases the risk of mechanical side effects compared to ablation in air. Strategies to minimize these effects have been discussed in Section 14.8. On the other hand, for dissection procedures where precision and smoothness of the cut is less important than cutting efficiency and speed, the generation of disruptive effects is advantageous. In these cases, the highest ablation efficiency for a given average laser power can be achieved using a series of short pulses with high peak power and radiant exposures well above the ablation threshold.

14.11 Outlook and Challenges

During the past years, a framework has been created to identify relevant mechanisms governing pulsed laser tissue ablation techniques and to facilitate a rational formulation of irradiation strategies for specific tasks. Nevertheless, much work remains to further improve the understanding of pulsed laser ablation processes. While the

detailed experimental study of the kinetics of tissue ablation combined with a careful consideration of the thermodynamics of phase transitions has provided a fairly complete qualitative understanding, predictive quantitative models are still absent to support it. Most current mechanistic models are limited to simple phase transitions such as surface vaporization and boiling. The challenge is to devise models that link material removal with both the complex phase change processes driving ablation (surface vaporization, phase explosion, vapor explosion, photodissociation, confined boiling etc.) as well as the dynamic tissue optical and mechanical properties that modulate the temperatures and pressures generated within the tissue prior to fracture. Moreover, once material removal is initiated, such models must consider that the recoil forces associated with ablation and the flow of the plume can both influence material removal.

Obviously, all these demands cannot be met by a simple comprehensive analytical model, and researchers must resort to computational approaches to fulfill these requirements. However, much information on dynamic material properties required for a faithful modeling is still missing. These data include dynamic optical properties (absorption and scattering coefficients), thermal properties (heat capacity, Grüneisen coefficient), and mechanical properties (elastic and shear modulus, ultimate tensile strength) that may depend on the magnitude and kinetics of the temperature, pressure, and strain rates achieved during the ablation process. This lack of data is now recognized to be even more important than was thought just a few years ago as recent experimental studies demonstrate that the temperatures and pressures involved in most ablation processes are more extreme than assumed previously [27]. Vogel and Venugopalan [1] provided a summary of the results of many of those studies, e.g. on surface temperature during ablation. We now know that pulsed laser tissue ablation is regularly associated with a temperature rise of hundreds to thousands K, recoil pressures of several hundred MPa, and strain rates on the order of 10^5–10^7 s^{-1}. Several of the above mentioned dynamic material properties are presently known only at the lower margin of this parameter space.

Molecular dynamics (MD) simulations are a powerful approach to gain insight into the pulsed laser ablation dynamics on a microscopic level with sub-nanosecond time resolution. However, to span the spatial and temporal scales usually encountered in realistic applications of tissue ablation, computational facilities are required that are currently still unavailable. To make such computations more tractable, MD simulations of laser ablation typically do not model the system on an atomic level but instead adopt a "breathing sphere" model where groups of atoms are treated as units [161, 164]. This simplification provides a realistic rate of vibrational relaxation of excited molecules, but only allows for an approximate treatment of photochemical reactions. However, it has enabled the calculation of the sequence of events initiated by instantaneous energy deposition on temporal and spatial scales as large as 2 ns and 1 μm, respectively, while maintaining high temporal (<5 ps) and spatial (<5 nm) resolution [166]. To simulate the more common case of laser energy deposition during the ablation process and incorporate plume dynamics, shielding, and recoil-induced material ejection, the MD simulations must be combined with continuum mechanical approaches or/and Monte Carlo simulations on a mesoscopic level.

The first steps in this direction have been already made [164] but there is still a long way to go to establish numerical tools that can treat ablation dynamics on microscopic, mesoscopic and macroscopic level and link the results to experimentally observable parameters.

References

1. Vogel A and Venugopalan V. Mechanisms of pulsed laser ablation of biological tissues. *Chem. Rev.*, 103:577–644 (2003).
2. Vogel A, Noack J, Hüttman, G, and Paltauf, G. Mechanisms of femtosecond laser nanosurgery of cells and tissues. *Appl. Phys. B*, 81:1015–1047 (2005).
3. Silver FH. *Biological materials: Structure, mechanical properties, and modeling of soft tissue.* New York University, New York and London, 229 pp (1987).
4. Nimni ME and Harkness RD. Molecular structure and functions of collagen. In: ME Nimni (ed) *Collagen. Vol. I Biochemistry*. CRC, Boca Raton, pp. 1–77 (1988).
5. Linsenmayer TF. Collagen. In: ED Hay (ed) *Cell biology of extracellular matrix.* Plenum, New York and London, , 2nd edition, pp. 7–44 (1991).
6. Vaezy S and, Clark JI. A quantitative analysis of transparence in the human sclera and cornea using Fourier methods. *J. Microsc.*, 163:85–94 (1991).
7. Nomura S, Hiltner A, Lando JB, and Baer E. Interaction of water with native collagen. *Biopolymers*, 16:231–246 (1977).
8. Cheong W-F. Summary of optical properties. In: AJ Welch, MJC van Gemert (eds) *Optical-thermal response of laser-irradiated tissue.* Plenum, New York, pp. 275–304 (1995).
9. Carp A, Prahl SA, and Venugopalan V. Radiative transport in the delta-P_1 approximation: Accuracy of fluence rate and optical penetration depth predictions in turbid semi-infinite media. *J. Biomed. Opt.*, 9:632–647 (2004).
10. Vodopyanov KL. Saturation studies of H_2O and HDO near 3400 cm^{-1} using intense picosecond laser pulses. *J. Chem. Phys*., 94:5389–5393 (1991).
11. Walsh JT and Cummings JP. Effect of the dynamic optical properties of water on midinfrared laser ablation. *Lasers Surg. Med.*, 15:295–305 (1994).
12. Shori R, Walston AA, Stafsudd OM, Fried D, and Walsh JT. Quantification and modeling of the dynamic changes in the absorption coefficient of water at 2.94 μm. *IEEE J. Sel. Top. Quant. Electron*., 7:959–970 (2002).
13. Staveteig PT and Walsh JT. Dynamic 193-nm optical properties of water. *Appl. Opt*., 35:3392–3403 (1996).
14. Yamada H. *Strength of biological materials*. Robert E. Krieger, Huntington, NY (1970).
15. Vogel HG. Influence of age, treatment with corticosteroids and strain rate on mechanical properties of rat skin. *Biochim. Biophys. Acta*, 286:79–83 (1972).
16. Haut RC. The effects of orientation and location on the strength of dorsal rat skin in high and low speed tensile failure experiments. *Trans. ASME Biomed. Eng.,* 111, 136–140 (1989).
17. Dombi GW, Haut RC, and Sullivan WG. Correlation of high-speed tensile strength with collagen content in control and lathyritic rat skin. *J. Surg. Res.*, 54: 21–28 (1993).
18. Stryer L. *Biochemistry*. WH Freeman and Company, San Francisco (1987).
19. Thomsen S. Pathological analysis of photothermal and photomechanical effects of laser-tissue interactions. *Photochem. Photobiol*., 53:825–835 (1991).
20. Allain JC, Le Lous M, Cohen-Solal L, Bazin S, and Maroteaux P. Isometric tensions developed during the hydrothermal swelling of rat skin. *Connect. Tissue Res*., 7:127–133 (1980).
21. Kampmeier J, Radt B, Birngruber R, and Brinkmann R. Thermal and biomechanical parameters of porcine cornea. *Cornea*, 19:355–363 (2000).
22. Asiyo-Vogel MN, Brinkmann R, Notbohm H, Eggers R, Lubatschowski H, Laqua H, and Vogel A. Histologic analysis of thermal effects of laser thermokeratoplasty and corneal

ablation using Sirius-red polarization microscopy. *J. Cataract Refract. Surg.*, 23:515–526 (1997).
23. Le Lous M, Flandin F, Herbage D, and Allain JC. Influence of collagen denaturation on the chemorheological properties of skin, assessed by differential scanning calorimentry and hydrothermal isometric tension measurement. *Biochim. Biophys. Acta*, 717:295–300 (1982).
24. Henriques FC. Studies of thermal injury. V. The predictability and the significance of thermally induced rate processes leading to irreversible epidermal injury. *Arch. Pathol.*, 43, 489–502 (1947).
25. Simankowskii .M, Mackanos MA, Irani AR, O'Connel-Rodwell CE, Contag CH, Schwettman HA, and Palanker DV. Cellular tolerance to pulsed hyperthermia. *Phys. Rev. E*, 74:011915, 1–7 (2006).
26. Harris DM, Fried D, Reinisch L, Bell T, Schlachter D, From L, and Burkart J. Eyelid resurfacing. *Lasers Surg. Med.*, 25:107–122 (1999).
27. Apitz I and Vogel A. Material ejection in nanosecond Er:YAG laser ablation of water, liver, and skin. *Appl. Phys. A*, 81:329–228 (2005).
28. Carslaw HS and Jaeger JC. *Conduction of heat in solids.* Oxford University, Oxford, 2nd edition (1959).
29. Anderson RR and Parrish JA. Selective photothermolysis – precise microsurgery by selective absorption of pulsed radiation. *Science*, 220:524–527 (1983).
30. Jacques S. Role of tissue optics and pulse duration on tissue effects during high-power laser irradiation. *Appl. Opt.*, 32:2447–2454 (1993).
31. Venugopalan V, Nishioka NS, and Mikic BB. The thermodynamic response of soft biological tissues to pulsed ultraviolet laser radiation. *Biophys. J.*, 69:1259–1271 (1995).
32. Venugopalan V, Nishioka NS, and Mikic BB. Thermodynamic response of soft biological tissues to pulsed infrared laser radiation. *Biophys. J.*, 70:2981–2933 (1996).
33. Paltauf G and Dyer PE. Photomechanical processes and effects in ablation. *Chem. Rev.*, 103:487–518 (2003).
34. Sigrist MW. Laser generation of sound waves in liquids and gases. *J. Appl. Phys.*, 60:R83–R121 (1986).
35. Bushnell JC and McCloskey DJ. Thermoelastic stress generation in solids. *J. Appl. Phys.*, 39:5541–5546 (1968).
36. Itzkan I, Albagli D, Dark ML, Perelman LT, von Rosenberg C, and Feld M. The thermoelastic basis of short pulsed laser ablation of biological tissue. *Proc. Natl. Acad. Sci. USA*, 92:1960–1964 (1995).
37. Paltauf G and Schmidt-Kloiber H. Microcavity dynamics during laser-induced spallation of liquids and gels. *Appl. Phys. A*, 62:303–311 (1996).
38. Frenz M, Paltauf, G, and Schmidt-Kloiber H. Laser-generated cavitation in absorbing liquid induced by acoustic diffraction. *Phys. Rev. Lett.*, 76:3546–3549 (1996).
39. Paltauf G and Schmidt-Kloiber H. Photoacoustic cavitation in spherical and cylindrical absorbers. *Appl. Phys. A*, 68:525–531 (1999).
40. Miotello A and Kelly R. Critical assessment of thermal models for laser sputtering at high fluences. *Appl. Phys. Lett.*, 67:3535–3537 (1995).
41. Schrage RW. *A theoretical study of interphase mass transfer*. Columbia University, New York (1953).
42. Yablon AD, Nishioka NS, Mikić, BB, and Venugopalan V. Physical mechanisms of pulsed infrared laser ablation of biological tissues. *Proc. SPIE*, 3343:69–77 (1998).
43. Venugopalan V, Nishioka NS, and Mikic BB. The effect of laser parameters on the zone of thermal injury produced by laser ablation of biological tissue. *Trans. ASME J. Biomech. Eng.*, 116:62–70 (1994).
44. Miotello A and Kelly R. Laser-induced phase explosion: New physical problems when a condensed phase approaches the thermodynamic critical temperature. *Appl. Phys. A*, 69:S67–S73 (1999).
45. Debenedetti P. *Metastable liquids: Concepts and principles.* Princeton University, Princeton, NJ (1996).
46. Skripov VP. *Metastable liquids.* Wiley, New York (1974).

47. Skripov VP, Sinitsyn, EN, Pavlov, PA, Ermakov, GV, Muratov, GN, Bulanov, NV, and Baidakov G. *Thermophysical properties of liquids in the metastable (superheated) state*. Gordon and Breach Science, New York (1988).
48. Apfel RE. Water superheated to 279.5°C at atmospheric pressure. *Nature Phys. Sci.*, 238:63–64 (1972).
49. Martynyuk MM. Vaporization and boiling of liquid metal in an exploding wire. *Sov. Phys. Tech. Phys.*, 19:793–797 (1974).
50. Martynyuk MM. Phase explosion of a metastable fluid. *Combust. Explos. Shock Waves*, 13:178–191 (1977).
51. Frenkel II. *Kinetic theory of liquids.* Dover, New York (1955).
52. Kiselev SB. Kinetic boundary of metastable states in superheated and stretched liquids. *Physica A*, 269:252–268 (1999).
53. Majaron B, Plestenjak P, and Lukac M. Thermo-mechanical laser ablation of soft biological tissue: Modelling the micro-explosions. *Appl. Phys. B*, 69:71–80 (1999).
54. Verdaasdonk RM, Borst C, and van Germert MJC. Explosive onset of continuous wave laser tissue ablation. *Phys. Med. Biol.*, 35:1129–1144 (1990).
55. LeCarpentier GL, Motamedi M, McMath LP, Rastegar S, and Welch AJ. Continuous wave laser ablation of tissue: Analysis of thermal and mechanical events. *IEEE Trans. Biomed. Eng.*, 40:188–200 (1993).
56. Frenz M, Romano V, Zweig AD, Weber HP, Chapliev NI, and Silenok AV. Instabilities in laser cutting of soft media. *J. Appl. Phys.*, 66:4496–4503 (1989).
57. Zweig AD. A thermo-mechanical model for laser ablation. *J. Appl. Phys.*, 70:1684–1691 (1991).
58. Lu Q. Thermodynamic evolution of phase explosion during high-power nanosecond laser ablation. *Phys. Rev. E*, 67:016410 (5 pages) (2003).
59. Dingus RS and Scammon RJ. Grüneisen-stress induced ablation of biological tissue. *Proc. SPIE*, 1427:45–54 (1991).
60. Dingus RS, Curran DR, Oraevsky AA, and Jacques SL. Microscopic spallation process and its potential role in laser-tissue ablation. *Proc. SPIE*, 2134A:434–445 (1994).
61. Paltauf G and Schmidt-Kloiber H. Model study to investigate the contribution of spallation to pulsed laser ablation of tissue. *Lasers Surg. Med.*, 16:277–287 (1995).
62. Oraevsky AA, Jacques SL, Esenaliev RO, and Tittel FK. Pulsed laser ablation of soft tissue, gels, and aqueous solutions at temperatures below 100°C. *Lasers Surg. Med.*, 18:231–240 (1996).
63. Duvall GE and Fowles GR. Shock waves. In: RS Bradley (ed) *High pressure physics and chemistry.* Academic, New York, pp. 209–291 (1963).
64. Zel'dovich YB and Raizer YP. *Physics of shock waves and high temperature hydrodynamic phenomena*, Vol. I and II. Academic, New York and London (1966).
65. Lu Q, Mao SS, Mao X, and Russo RE. Delayed phase explosion during high-power nanosecond laser ablation of silicon. *Appl. Phys. Lett.*, 80:3072–3074 (2002).
66. Vogel A and Noack J. Shock wave energy and acoustic energy dissipation after laser-induced breakdown. *Proc. SPIE*, 3254:180–189 (1998).
67. Albagli D, Perelman LT, Janes GS, von Rosenberg C, Itzkan I, and Feld M. Inertially confined ablation of biological tissue. *Lasers Life Sci.*, 6:55–68 (1994).
68. Kawamura Y, Toyoda K, and Namba S. Effective deep ultraviolet photoetching of polymethyl methacrylate by an excimer laser. *Appl. Phys. Lett.*, 40:374–375 (1982).
69. Srinivasan R and Mayne-Barton V. Self-developing photoetching of poly(ethylene terephthalate) films by far-ultraviolt excimer laser radiation. *Appl. Phys. Lett.*, 41:576–578 (1982).
70. Trokel SL, Srinivasan R, and Braren B. Excimer laser surgery of the cornea. *Am J. Ophthalmol.*, 96:710–715 (1983).
71. Srinivasan R and Leigh W. Ablative photodecomposition. Action of far-ultraviolet (193 nm) laser radiation on poly(ethylene terephthalate) films. *Am. J. Chem. Soc.*, 104:6784–6785 (1982).

72. Srinivasan R. Ablation of polymers and biological tissues using ultraviolet lasers. *Science*, 234:559–565 (1986).
73. Jellinek HHG and Srinivasan R. Theory of etching polymers by far-ultraviolet, high-intensity pulsed laser and long-term radiation. *J. Phys. Chem.*, 88:3048–3051 (1984).
74. Palmer BJ, Keyes T, Clarke RH, and Isner JM. Theoretical study of ablative photodecomposition in polymeric solids. *J. Phys. Chem*., 93:7509–7516 (1989).
75. Garrison BJ and Srinivasan R. Laser ablation of organic polymers: Microscopic models for photochemical and thermal processes. *J. Appl. Phys.*, 57:2909–2914 (1985).
76. Srinivasan R and Braren B. Ultraviolet ablation of organic polymers. *Chem. Rev.*, 89:1303–1316 (1989).
77. Yeh C. Laser ablation of polymers. *J. Vac. Sci. Technol.*, A4:653–658 (1986).
78. Isner JM and Clarke RH. The paradox of thermal ablation without thermal injury. *Lasers Med. Sci.*, 2:165–173 (1987).
79. Wayne CE and Wayne RP. *Photochemistry.* Oxford University, Oxford (1996).
80. Ashfold MNR and Cook PA. Photochemistry with VUV photons. In:RA Meyers (ed) *Encyclopedia of physical science and technology*. Academic, 3rd edition (2002).
81. Davis GM, Gower MC, Fotakis C. Efthimiopoulos T, and Argyrakis P. Spectroscopic studies of ArF laser photoablation of PMMA. *Appl. Phys. A*, 36:27–30 (1985).
82. Dyer PE and Sidhu J. Excimer laser ablation and thermal coupling efficiency to polymer films. *J. Appl. Phys.*, 57:1420–1422 (1985).
83. Gorodetsky G, Kazyaka T G, Melcher RL, and Srinivasan R. Calorimetric and acoustic study of ultraviolet laser ablation of polymers. *Appl. Phys. Lett.*, 46:828–830 (1985).
84. Brunco DP, Thompson MO, Otis CE, and Goodwin PM. Temperature measurements of polyimide during KrF excimer laser ablation. *J. Appl. Phys*., 72:4344–4350 (1992).
85. Otis CE, Braren B, Thompson MO, Brunco D, and Goodwin PM. Mechanisms of excimer laser ablation of strongly absorbing systems. *Proc. SPIE*, 1856:132–142 (1993).
86. Sutcliffe E and Srinivasan R. Dynamics of UV laser ablation of organic polymer surfaces. *J. Appl. Phys.*, 60:3315–3322 (1986).
87. Srinivasan R. Ablation of polyimide (Kapton^{TM}) films by pulsed (ns) ultraviolet and infrared (9.17 μm) laser. *Appl. Phys. A*, 56:417–423 (1993).
88. Lippert T and Dickinson JT. Chemical and spectroscopic aspects of polymer ablation: Special features and novel directions. *Chem. Rev.*, 103:453–485 (2003).
89. Pettit GH. The physics of ultraviolet laser ablation. In: RW Waynant (ed) *Lasers in medicine*. CRC, Boca Raton, pp. 109–133 (2002).
90. Andrew JE, Dyer PE, Forster D, and Key PH. Direct etching of polymeric materials using a XeCl laser. *Appl. Phys. Lett*., 43:717–719 (1983).
91. Küper S, Madaressi S, and Stuke M. Photofragmentation pathways of a PMMA model compound under UV excimer laser ablation conditions. *J. Phys. Chem.*, 94:7514–7518 (1990).
92. Kitai MS, Popkov, VL, Semchishen VA, and Kharizov AA. The physics of UV laser cornea ablation. *IEEE J. Quant. Electron.*, 27:302–307 (1991).
93. Arnold N and Bityurin N. Model for laser-induced thermal degradation and ablation of polymers. *Appl. Phys. A*, 68:615–625 (1999).
94. Nikogosyan DN and Görner H. Laser-induced photodecomposition of amino-acids and peptides: Extrapolation to corneal collagen. *IEEE J. Sel. Top. Quant. Electron*., 5:1107–1115 (1999).
95. Bityurin N, Luk'yanchuk BS, Hong MH, and Chong TC. Models for laser ablation of polymers. *Chem Rev.*, 103:519–552 (2003).
96. Schmidt H, Ihlemann J, Wolff-Rottke B, Luther K, and Troe J. Ultraviolet laser ablation of polymers: Spot size, pulse duration, and plume attenuation effects explained. *J. Appl. Phys.*, 83:5458–5468 (1998).
97. Yingling YG, Zhigilei LV, and Garrison BJ. The role of photochemical fragmentation in laser ablation: A molecular dynamics study. *J. Photochem. Photobiol*., 145:173–181 (2001).

98. Oraevsky AA, Jacques SL, Pettit GH, Saidi IS, Tittel FK, and Henry PD. XeCl laser ablation of atherosclerotic aorta: Optical properties and energy pathways. *Lasers Surg. Med.*, 12:585–597 (1992).
99. Vogel A, Apitz I, Freidank S, and Dijkink R. Sensitive high-resolution white-light Schlieren technique with large dynamic range for the investigation of ablation dynamics. *Opt. Lett*., 31:1812–1814 (2006).
100. Hecht E and Zajac A. *Optics*. Addison Wesley, Reading, MA (1977).
101. Walsh JT and Deutsch TF. Measurement of Er:YAG laser ablation plume dynamics. *Appl. Phys. B*, 52:217–224 (1991).
102. Bor Z, Hopp B, Rácz B, Szabó G, Ratkay I, Süveges I, Füst A, and Mohay J. Plume emission, shock wave and surface wave formation during excimer laser ablation of the cornea. *Refract. Corneal Surg. (Suppl.)*, 9: S111–S115 (1993).
103. Cummings JP and Walsh JT. Q-switched ablation of tissue: Plume dynamics and the effect of tissue mechanical properties. *Proc. SPIE*, 1646:242–253 (1992).
104. Krueger RR and Trokel SL. Quantitation of corneal ablation by ultraviolet laser light. *Arch. Ophthalmol.*, 103:1741–1742 (1985).
105. Noack J, Tönnies R, Hohla C, Birngruber R, and Vogel A. Influence of ablation plume dynamics on the formation of central islands in excimer laser photorefractive keratectomy. *Ophthalmology*, 104:823–830 (1997).
106. Arnold N, Gruber J, and Heitz J. Spherical expansion of the vapor plume into ambient gas: An analytical model. *Appl. Phys. A*, 69:S87–S93 (1999).
107. Brode HL. Blast wave from a spherical charge. *Phys. Fluids*, 2:217–229 (1959).
108. Taylor G. The formation of a blast wave by a very intense explosion. I Theoretical discussion. *Proc. Roy. Soc. A*, 201:159–186 (1950).
109. Sedov LI. *Similarity and dimensional methods in mechanics*. Academic, New York (1959).
110. Landau LD and Lifschitz EM. *Fluid mechanics*. Pergamon, Oxford, 2nd edition (1987).
111. Sakurai A. On the propagation and structure of a blast wave, I. *J. Phys. Soc. Japan*, 8:662–671 (1953).
112. Sakurai A. On the propagation and structure of a blast wave, II. *J. Phys. Soc. Japan*, 9:256–266 (1954).
113. Freiwald DA and Axford RA. Approximate spherical blast theory including source mass. *J. Appl. Phys.*, 46:1171–1174 (1975).
114. Kelly R and Braren B. On the direct observation of the gas dynamics of laser-pulse sputtering of polymers. Part I: Analytical considerations. *Appl. Phys. B*, 53:160–169 (1991).
115. Kelly R and Miotello A. Pulsed-laser sputtering of atoms and molecules. Part I: Basic solutions for gas-dynamic effects. *Appl. Phys. B*, 57:145–158 (1993).
116. Dyer PE and Sidhu J. Spectroscopic and fast photographic studies of excimer laser polymer ablation. *J. Appl. Phys.*, 64:4657–4663 (1988).
117. Jones DL. Intermediate strength blast wave. *Phys. Fluids*, 11:1664–1667 (1968).
118. Stauter C, Gérard P, Fontaine J, and Engel T. Laser ablation acoustical monitoring. *Appl. Surf. Sci.*, 109/110:174–178 (1997).
119. Freiwald DA. Approximate blast wave theory and experimental data for shock trajectories in linear explosive-driven shock tubes. *J. Appl. Phys.*, 43:2224–2226 (1972).
120. Chen Z, Bogaerts A, and Vertes A. Phase explosion in atmospheric pressure infrared laser balation from water-rich targets. *Appl. Phys. Lett.*, 89:041503 (2006).
121. Nahen K and Vogel A. Shielding by the ablation plume during Er:YAG Laser ablation. *J. Biomed. Opt.*, 7:165–178 (2002).
122. Nahen K and Vogel A. Investigations on acoustic on-line monitoring of IR laser ablation of burned skin. *Lasers Surg. Med.*, 25:69–78 (1999).
123. Nahen K and Vogel A. Acoustic signal characteristics during IR laser ablation, and their consequences for acoustic tissue discrimination. *Proc. SPIE*, 3914:166–176 (2000).
124. Vogel A, Kersten B, and Apitz I. Material ejection in free-running Er:YAG laser ablation of water, liver and skin by phase explosion, confined boiling, recoil-induced expulsion and flow-induced suction. *Proc. SPIE*, 4961:40–47: (2003).

125. Hibst R. *Technik, Wirkungsweise und medizinische Anwendungen von Holmium- und Erbium-Lasern*. Ecomed, Landsberg, pp. 1–120 (in German) (1996).
126. Puliafito CA, Stern D, Krueger RR, and Mandel ER. High-speed photography of excimer laser ablation of the cornea. *Arch. Ophthalmol.*, 105:1255–1259 (1987).
127. Bäuerle D. *Laser processing and chemistry.* Springer, Berlin (2000).
128. Cummings JP and Walsh JT. Tissue tearing caused by pulsed laser-induced ablation pressure. *Appl. Opt.*, 32:494–503 (1993).
129. Phipps CR, Harrison RF, Shimada T, York GW, Turner TP, Corlis XF, Steele HS, Haynes LC, and King TR. Enhanced vacuum laser-impulse coupling by volume absorption at infrared wavelengths. *Lasers Particle Beams*, 8:281–297 (1990).
130. Dyer PE and Al-Dhahir RK. Transient photoacoustic studies of laser tissue ablation. *Proc. SPIE*, 1202:46–60 (1990).
131. Könz F, Frenz M, Pratisto H, Weber HP, Lubatschowski H, Kermani O, Ertmer W, Altermatt HJ, and Schaffner T. Thermal and mechanical damage of corneal tissue after free-running and Q-switched mid-infrared laser ablation. *Proc. SPIE*, 2077:78–86 (1994).
132. Esenaliev RO, Oraevsky AA, Letokhov VS, Karabutov AA, and Malinsky TV. Studies of acoustical and shock waves in the pulsed laser ablation of biotissue. *Lasers Surg. Med.*, 13:470–484 (1993).
133. Engel OGJ. Crater depth in fluid impacts. *J. Appl. Phys.*, 37:1798–1808 (1966).
134. Prosperetti A and Oguz HN. The impact of drops on liquid surfaces and the underwater noise of rain. *Ann. Rev. Fluid Mech.*, 25:577–602 (1993).
135. Walsh JT and Deutsch TF. Er:YAG laser ablation of tissue: Measurement of ablation rates. *Lasers Surg. Med.*, 9:327–337 (1989).
136. Duck FA. *Physical properties of tissue*. Academic, London (1990).
137. Fung YC. *Biomechanics: Mechanical properties of living tissues.* Springer, New York (1993).
138. Watanabe S, Flotte TJ, McAucliffe DJ, and Jacques SL. Putative photoacoustic damage in skin induced by pulsed ArF excimer laser. *J. Invest. Dermatol.*, 90, 761–766 (1988).
139. Yashima Y, McAuclffe DJ, Jacques SL, and Flotte TJ. Laser-induced photoacoustic injury of skin: Effect of inertial confinement. *Lasers Surg. Med.*, 11:62–68 (1991).
140. Pini R, Rossi F, Salimbeni S, Siano S, Vannini M, Carones F, Trabucci G, Brancato R, and Gobbi PG. Experimental investigation on acoustic phenomena induced inside the eyeball by excimer laser ablation of the cornea. *Proc. SPIE*, 2632:25–29 (1996).
141. Nishioka NS and Domankevitz Y. Reflectance during pulsed holmium laser irradiation of tissue. *Lasers Surg. Med.*, 9:375–381 (1989).
142. Kaufmann R and Hibst R. Pulsed erbium:YAG laser ablation in cutaneous surgery. *Lasers Surg. Med.*, 19:324–330 (1996).
143. Venugopalan V, Nishioka NS, and Mikic BB. The effect of CO_2 laser pulse repetition rate on tissue ablation rate and thermal damage. *IEEE Trans Biomed. Eng.*, 38:1049–1052 (1991).
144. Walsh JT and Deutsch TF. Pulsed CO_2 laser tissue ablation: Measurement of the ablation rate. *Lasers Surg. Med.*, 8:264–275 (1988).
145. Shimmick JK, Telfair WB, Munnerlyn CR, Bartlett JD, and Trokel SL. Corneal ablation profilometry and steep central islands. *J. Refract. Surg.*, 13:235–245 (1997).
146. Zhigilei LV and Garrison BJ. Molecular dynamics simulation study of the fluence dependence of particle yield and plume composition in laser desorption and ablation of organic solids. *Appl. Phys. Lett.*, 74:1341–1343 (1999).
147. Yingling YG, Zhigilei LV, Garrison BJ, Koubenakis A, Labrakis J, and Georgiou S. Laser ablation of bicomponent systems: A probe of molecular ejection mechanisms. *Appl. Phys. Lett.*, 78:1631–1633 (2001).
148. Cross FW, Al-Dhahir RK, Dyer PE, and MacRobert AJ. Time-resolved photoacoustic studies of vascular tissue ablation at three laser wavelengths. *Appl. Phys. Lett.*, 50:1019–1021 (1987).

149. Domankevitz Y and Nishioka NS. Measurement of laser ablation threshold with a high-speed framing camera. *IEEE J. Quant. Electon.*, 26:2276–2278 (1990).
150. Walsh JT and Deutsch TF. Pulsed CO_2 laser ablation of tissue: Effect of mechanical properties. *IEEE Trans. Biomed. Eng.*, 36(12):1195–1201 (1989).
151. Deutsch TF and Geis MW. Self-developing UV photoresist using excimer laser exposure. *J. Appl. Phys.*, 54:7201–7204 (1983).
152. Puliafito CA, Wong K, and Steinert RF. Quantitative and ultrastructural studies of excimer laser ablation of the cornea at 193 and 248 nanometers. *Lasers Surg. Med.*, 7:155–159 (1987).
153. Pettit GH and Ediger MN. Corneal tissue absorption coefficients for 193- and 213-nm ultraviolet radiation. *Appl. Opt.*, 35:3386–3391 (1996).
154. Langerholc J. Moving phase-transitions in laser-irradiated biological tissue. *Appl. Opt.*, 18:2286–2293 (1979).
155. McKenzie AL. A three-zone model of soft-tissue damage by a CO_2 laser. *Phys. Med. Biol.*, 31:967–983 (1986).
156. McKenzie AL. An extension of the three-zone model to predict depth of tissue damage beneath Er:YAG and Ho:YAG laser excisions. *Phys. Med. Biol.*, 34:107–114 (1989).
157. Partovi F, Izatt JA, Cothren RM, Kittrell C, Thomas JE, Strikwerda S, Kramer JR, and Feld MS. A model for thermal ablation of biological tissue using laser radiation. *Lasers Surg. Med.*, 7:141–154 (1987).
158. Gerstmann, M, Sagi, A, Avidor-Zehavi, A, Katzir, A, and Akselrod, S. Model simulation of biological damage in tissue exposed to CO_2 laser irradiation. *Opt. Eng.*, 32:291–297 (1993).
159. Olmes A, Franke HG, Bänsch E, Lubatschowski H, Raible M, Dziuk G, and Ertmer W. Modeling of infrared soft-tissue photoablation process. *Appl. Phys. B*, 65:659–666 (1997).
160. Zweig AD and Weber HP. Mechanical and thermal parameters in pulsed laser cutting of tissue. *IEEE J. Quant. Electron.*, 23:1787–1793 (1987).
161. Zhigilei LV, Kodali PBS, and Garrison BJ. Molecular dynamics model for laser ablation and desorption of organic solids. *J. Phys. Chem. B*, 101:2028–2037 (1997).
162. Zhigilei LV, Kodali PBS, and Garrison BJ. A microscopic view of laser ablation. *J. Phys. Chem. B*, 102:2845–2853 (1998).
163. Zhigilei LV and Garrison BJ. Microscopic mechanisms of laser ablation of organic solids in the thermal and stress confinement irradiation regimes. *J. Appl. Phys.*, 88:1281–1298 (2000).
164. Zhigilei LV, Leveugle E, Garrison B, Yingling YG, and Zeifman MI. Computer simulations of laser ablation of molecular substrates. *Chem. Rev.*, 103:321–347 (2003).
165. Dou Y, Winograd N, Garrison B, and Zhigilei L. Substrate-assisted initiation od proteins embedded in water films. *J. Phys. Chem. B*, 107:2362–2365 (2003).
166. Knochenmuss R and Zhigilei LV. Molecular dynamics model of Ultraviolet matrix-assisted laser desorption/ionization including ionization process. *J. Phys. Chem. B*, 109:22947–22957 (2005).
167. Leveugle E and Zhigilei LV. Microscopic mechanisms of short pulse laser spallation of molecular solids. *Appl. Phys. A*, 79:753–756 (2004).
168. Leveugle E, Ivanov DS, and Zhigilei LV. Photomechanical spallation of molecular and metal targets: Molecular dynamics study. *Appl. Phys. A*, 79:1643–1655 (2004).
169. Krauss JM, Puliafito CA, and Steinert RF. Laser interactions with the cornea. *Surv. Ophthalmol.*, 31:37–53 (1986).
170. Puliafito CA, Steinert RF, Deutsch TF, Hillenkamp F, Dehm EJ, and Adler CM. Excimer laser ablation of the cornea and lens. *Ophthalmology*, 92:741–748 (1985).
171. Lane RJ, Linsker R, Wynne JJ, Torres A, and Geronemus RG. Ultraviolet-laser ablation of skin. *Arch. Dermatol.*, 121:609–617 (1985).
172. Seiler T, Marshall J, Rotherty S, and Wollensak J. The potential of an infrared hydrogen fluoride (HF) laser (3,0 μm) for corneal surgery. *Lasers Ophthalmol.*, 1:49–60 (1986).
173. Loertscher H, Mandelbaum S, Parrish RK III, and Parel JM. Preliminary report on corneal incisions created by a hydrogen fluoride laser. *Am. J. Ophthalmol.*, 102:217–221 (1986).

174. Bende T, Kriegerowski M, and Seiler T. Photoablation in different ocular tissues performed with an Erbium:YAG laser. *Lasers Light Ophthalmol.*, 2:263–269 (1989).
175. Wolbarsht ML. Laser surgery: CO_2 or HF. *IEEE J. Quant. Electron.*, 20:1427–1432 (1984).
176. Telfair WB, Bekker C, Hoffman HJ, Yoder PR Jr, Norquist RE, Eiferman RA, and Zenzie HH. Histological comparison of corneal ablation with Er:YAG laser, Nd:YAG optical parametric oscillator, and excimer laser. *J. Refract. Surg.*, 16:40–50 (2000).
177. Yablon AD, Nishioka NS, Mikic BB, and Venugopalan V. Measurement of tissue absorption coefficients by use of interferometric photothermal spectroscopy. *Appl. Opt.*, 38:1259–1272 (1999).
178. Srinivasan R, Dyer PE, and Braren B. Far-Ultraviolet laser ablation of the cornea: Photoacoustic studies. *Lasers Surg. Med.*, 6:514–519 (1987).
179. Ishihara M, Arai T, Sato S, Morimoto Y, Obara M, and Kikuchi M. Measurement of the surface temperature of the cornea during ArFexcimer laser ablation by thermal radiatiometry with a 15 nanosecond time response. *Lasers Surg. Med.*, 30:54–59 (2002).
180. Kermani O, Koort HJ, Roth E, and Dardenne MU. Mass-spectroscopic analysis of excimer laser ablated material from human corneal tissue. *J. Cataract Refract. Surg.*, 14:638–641 (1988).
181. Hahn DW, Ediger MN, and Petttit GH. Dynamics of ablation plume particles generated during excimer laser corneal ablation. *Lasers Surg. Med.*, 16:384–389 (1995).
182. Isner JM, deJesus S R, Clarke RH, Gal D, Rongione AJ, and Donaldson RF. Mechanism of laser ablation in an absorbing flow field. *Lasers Surg. Med.*, 8:543–554 (1988).
183. Deckelbaum LI. Coronary laser angioplasty. *Lasers Surg. Med.*, 14:101–110 (1994).
184. Köstli KP, Frenz M, Weber HP, Paltauf G, and Schmidt-Kloiber H. Optoacoustic measurements of water, bone and cartilage with an infrared-OPO. *Proc. SPIE*, 3601:310–318 (1999).
185. Dyer PE, Khosroshahi ME, and Tuft SJ. Studies of laser-induced cavitation and tissue ablation in saline using a fibre-delivered pulsed HF laser. *Appl. Phys. B*, 56:84–93 (1993).
186. Brinkmann R, Hansen C, Mohrenstecher D, Scheu M, and Birngruber R. Analysis of cavitation dynamics during pulsed laser tissue ablation by optical on-line monotoring. *IEEE J. Sel. Top. Quant. Electron.*, 2:826–835 (1996).
187. Frenz M, Könz F, Pratisto H, Weber HP, Silenok AS, and Konov VI. Starting mechanisms and dynamics of bubble formation induced by a Ho:Yttrium aluminium garnet laser in water. *J. Appl. Phys.*, 84:5905–5912 (1998).
188. Ith M, Pratisto H, Altermatt HJ, Frenz M, and Weber HP. Dynamics of laser-induced channel formation in water and influence of pulse duration on the ablation of biotissue under water with pulsed erbium-laser radiation. *Appl. Phys. B*, 59:621–629 (1994).
189. Vogel A, Engelhardt R, Behnle U, and Parlitz U. Minimization effects in pulsed laser ablation illustrated in laser angioplasty. *Appl. Phys. B*, 62:173–182 (1996).
190. Pratisto H, Frenz M, Ith M, Altermatt HJ, Jansen ED, and Weber HP. Combination of fiber-guided pulsed erbium and holmium laser radiation for tissue ablation under water. *Appl. Opt.*, 35:3328–3337 (1996).
191. Jansen ED, Asshauer T, Frenz M, Motamedi M, Delacrétaz G, and Welch AJ. Effect of pulse duration on bubble formation and laser-induced pressure waves during holmium laser ablation. *Lasers Surg. Med.*, 18:278–293 (1996).
192. Vogel A, Schmidt P, and Flucke B. Minimization of thermomechanical side effects in IR ablation by use of multiply Q-switched laser pulses. *Med. Laser Appl.*, 17:15–20 (2002).
193. Jansen ED, van Leeuwen TG, Motamedi M, Borst C, and Welch AJ. Partial vaporization model for pulsed mid-infrared laser ablation of water. *J. Appl. Phys*., 78:564–571 (1995).
194. Skripov VP and Pavlov PA. Superheating and explosive boiling of liquids. *Sov. Tech. Rev. B. Therm. Phys.*, 2:171–207 (1989).
195. Skripov VP. Thermodynamic stabilities of superheated and supercooled liquids. *Fluid Mech. Res.*, 21:43–50 (1992).

196. Paltauf G, Schmidt-Kloiber H, and Frenz M. Photoacoustic waves excited in liquids by fiber-transmitted laser pulses. *J. Acoust. Soc. Am.*, 104:890–897 (1998).
197. Dingus RS. Laser-induced contained vaporization in tissue. *Proc SPIE*, 1646:266–274 (1992).
198. Vogel A, Noack J, Nahen K, Theisen D, Busch S, Parlitz U, Hammer DX, Noojin GD, Rockwell BA, and Birngruber R. Energy balance of optical breakdown in water at nanosecond to femtosecond time scales. *Appl. Phys. B*, 68:271–280 (1999).
199. van Leeuwen T G, Meertens JH, Velema E, Post MJ, and Borst C. Intraluminal vapor bubble induced by excimer laser pulse causes microsecond arterial dilitation and invagination leading to extensive wall damage in the rabbit. *Circulation*, 87:1258–1263 (1993).
200. Cross FW, Al-Dhahir RK, and Dyer PE. Ablative and acoustic response of pulsed UV laser-irradiated vascular tissue in a liquid environment. J. *Appl. Phys.*, 64:2194–2201 (1988).
201. Brujan EA, Nahen K, Schmidt P, and Vogel A. Dynamics of laser-induced cavitation bubbles near an elastic boundary. *J. Fluid Mech*., 433:251–281 (2001).
202. Brujan EA, Nahen K, Schmidt P, and Vogel A. Dynamics of laser-induced cavitation bubbles near elastic boundaries: Influence of the elastic modulus. *J. Fluid Mech*., 433:283–314 (2001).
203. Vogel A, Hentschel W, Holzfuss J, and Lauterborn W. Cavitation buble dynamics and acoustic transient generation in ocular surgery with pulsed neodymium:YAG lasers. *Ophthalmology*, 93:1259–1269 (1986).
204. Vogel A, Schweiger P, Frieser A, Asiyo M, and Birngruber R. Intraocular Nd:YAG laser surgery: Light-tissue interaction, damage range, and reduction of collateral effects. *IEEE J. Quant. Electron.*, 26:2240–2260 (1990).
205. Lin C, Stern D, and Puliafito CA. High-speed photography of Er.YAG laser ablation in fluid. *Invest. Ophthalmol. Vis. Sci.*, 31:2546–2550 (1990).
206. Vogel A, Busch S, and Parlitz U. Shock wave emission and cavitation bubble generation by picosecond and nanosecond optical breakdown in water. *J. Acoust. Soc. Am.*, 100:148–165 (1996).
207. Brennen CE. *Cavitation and bubble dynamics.* Oxford University, Oxford, New York (1995).
208. Vogel A, Lauterborn W, and Timm R. Optical and acoustic investigation of the dynamics of laser-produced cavitation bubbles near a solid boundary. *J. Fluid Mech.*, 206:299–338 (1989).
209. Palanker D, Vankov A, Miller J, Friedman M, and Strauss M. Prevention of tissue damage by water jet during cavitation. *J. Appl. Phys*., 94:2654–2661 (2003).
210. Chapyak EJ and Godwin RP. Simulations of laser thrombolysis. *Proc. SPIE*, 3590:328–335 (1999).
211. Frenz M, Pratisto H, Könz F, Jansen ED, Welch AJ, and Weber HP. Comparison of the effects of absorption coefficient and pulse duration of 2.12 μm and 2.79 μm radiation on laser ablation of tissue. *IEEE J. Quant. Electron*., 32:2025–2036 (1996).
212. Srinivasan R, Casey KG, and Haller JD. Subnanosecond probing of the ablation of soft plaque from arterial wall by 308 nm laser pulses delivered through a fiber. *IEEE J. Quant. Electron.*, 26:2279–2283 (1990).
213. Brinkmann R, Dröge G, Schröer F, Scheu M, and Birngruber R. Ablation dynamics in laser sclerostomy ab externo by means of pulsed lasers in the mid-infrared spectral range. *Ophthalm. Surg. Lasers*, 28:853–865 (1997).
214. Brinkmann R, Theisen D, Brendel T, and Birngruber R. Single pulse 30-J holmium laser for myocardial revascularisation – a study on ablation dynamics in comparison to CO_2 Laser. *IEEE J. Sel. Top. Quant. Electron.*, 5:969–980 (1999).
215. Jansen ED, van Leeuwen TG, Verdaasdonk RM, Le TH, Motamedi M, Welch AJ, and Borst C. Influence of tissue mechanical strength during UV and IR laser ablation in vitro. *Proc. SPIE*, 1882:139–146 (1993).
216. Frenz M, Mischler C, Romano V, Forrer M, Müller OM, and Weber HP. Effect of mechanical tissue properties on thermal damage in skin after IR-laser ablation. *Appl. Phys. B*, 52:251–258 (1991).

217. Jansen ED, Frenz M, Kadipasaoglu KA, Pfefer T, Altermatt HJ, Motamedi M, and Welch AJ. Laser-tissueinteraction during transmyocardial laser revascularisation. *Proc. SPIE*, 2671:49–57 (1996).
218. Müller G, Dörschel K, and Kar H. Biophysics of the photoablation process. *Lasers Med. Sci.*, 6:241–254 (1991).
219. Hibst R and Kaufmann R. Effects of laser parameters on pulsed Er-YAG laser skin ablation *Lasers Med. Sci.*, 6:391–397 (1991).
220. Schomacker KT, Walsh JT, Flotte TJ, and Deutsch TF. Thermal damage produced by high-irradiance continuous wave CO_2 laser cutting of tissue. *Lasers Surg. Med.*, 10:74–84 (1990).
221. Neev J, Da Silva LB, Feit MD, Perry MD, Rubenchik AM, and Stuart BC. Ultrashort pulse lasers for hard tissue ablation. *IEEE J. Sel. Top. Quant. Electron*., 2:790–800 (1996).
222. Majaron B, Srinivas SM, Huang HL, and Nelson JS. Deep coagulation of dermal collagen with repetitive Er:YAG laser irradiation. *Lasers Surg. Med.*, 26:215–222 (2000).

Chapter 15
Pulsed Laser Tissue Interaction

Joseph T. Walsh, Ton G. van Leeuwen, E. Duco Jansen, Massoud Motamedi, and Ashley J. Welch

15.1 Introduction

Pulsed lasers, by virtue of their ability to deliver energy in a spatially and temporally confined fashion, are able to micromachine biological tissues. The clinical success of pulsed laser treatment, however, is often limited by the extent of damage that is caused to the tissue in the vicinity of the ablation crater. In general, pulsed ablation is a trade off between thermal damage to surrounding tissue, caused by relatively long pulses (>100 ms), and mechanical damage to surrounding tissue, caused by relatively short pulses (<1 ms). To identify the origin of pulsed laser induced damage, the possible laser tissue interactions and ablation are discussed here and in Chapter 14. The purpose of this chapter is to provide the reader with a condensed overview of the parameters that must be considered in the process of pulsed laser ablation of soft tissue. In this chapter, pulsed infrared ablation of biological soft tissue is used as a paradigm to illustrate the concepts and design considerations. Generally speaking, the absorption of laser light may lead to photothermal, photomechanical or photochemical interactions with the irradiated tissue [1–5]. The vast majority of therapeutic laser-tissue interactions is based on photothermal interactions where laser energy is converted into heat. Subsequent to thermalization of the absorbed optical energy, heat transfer mechanisms, in particular conduction, allow thermal diffusion from high temperature areas to surrounding regions. When laser penetration depth is less than the laser spot radius, the thermal diffusion time, τ_{th}, can be defined as:

$$\tau_{\text{th}} = \frac{\delta^2}{4\alpha} \tag{15.1}$$

with δ (mm) the optical penetration depth of laser light in tissue ($\delta = 1/\mu_a$) and with thermal diffusivity, $\alpha = 0.15\,(\text{mm}^2/\text{s})$ for water at 37°C (see Chapter 10). For

J.T. Walsh (✉)
Biomedical Engineering Department, Northwestern University, Evanston, IL, USA
e-mail: jwalsh@northwestern.edu

A.J. Welch, M.J.C. van Gemert (eds.), *Optical-Thermal Response of Laser-Irradiated Tissue*, 2nd ed., DOI 10.1007/978-90-481-8831-4_15,

laser pulses shorter than the thermal diffusion time (i.e. $\tau_p < \tau_{\text{th}}$), the distribution of thermal energy is determined by the laser light distribution (thermal confinement) [4, 5]. If the laser pulse duration exceeds the thermal diffusion time, the thermal energy propagates into the tissue during the laser pulse [4–7].

Photomechanical interactions include the generation of pressure waves and the effects of explosive vaporization of tissue during the ablation process. Pressure waves are generated by one of three mechanisms: (1) the thermoelastic expansion of tissue, due to heating of the target tissue by the pulsed laser light; (2) recoil caused by ejection of the ablated material as dictated by Newton's third law; or (3) the collapse of vapor cavities (bubbles) when the ablation process takes place in a liquid environment. Laser induced stress waves, which propagate with the speed of sound ($\sigma = 1500\,(\text{m/s}) = 1.5\,(\text{mm}/\mu sec)$ in water or soft tissues) or even faster as in the case of shockwaves [8], have the ability to cause significant mechanical damage in tissue even in areas that are beyond the optical interaction zone. Generation by thermoelastic expansion requires stress confiment but can occur at radiant exposures well below the ablation threshold (in fact this phenomenon lies at the basis of photoacoustic technologies described elsewhere in this book. See Chapter 19). The time it takes laser-induced pressure increase to propagate out of the irradiated volume is known as the stress confinement time (τ_{str}) and is given by:

$$\tau_{\text{str}} = \frac{\delta}{\sigma} \tag{15.2}$$

where δ the penetration depth of laser light in tissue, and σ is the speed of sound in tissue. If the laser pulse duration is shorter than stress confinement time (i.e. $\tau_p < \tau_{\text{str}}$), large peak stresses can be reached as temperature-induced pressure increases cannot diffuse out of the irradiated volume, resulting in pressure build-up and the generation of high amplitude pressure transients. These stresses may contribute to the ablation mechanism and may inflict damage to adjacent tissue [5, 9, 10]. It can be seen from Eqs. (15.1) and (15.2) that confinement (stress or thermal) depends on the laser pulse length and penetration depth $1/\mu_a$. In Fig. 15.1 the pulse length versus penetration depth is plotted for various lasers. This figure provides an easy way to determine for any laser source what, if any, confinement can be expected. It should be noted that the stress and thermal confinement zones in reality represent gradual changes from one zone to the other. As an example, it can be seen that for a free-running holmium laser with a pulse duration of several hundred microseconds and penetration depth of 300–400 μm, the interaction will be thermally confined but not stress confined. If stress confinement is desired, the pulse duration for this laser must be shortened to approximately 100 ns which can be accomplished by Q-switching.

In contrast, the other two potential sources of laser-induced stress waves, ablative recoil and bubble collapse, do not require stress confinement but are dependent on the ablation process itself.

Even in the absence of pressure wave generation, a major photomechanical interaction during mid-infrared laser ablation originates from tissue water vaporization, resulting in explosive removal of tissue structures, described as the "popcorn

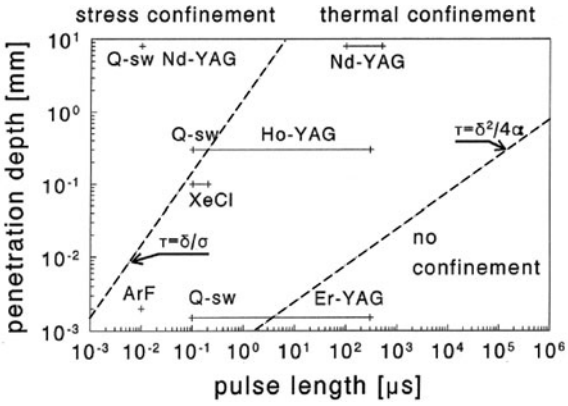


Fig. 15.1 Penetration depth δ of laser light in tissue as a function of laser pulse length τ_p. The photomechanical effects are classified for various laser types. The stress confined region is given by $\tau_p, (\delta/\sigma)$, with σ the speed of sound. Note that the XeCl excimer laser and the long pulsed holmium-YAG laser are classified only in the thermal confinement zone ($\delta/\sigma < \tau_p < \delta^2/4\alpha$, with α the thermal diffusivity.) *Legend*: Q-sw: Q-switched, Nd-YAG: neodymium-YAG laser (wavelength $\lambda = 1064$ nm), Ho-YAG; holmium-YAG laser ($\lambda = 2.12$ µm), ArF: argon-fluoride excimer laser ($\lambda = 193$ nm), XeCl: xenon choloride excimer laser ($\lambda = 308$ nm), Er-YAG: erbium-YAG laser ($\lambda = 2.94$ µm). (Adapted from [5])

effect" for the cw lasers [11, 12]. Similarly, for pulsed laser ablation, this explosive vaporization-driven process may result in tissue mechanical damage that is manifested as fissures, tears, damage to extracellular matrix and cellular damage. Generally, the shorter the pulse duration, the more violent the explosion and subsequent mechanical damage.

Photochemical ablation involves molecular bond breaking. The absorbed photon excites the molecule to a higher electronic, vibrational or rotational state. If the photon energy is high enough to excite the molecule to an electronic state which is not bound, the molecule dissociates [13]. Then, it may form radicals (as in photodynamic therapy) [14, 15] or it may break (as proposed for far ultraviolet laser ablation) [16, 17]. This phenomenon plays a significant role for wavelengths well below 300 nm when the photon energy exceeds 4.13 eV. In the application of the ArF laser at 193 nm in vision correction procedures, photochemical decomposition is believed to play a major role. For infrared ablation, the focus in this chapter, photochemical interactions are insignificant and will not be discussed.

15.1.1 Pulsed Infrared Laser Ablation

Infrared laser ablation is believed to be a process of water vaporization and photothermal disruption of tissue. The main absorber of the IR wavelength is tissue water which is heated and vaporized (see Chapter 14, Fig. 14.2a). This process is similar to ablation with cw irradiation and is associated with a rapid subsurface

pressure buildup with possible superheating. Eventually an explosion follows which is accompanied by an acoustical transient or shock wave and rapid ejection of tissue fragments [18]. The influence of laser parameters on the ablation process, temperatures and pressures associated with the ablation process has been the subject of many studies and is reasonably well understood. Nevertheless, unanswered questions and controversies remain. To what extent observed mechanical damage can be attributed to photothermal and photomechanical processes (i.e. bubble formation) or to photoacoustical mechanisms yielding pressure transients with amplitudes as high as 1000 bars [19–22], remains to be determined.

15.1.2 Thermal Damage

With a few notable exceptions (that involve the desire to coagulate tissue while ablating) induction of thermal damage to tissue adjacent to the ablation crater, is an undesirable side effect of laser ablation. As laser irradiation is delivered to tissue, absorption of laser energy by tissue leads to an increase in temperature. If this situation exists long enough, heat is conducted from directly heated sites to cooler areas in the tissue. This is one of the main reasons to favor pulsed lasers over cw lasers for tissue ablation. As was stated earlier in this chapter, there is the desire to produce a thermal event that is shorter than the thermal relaxation time of the tissue in order to reduce the zone of thermal damage. In this case, the laser pulse is too short to allow significant thermal diffusion to take place during the short exposure. Secondly, the delivery of sufficient energy to ablate tissue with each pulse removes hot tissue before heat is transferred to surrounding tissue.

In the region where pulsed laser energy is directly absorbed but where the energy density is below the ablation threshold, thermal damage is caused by a temperature rise resulting from non-ablative laser light or heat accumulation induced by subsequent laser pulses. Thus, even after pulsed tissue ablation, thermal damage may be significant. Depending on the temperature time history of the heated tissue, cells may die. After one to three days survival, thermally induced necrosis can be observed by microscopic analysis of histologic slices. In vitro, under light microscopy, a marker for thermal damage is a zone of vacuolization around the ablation crater. This zone corresponds roughly with a temperature of close to 100°C, although it should be kept in mind that thermal damage to biological tissues is always a temperature-time governed process. A more subtle way of assessing thermal damage is by looking at birefringence loss, as has been described by Thomsen et al. in Chapter 13 of this book [23, 24].

15.1.3 Multiple Origins of Tissue Damage

It is highly likely that laser-induced damage is caused by a combined effect. Pressure waves are often accompanied by bubble formation. At the start of bubble expansion and upon collapse in a liquid environment, pressure waves are generated.

To complicate the process, bubble formation, which is often referred to by the more generic term cavitation, can be initiated by tensile stresses (e.g. from reflected pressure waves), plasma formation or water vaporization. The first is well known from fast rotating ship propellers [25]. The threshold for plasma formation (and hence cavitation) is typically intensity-related (10^8–10^{12} W/cm^2) [8, 26, 27]. A detailed discussion of this phenomenon can be found in Chapter 14. Cavitation associated with infrared laser ablation is most commonly caused by direct heating and vaporization of water [28]. During bubble expansion, the water vapor may condense. The released condensation energy may induce thermal damage to surrounding tissue. Upon bubble collapse, additional pressure transients are generated. The peak pressures of these pressure waves are often higher than those generated at the start of the laser pulse but are strongly dependent on bubble size and geometry with more spherical cavities generating the largest collapse pressures [28, 29]. Also, bubble collapse in the vicinity of solid boundaries can create high velocity fluid jets that can impinge on that boundary, creating additional damage. After the bubble has collapsed, it may re-expand (to a smaller volume than that of the initial bubble) and collapse again. An example of explosive bubble formation and collapse is shown in Fig. 15.2.

Damage observed after in vitro pulsed laser ablation has been described as acoustic damage [31–35]. However, considering the energy and damage range, it is unlikely that pressure waves cause macroscopic tissue rupturing. A detailed treatise of the energy balance of the various components of the ablation process can be found in Chapter 14. In brief, the potential energy of the bubble is defined as the product of

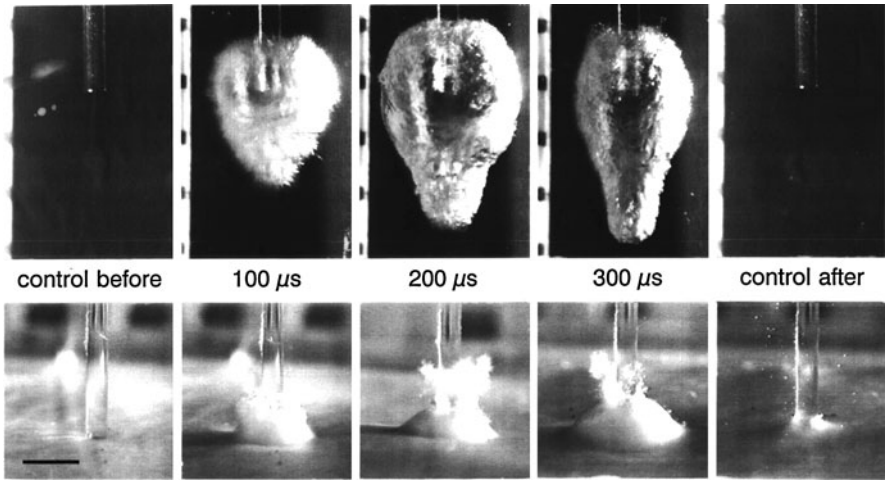


Fig. 15.2 Explosive bubble formation and collapse due to a Holmium laser pulse (500 μs duration and 500 mJ pulse energy) delivered via a 320 μm fiber tip. In the top panels, the bubble formation in water is shown. Please note the pear shape of the bubble, which is due to the vaporization of the water during the pulse. In the lower panels, the bubble formation in (pig aortic) tissue can be observed as an elevation of the tissue surface. The bubble formation in the tissue causes mechanical damage in the form of fissures [30]

the volume of the bubble and the environmental pressure [36]. The loss of potential energy during each cycle of collapse and re-expansion is (partly) emitted as energy of the pressure wave [36]. Consequently, the potential bubble energy which is possibly hazardous, is larger than the energy in the shockwave. Furthermore, shockwaves can only induce small translations (<1 μm) of tissue structures, which might be too small to cause the macroscopic fissures in tissue [37]. It is more likely that shock and pressure waves, in particular the tensile components, have an effect at the subcellular level (on organelles) [4, 5, 9]. Flotte et al. nicely demonstrated from measuring red blood cell lysis, that the largest contribution to damage to the red blood cells was due to bubble formation and not due to pressure waves [38]. Thus, it is to be expected that the contribution of bubble formation during pulsed laser ablation to damage of adjacent tissue is larger than the acoustically induced damage [82].

15.2 General Design Considerations for IR Ablation

There are several irradiation parameters that must be selected in the design of a pulsed laser based system for the removal of tissue. Typically, one selects the wavelength first. Unfortunately, it is usually the case that not all parameters can be selected independently. Figure 15.3 shows the inter-relationship among irradiance, spot size, and pulse duration for a constant pulse energy. If commercially available systems are to be used, the wavelength selection constrains the selection of the other irradiation parameters. For example, a 5 μs long pulse from a Holmium:YAG laser is not available today, although these boundaries continue to be pushed by the development of solid state lasers (diode lasers, diode-pumped solid state (DPSS) lasers and OPO/OPA systems). On the following pages the irradiation parameters are discussed roughly in order of importance. After selection of the wavelength, a laser system is specified that will emit sufficient energy with an appropriate pulse duration. Finally, the size of the irradiated area, the pulse repetition rate, and the beam profile are selected.

15.2.1 Wavelength Considerations

A pulsed laser used for ablative purposes must emit laser energy that can get to the lesion, remove the lesion, and leave a surface that, ideally, permits normal wound healing.

With the exception of those procedures that target tissues on the surface of the body (i.e. skin, cornea), the surgical procedure must be minimally invasive – i.e. the laser energy must be transmitted through a fiber. Commercially available fibers can transmit radiation from approximately 300 nm to 2.5 μm. Thus, certain UV-emitting (excimer) and IR-emitting solid-state and gas lasers cannot be used without special consideration to the delivery system. Practically speaking, this consideration limits the choice of wavelength for procedures where fiberoptic delivery is required to those wavelengths that can in fact be transmitted through fibers, despite the fact that

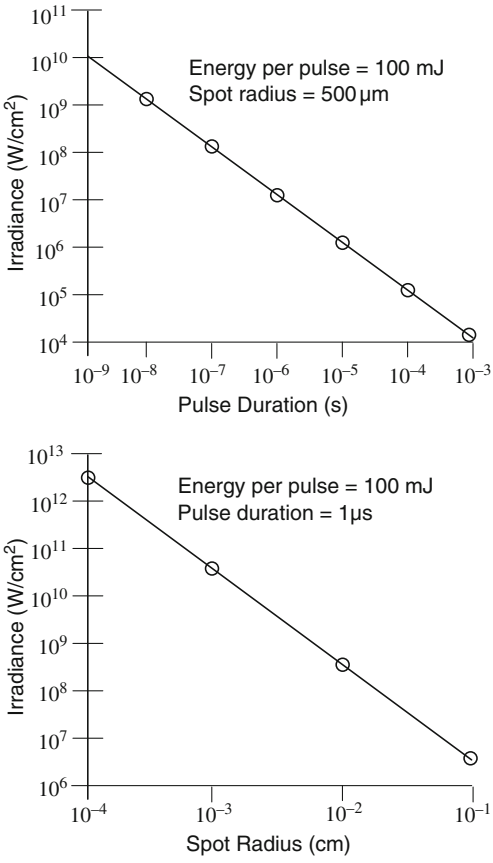


Fig. 15.3 The interrelationship among irradiance, spot radius, and pulse duration for a constant pulse energy of 100 mJ/pulse

from a pure ablative precision point of view far UV (e.g. ArF at 193 nm) or mid-IR (Er:YAG at 2,940 nm) lasers maybe more desirable.

In principle, a laser at virtually any wavelength can be used to remove either soft or hard tissue. The key issue is long-term recovery. If a wavelength is selected that is poorly absorbed by the target tissue, then the incident light penetrates deeply within the tissue causing unwanted thermal damage which, in turn, may impede the normal healing process and/or result in suboptimal clinical outcomes.

Consequently, it is desirable to select a laser that emits radiation which is strongly absorbed by the target tissue. The penetration depth of the radiation controls the volume heated and thus the energy required for ablation. This, in turn, determines the overall thermal load to the tissue. Moreover, it dictates the depth of damage and depth of the cut per pulse. A shallow penetration depth generally means the residual thermal damage is limited to the structures near the tissue surface and only the tissue near the surface is ablated. There are two regions of the spectrum in which absorption of optical energy is strong: the UV and the IR. In general, wavelengths between 600 and 1100 nm penetrate deeply into tissue. Proteins and nucleic acids absorb

strongly in the UV. Water is the major chromophore in the IR with absorption peaks near 1.93 μm ($\mu_a \approx 80\ \text{cm}^{-1}$) and 2.94 μm ($\mu_a \approx 13,000\ \text{cm}^{-1}$) [39]. Similarly, there is very strong absorption by tissue in the UV at wavelengths less than 200 nm; in particular, at 193 nm where the ArF excimer laser emits, μ_a is approximately 2700 cm^{-1} [40].

At the high irradiances seen during pulsed laser ablation of tissue there is evidence that the absorption coefficient of the tissue changes during the laser pulse. The issue of dynamic changes in optical absorption is covered more extensively in Section 15.3 of this chapter and in Chapter 9.

15.2.2 Pulse Duration Considerations

The pulse duration of the laser controls a number of events. First, if the pulse duration is short enough, there will be little diffusion of the thermalized optical energy during the pulse. In general, the pulse duration, τ_p, must be shorter than the thermal relaxation time, τ_{th}, of the laser heated layer of tissue in order to confine the thermal energy as discussed earlier in this chapter (Section 15.1). To a first approximation, for wavelengths where the penetration depth ($1/\mu_a$) is smaller than the laser spot size (typically for wavelengths with a shallow penetration depth), the thermal relaxation time is equal to the axial time constant (Eq. (15.1)).

Thus, for a 10 μm penetration depth the thermal energy of a single pulse will be confined to the site of absorption if the pulse duration is less than 200 μsec (see Table 15.1).

If the pulse duration is short enough to confine the thermalized energy, the ablation process based upon tissue heating will be maximally efficient: lengthening the pulse so $\tau_p > \tau_{\text{th}}$ will decrease the ablation efficiency and increase the percentage of energy left to damage the non-ablated tissue at the cut edge [41]. Thus, the pulse duration controls the efficiency of the ablation process and the extent of the residual thermal damage zone. Once the pulse is short enough, then the temperature rise, T, as a function of distance, z, into the tissue at time τ_p is linearly related to the pulse energy:

$$T(z, t = \tau_p) = \frac{W}{\rho c} = \frac{\mu_a H_0 e^{-\mu_a z}}{\rho c} \tag{15.3}$$

where $W = Q/V$ is the energy deposited per unit volume (J/cm^3), ρ is the tissue density (g/cm^3), c is the specific heat (J/g K), and H_0 is the radiant exposure per

Table 15.1 Relationship between absorption coefficient and thermal relaxation time (Eq. (15.1)) of heated layer (spot size $>> 1/\mu_a$) $\propto\approx 0.13$ [mm^2/s]

$\mu_a(\text{cm}^{-1})$	τ_{th}
10^4	2 μs
10^3	200 μs
10^2	20 ms

pulse at the tissue surface (J/cm^2). This expression is valid if absorption dominates over scattering, surface reflectivity is negligible, and μ_a, ρ, and c do not vary with energy deposited per unit volume. After the pulse, the tissue temperature is controlled by diffusion of the thermal energy in the tissue and removal of energy by ablation of hot tissue from the surface, i.e. ablative cooling.

Pulses that are longer than the thermal diffusion time constant do not look impulsive; and the tissue temperature during the laser pulse is determined by the complicated relationship between energy deposition, thermal diffusion, and ablative cooling. If one neglects ablative cooling, then the temperature rise can be calculated readily from the superposition of the temperature rise induced by overlapping, appropriately-weighted impulse responses (see also Chapter 10).

Not only does one wish to confine the thermalized optical energy to the site of absorption but one also wishes to avoid pressure build up by allowing dissipation of laser-induced increases in pressure such that large amplitude stress waves cannot build up. This is shown in Eq. (15.2). As long as the laser pulse duration is longer than the time constant for stress relaxation, thermoelastic pressures will not be generated. The obvious exception to this are those applications where the efficacy of the process relies on the pressure transients themselves (some cases of laser lithotripsy or pressure wave enhanced delivery of drugs or therapeutic genes).

Independent of the thermoelastic expansion, short pulse ablation is explosive in nature and will result in plume formation (plume velocities greater than the speed of sound in air have been measured by several investigators [42–44]).

There are some practical limits on how short the pulse duration can be. If a constant energy pulse is shortened to the point where the incident irradiance is approximately 10^8 W/cm^2, then plasma formation is likely. Plasma formation decreases ablation efficiency during pulsed laser ablation of tissue (see Chapter 14).

Shortening of the laser pulse also plays a role in optical fiber damage. In brief, optical fiber damage is irradiance dependent. Thus, if a certain energy per pulse is needed to achieve efficient ablation, then the pulse must be long enough so that the irradiance for fiber damage is not exceeded.

15.2.3 Energy or Power Density Considerations

15.2.3.1 First-Order Approximations

The amount of tissue removed is critically dependent upon the amount of energy delivered to the tissue. There are two important parameters: the pulse energy, Q_p (J), and the peak power, P_{peak} (W). For a temporally-uniform pulse, the two are related by a third parameter, the pulse duration, τ_p:

$$P_{\text{peak}} = \frac{Q_p}{\tau_p} \tag{15.4}$$

Thus, the three parameters cannot be independently varied. The choice of which variable best describes the ablation process affects how one thinks about the process.

In general, if $\tau_p < \tau_{th}$, it is appropriate to consider the energy delivered because the energy deposited per unit volume is the determinant factor in the ablation process. If $\tau_p > \tau_{th}$, it is appropriate to consider the power delivered because the energy deposited per unit time (i.e. the power) per unit volume drives the ablation process. Most biological soft tissue is composed mainly (70–80%) of water. Room temperature water vaporizes when approximately 2,500 J/cm^3 are deposited adiabatically. Energetically, this heat of ablation, W_{abl}, derives from the energy density required to increase the temperature of the water from the starting temperature (37°C) to its boiling point (100°C at 1 atm) and the energy required for the phase transition from liquid water to water vapor:

$$W_{abl} = \rho c \Delta T + \rho L_v \tag{15.5}$$

Thus, complete vaporization of tissue water requires approximately 2,500 J/cm^3 • 0.8 = 2,000 J/cm^3. However, it is unlikely that pulsed laser ablation of tissue is driven by a vaporization process governed solely by the latent heat of vaporization. In brief, if the energy deposition occurs rapidly, water does not undergo a phase change at 100°C with an immediate volume expansion of the water. Rather the energy is deposited approximately isovolumetrically and the water temperature and pressure rise rapidly. Subsequently, volumetric expansion of the tissue occurs due to the high pressures within the tissue – i.e. there is explosive ablation (a so-called "phase explosion") of the tissue. Depending upon the mechanical properties of tissue, this explosive removal process, which will include the ejection of droplets and particles of matter in their condensed state, can be more energetically efficient than a model that assumes complete vaporization of the irradiated volume.

Given the energy in a pulse, Q_p, one can calculate the mean energy that is deposited per unit volume at the tissue surface:

$$W|_{z=0} = \frac{Q}{V}|_{z=0} = \frac{(1-R)(Q_p\,\mu_a)}{A} = (1-R)\,\mu_a\,H_0 \tag{15.6}$$

where A is the area irradiated, H_0 is the incident radiant exposure, R is the fraction of incident energy remitted at the tissue surface, and one assumes that absorption dominates over scattering. It should be noted that the index of refraction of tissue is approximately 1.34; thus, one would expect a priori that R due to Fresnel reflections is small compared with the energy deposited within the $1/\mu_a$ depth. However, measurements of R indicate that as much as 40% of incident 2.1 μm laser energy can propagate back from the tissue surface [45]. The remitted light may be Fresnel reflection, scattering by the ablation plume, or remission of light scattered from within the tissue. Nonetheless, it is clear that R can be a significant component in Eq. (15.6) and should be considered in the calculation of the energy deposited at the tissue surface.

There are two first-order models of ablation that form the foundation upon which pulsed laser ablation can be understood. The first model, known as the steady state model, assumes that incident energy is absorbed: and once the tissue is heated to a

threshold energy deposited per unit volume the tissue is removed during the pulse and no longer absorbs incident radiation. This approach is essentially the same as the modeling of continuous wave ablation with the understanding that pulsed laser ablation is merely continuous ablation with a very short radiation time. Thus, in essence, the first model forces there to be a linear relationship between ablation rate and incident energy. The second model, referred to as the blow-off model, assumes that absorption of the incident radiation follows Beer's Law for the entire pulse and that ablation occurs when a differential volume of tissue absorbs more than a threshold energy per unit volume. In this model it is assumed that no material is removed during the laser pulse. Thus, in essence, the second model forces there to be a logarithmic relationship between ablation rate and incident energy. Neither model is accurate over a wide range of tissue and irradiation parameters; second-order effects must be considered for model accuracy to be improved.

For an ablation process that is best fit by a linear relationship between the incident radiant exposure per pulse, H_0, and the mass of tissue removed per pulse, the steady state model can be applied, and we can write

$$\text{Mass loss} = A\,\varepsilon_{\text{slope}}\,(H_0 - H_{\text{threshold}}) \text{ for } \; H_0 > H_{\text{threshold}} \tag{15.7}$$

where A is the area irradiated, $\varepsilon_{\text{slope}}$ is the slope efficiency of ablation, and $H_{\text{threshold}}$ is the threshold radiant exposure for ablation. One can, alternatively, write a similar expression for the relationship between the incident radiant exposure per pulse, H_0, and the etch depth per pulse:

$$\text{Etch depth} = \frac{\varepsilon_{\text{slope}}(H_0 - H_{\text{threshold}})}{\rho} = \frac{\text{mass loss}}{A\rho} \text{ for } H_0 > H_{\text{threshold}} \tag{15.8}$$

From the slope of either an etch depth versus radiant exposure per pulse curve or mass loss versus radiant exposure curve, one calculates the slope efficiency, $\varepsilon_{\text{slope}}$. The slope efficiency is a measure of the additional tissue that will be removed if the incident energy is increased (see Fig 15.4). Thus, the slope efficiency is a measure of the energy invested in the ablation process; and from the inverse of the slope, one finds the heat of ablation. Figure 15.4 shows typical ablation data for infrared radiation absorbed near the surface; note the intercept of the regression line is the threshold radiant exposure for ablation, H_{th}, i.e. the minimum incident energy per unit area to achieve tissue removal. Note further that the product $\mu_a H_{\text{th}}[\text{J/cm}^3]$ is the threshold energy deposition per unit volume required for ablation. As a first-order approximation, for pulses that are short compared to τ_{th}, this product is related to the heat of vaporization of water. Note that the data shown in Fig. 15.4 indicate that $H_{\text{th}} = 1\ \text{J/cm}^2$. Given that $\mu_a \approx 720\ \text{cm}^{-1}$, i.e. assuming that water is the dominant absorber at 10.6 μm and that the tissue is approximately 70% water, one calculates that the actual threshold energy density required for ablation is approximately $720\ \text{J/cm}^3$.

The slope efficiency differs from the absolute efficiency, $\varepsilon_{\text{absolute}}$, of the ablation process. The absolute efficiency, defined as the ratio of mass removed to the

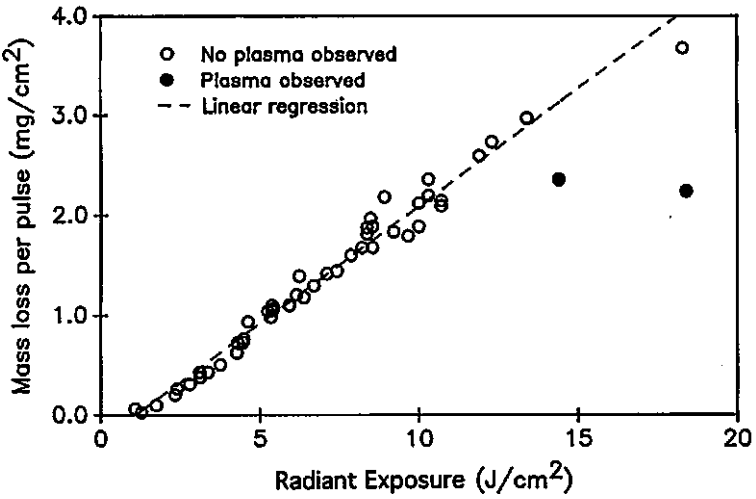


Fig. 15.4 The mass of guinea pig skin removed per pulse by a TEA CO_2 laser ($\lambda = 10.6\mu$m, $\tau_p = 2\,\mu$s, spot size ≈ 1.6 mm $\times$ 2.0 mm rectangular, beam profile approximately flat top) versus the incident radiant exposure per pulse. The heat of ablation is 4.26 kJ/cm^3 as determined from the inverse of the slope and multiplying by tissue density ρ. The threshold radiant exposure for ablation is 1.09 J/cm^2. A plasma that forms at higher radiant exposures consumes some of the incident radiation and causes ablation efficiency, $\varepsilon_{\text{slope}}$, to decrease. Assuming a tissue density of 1.0 g/cm^3, 4.0 mg/cm^2 equals 40 μm/pulse

energy delivered, considers both the energy invested in ablation and the energy lost to other processes. Figure 15.5 shows that the absolute efficiency changes with radiant exposure: it is zero for radiant exposures less than threshold and rises as the incident energy increases. Thus, care must be taken when interpreting the measurement of absolute efficiency at a single radiant exposure per pulse. The slope efficiency, however, is a singular value descriptive of the ablation process.

For an ablation process that is best fit by a semi-logarithmic relationship between tissue removal and incident radiant exposure, the blow-off model can be used and we can write:

$$\text{Mass loss} = \frac{A\rho}{\mu_a} \ln \frac{H_0}{H_{\text{th}}} \tag{15.9}$$

or

$$\text{Etch depth} = \frac{1}{\mu_a} \ln \frac{H_0}{H_{\text{th}}} \tag{15.10}$$

Equations (15.9) and (15.10) can be derived from Beer's Law, assuming that tissue is removed to the depth to which a threshold amount of energy is deposited. As a simple example, if $\mu_a = 10,000$ cm^{-1}, the spot diameter is 1 mm, $\rho = 1$ g/cm^3, and $H_{\text{th}} = 250$ mJ/cm^2, then for $H_0 = 2$ J/cm^2, a mass of 1.6 μg of tissue is removed per pulse. That is, about 2 μm of tissue are removed per pulse.

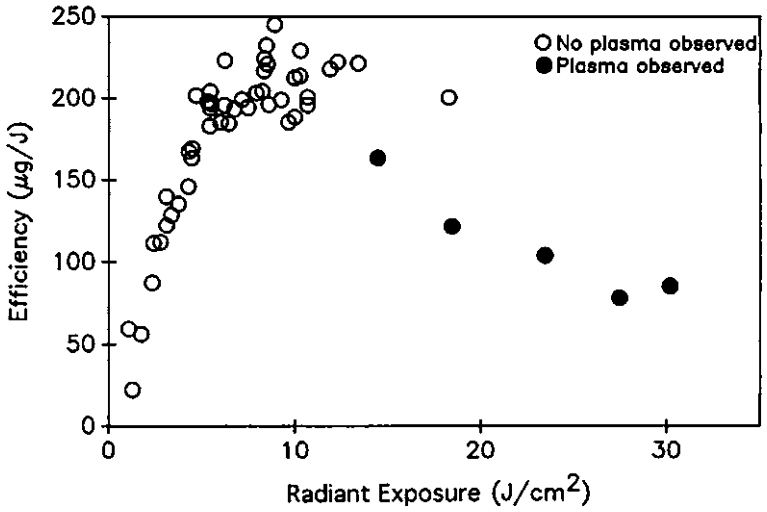


Fig. 15.5 The absolute efficiency of TEA CO_2 laser ablation versus radiant exposure per pulse for guinea pig skin. The maximum efficiency approaches the slope efficiency, 235 μg/J, calculated from the slope of the data shown in Fig. 15.4. A plasma that forms at higher radiant exposures consumed some of the incident radiation and causes ablation efficiency to decrease ($\lambda = 10.6\,\mu m$, $\tau_p = 2\,\mu s$, spot size $\approx 1.6 \times 2.0$ mm rectangular, beam profile approximately flat top) [46]

In many of the aforementioned examples, the experimental ablation depths do not follow the theoretically predicted values. More sophisticated models have been developed that are typically only valid for specific conditions (see Chapter 14). One of the basic assumptions in most models is that the vaporization of a thin layer of water takes place over the entire area of the fiber tip or spot size. Water vaporization threshold experiments as a function of the water temperature with Holmium, Thulium and even Excimer lasers as a function of the temperature of the water, showed that the threshold for vaporization was well below the expected values [47, 48, 83]. The measured thresholds indicated that at threshold only a small fraction of the thin layer is actually vaporized. Physically this means that, instead of vaporizing a layer of water over the entire area underneath the delivery fiber to form the vapor bubbles that are experimentally observed, vaporization takes place at a few nucleation sites. Consequently, the theoretical description of ablation depth is significantly altered (see Fig. 15.6).

A detailed theory of this so-called "partial vaporization" was developed and validated with experimental evidence by van Leeuwen and Jansen [48, 49]. In this model, the threshold for vaporization is determined by the fraction f of the irradiated area that is needed to start the ablation process:

$$H_{\text{th}} = \frac{\rho\,(c\Delta T + fL_v)}{\mu_a} \tag{15.11}$$

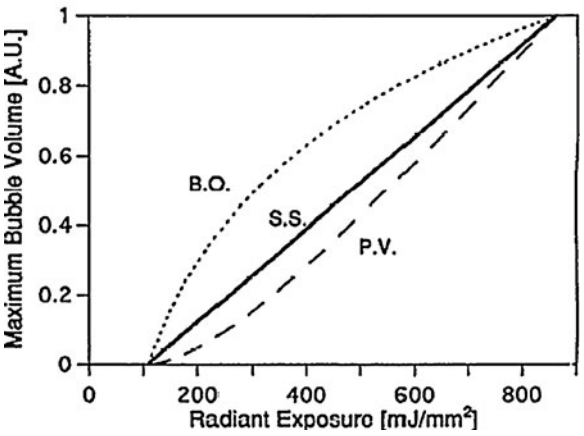


Fig. 15.6 Comparison of theoretical maximum bubble volume (arbitrary units a.u.) for the different models: *Dotted line*: partial vaporization model (P.V.); *Interrupted line*: blow off model (B.O.); *solid line*: steady state model (S.S.). For all models, the threshold radiant exposure for bubble formation was chosen at 8.2 mJ/mm^2 and the bubble volume was set to 1 at 80 mJ/mm^2 [49]

The absorbed energy between $z = 0$ and $z = d$ (at which the threshold energy deposition is reached) is not sufficient to vaporize the whole disk area underneath the fiber. It is hypothesized that the additional radiant energy of the laser pulse above the threshold is used for the partial vaporization of the heated disk and that this fraction f is given by

$$f(z) = \frac{\mu_a (H(z) - H_{th})}{\rho L_v} \tag{15.12}$$

Integrating f from $z = 0$ to $z = d$ leads to the volume of liquid water underneath the fiber tip that is vaporized:

$$\frac{V}{A} = \frac{H_{\text{th}} \left(\frac{H_0}{H_{\text{th}}} - 1 - \ln \left(\frac{H_0}{H_{\text{th}}} \right) \right)}{\rho L_v} \tag{15.13}$$

In this relation, the volume of vaporized tissue is described as a function of the radiant exposure and the threshold radiant exposure. This equation also shows that the vapor volume depends on the absorption coefficient of the medium at the irradiation wavelength since the threshold radiant exposure depends on the absorption coefficient.

In Fig. 15.6 the vaporized volume of water, a measure for the mass loss per pulse, as calculated by the steady state model (Eq. (15.7)), the blow off model (Eq. (15.9)), and the partial vaporization model (Eq. (15.13)) are compared.

The partial vaporization model describes the effect that the threshold radiant exposure is lower than the theoretically expected value given for the full vaporization of the irradiated area. Also, the model predicts that the ablated volume increases more than linearly for lower radiant exposures. Jansen et al. demonstrated that the volume of the induced bubble followed the predicted curve of the partial vaporization model, as depicted in Fig. 15.7 [49].

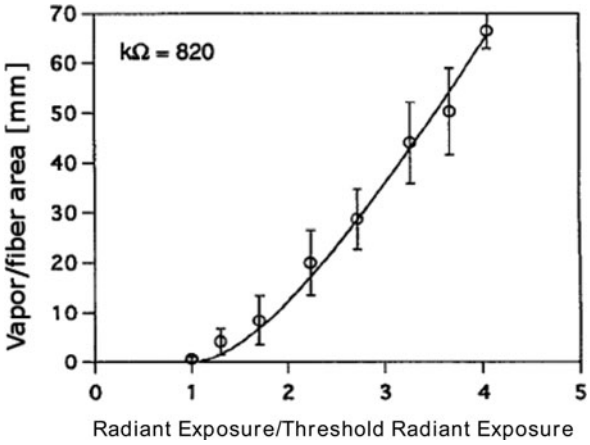


Fig. 15.7 Ratio of the maximum bubble volume to the surface area of the fiber as a function of the ratio of the radiant exposure to the theoretical threshold radiant exposure for bubble formation in water of 22°C. All data points were taken together in bins with an interval of 50 mJ/mm^2. The average and standard deviation is shown. The weighed least squares fit according to Eq. (15.13) is represented by the *solid line*. The fitting parameter, $k\Omega$, was fitted to be 820 ($r = 0.99$). (Reprinted with permission from [49])

In summary, the energy delivered to the tissue surface controls the amount of tissue removed. The removal process can be described by either a linear or semi-logarithmic relationship. There are, however, data that do not so nicely fit either the linear relationship or the semi-logarithmic relationship between mass loss and radiant exposure as described in Chapter 14. For such cases it is currently believed that either the mechanism of ablation changes with incident radiant exposure or the optical properties of the tissue are not constant during the laser pulse. Furthermore, because ablation is an explosive process, tissue mechanical properties play a significant role in the control of ablation.

15.2.4 Spot Size Considerations

The effect of spot size on ablation is the least studied of all irradiation parameters that affect ablation. There are several issues that arise when selecting the spot size. If the spot size is too small, then the aspect ratio of the crater is high. A high aspect ratio can make it difficult for material to travel from the base of an ablation crater to the surface. Further, in the process of traversing a high-aspect-ratio cut, the hot gases and liquids that are generated can heat and damage the crater walls [50].

If the spot size is too large, then the radiant exposure can be too low for effective ablation since the pulse energy from laser sources is inherently limited. If the radiant exposure is sufficient and if the ablation process results in pressure wave formation within the tissue, then for the same radiant exposure the spotsize may have profound effects on the laser-induced pressure waves generated [51].

One of the biggest problems is maintaining the appropriate spot size. In many applications where laser energy is delivered via fiberoptics, the divergence of the beam when it exits the fiber results in a spotsize that depends on the distance between the fiber tip and the target tissue. In particular in endoscopic procedures where depth vision is severely limited due to the inherent 2D nature of endoscopic visualization, this a significant issue. Since the radiant exposure depends on the spot radius squared, even minor changes in spotsize produce significant changes in radiant exposure. On the other hand, beam divergence in a liquid environment is smaller than in air due the index matching between the fiber core and the surrounding medium.

15.2.5 Repetition Rate Considerations

One goal of pulsed laser ablation is to remove the target tissue rapidly without causing undue damage to the non-ablated tissue. Basically, the faster the laser is pulsed, the faster tissue is removed. This is another way of expressing that if the absolute efficiency is pulse repetition rate independent, then the average power determines the removal rate. Ablation rate will not be pulse repetition rate independent if, for example, a pulse arrives before the plume produced by the previous pulse has left the beam path. In such a case, beam energy may be absorbed by the plume and not get to the tissue surface. Such a scenario would be inefficient.

A very rapid pulse repetition rate can also increase the zone of thermal damage compared with that produced by single or low repetition rate pulses produced by the same laser. Equation (15.1) states a relationship between the thermal relaxation time of a heated volume of tissue when the penetration depth is much smaller than the beam spot size. The equation accurately describes the time required for thermal energy to diffuse to tissue beyond the layer initially heated by the optical radiation. More precisely, if a pulse is shorter in duration than the thermal relaxation time, then little thermal diffusion will occur during the pulse and little tissue beyond the layer initially heated thermally damaged during the pulse. The thermal relaxation time, however, is not a good indicator of the time required for the heated layer of tissue to cool (see Chapter 10). Thus, although at first glance it may seem appropriate to use a pulse repetition rate as high as $1/\tau_{th}$, such a repetition rate will lead to increased thermal damage. Because the diffusion of thermal energy out of a heated layer of tissue does not follow an exponential decay but rather a more complicated and slower decay [52], it takes approximately 10 thermal time constants, i.e. $10 \times \tau_{th}$, before the temperature in the initial heated layer drops back to its initial value [52–57]. Further, it is probably more accurate to use the depth of the layer of tissue damaged by a single laser pulse rather than the $1/e$ penetration depth of the incident optical radiation when calculating τ_{th} for use in pulse repetition rate decisions. The $1/e$ penetration depth of the incident radiation is often not well correlated with the depth of the damage zone – a result that has been attributed to a number of causes including forced convection, turbulent flow of liquid layers within the ablation crater [58] as well as dynamic optical properties of tissue [59].

Thus, a "rule of thumb" estimate to avoid increased thermal damage due to accumulation of thermal energy is that the pulse repetition rate should be less than approximately $(10 \times \tau_{\text{th}})^{-1}$, where $\tau_{\text{th}} = z^2/4\alpha$ and z is the thickness of the damage zone induced by a single pulse of laser radiation rather than the $1/e$ penetration depth of the incident radiation [54]. For example, when using an Er:YAG laser emitting 2.94 μm radiation, the expected $1/e$ penetration depth based upon water absorption coefficients is about 1 μm. Thus, the pulse duration should be less than about 2 μsec to minimize thermal diffusion during the laser pulse. The thermal damaged zone produced in soft tissue by a 100 nsec long pulse of 2.94 μm radiation is approximately 15 μm deep at the base and along the side of the ablation crater [60]. This zone of damage is somewhat independent of beam diameter. Thus the rule-of-thumb estimate would indicate that a pulse repetition rate as high as 230 Hz is not likely to increase the zone of thermal damage beyond approximately 15 μm. The extension of thermal damage due to superposition of individual temperature responses has not been seen at repetition rates as high as 30 Hz [54]. Similarly, using a pulsed CO_2 laser, one can assume water is the dominant absorber in a tissue such as skin ($\mu_{a,\text{water}} \approx 840\ \text{cm}^{-1}$, thus $\mu_{a,\text{skin}} \approx 0.7\,\mu_{a,\text{water}} \approx 600\ \text{cm}^{-1}$) and predict that the pulse should be shorter than ≈500 μs in order to confine the thermal energy to the site of absorption. Further, thermal damage measurements taken from 50 μs long ablative pulses of CO_2 laser radiation of skin indicate thermal damage on the order of 100 μm along the crater base and side. A 100 μm wide zone cools in approximately (10 × 20 ms) = 200 ms. Thus, a simple "rule of thumb" approximation indicates that thermal damage is not extended if the pulse repetition rate is less than approximately 5 Hz, a result in good agreement with non-ablative work [52].

One should note that the $(10 \times \tau_z)^{-1}$ estimate is only an estimate. The cooling of a heated layer of tissue is a slow process and more than 10 time constants may be required for the heated layer to sufficiently cool, in particular when the overall exposure time increases, heat diffuses and the temperature gradient between the irradiated tissue and the surrounding tissue areas is reduced.

One effective way to increase the pulse repetition rate (and thus reduce the overall time it takes to deliver the required energy) is to deliver the high repetition rates in a spatially distributed manner using a beam scanner. This way the repetition rate of the laser may be hundreds or even thousands of Hz even though a given volume of tissue experiences only the repetition rate determined by how frequently the scanner delivers a pulse to that volume. Numerous examples of this approach have found their way into clinical practice.

In summary, if the pulse repetition rate is too high then plume absorption may decrease ablation efficiency and cumulative heating of the tissue may increase residual thermal damage. If the pulse repetition rate is too low, then the operative procedure may take too long.

The influence of pulse duration, wavelength, and a series of pulses upon ablation is illustrated in Fig. 15.8. The ablation crater and zone of thermal damage in the cornea using Q-switched and free-running Er:YAG and Er:YSGG lasers is presented in the figure for a sequence of 10 pulses.

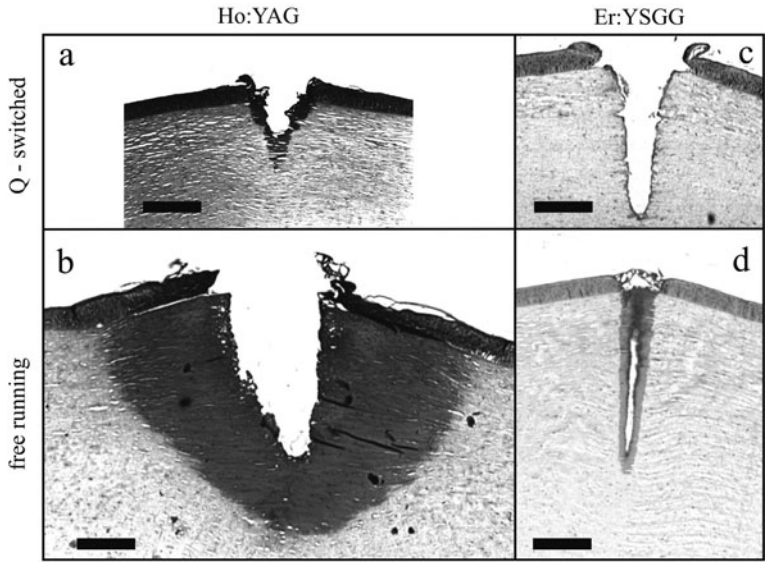


Fig. 15.8 (**a**) Histology of ablation crater performed in cornea with 10 pulses of a Q-switched Ho:YAG laser ($\tau = 40$ ns) and a radiant exposure of $H = 21.3$ J/cm^2. The coagulated zone adjacent to the 178 μm-deep crater is 50–120 μm. (**b**) Microsection of a laser crater produced with 10 free-running Ho:YAG laser pulses ($\tau = 250$ μs, H = 795 J/cm^2). The channel depth is 570 μm and the thermal tissue dmage 300–500 μm. (**c**) Crater in cornea after 10 Q-switched Er:YSGG pulses ($\tau = 40$ ns, $H = 15.9$ J/cm^2). Crater depth = 570 μm. The thermally damaged zone amounts to 4–8 μm. (**d**) Microsection of the cornea after delivery of 10 free-running Er:YSGG pulses ($\tau = 250$ μs, $H = 15.9$ J/cm^2), crater depth = 570 μm, thermal damage 10–50 μm. The fact that the crater diameter is smaller than the spot size (200 μm) is a histological artifact. The laser spot diameter was 200 μm on the cornea surface. Bars = 200 μm. (Reprinted with permission from [61])

15.2.6 Spatial Beam Profile Considerations

To an optical engineer the goal in laser cavity design is usually to create a beam that has a Gaussian profile. However, for the ablation of tissue a flat-top profile is more desirable. A flat-top beam ablates tissue leaving cleaner boundaries because there are no regions of sub-ablative energy in the beam wings. More importantly, a flat-top beam creates a generally flat-bottomed ablation crater; a Gaussian beam typically creates a crater that has a deep center region. The creation of flat-top beams directly out of a laser is generally not trivial, however; after the beam has traversed a multi-mode optical fiber, or a bundle of fibers, the spatial modes are usually well mixed. Getting a flat-top beam to the tissue surface is not a major technological problem, but if overlooked can compromise the efficacy of the system (see Chapter 17). It is also worthwhile to point out that calculation of radiant exposure in a laser beam is strongly dependent on the beam profile. It is standard practice to measure the output from a laser or fiberoptic using a energy meter and to measure the spotsize (using

a knife edge method, beam profiler or assuming the fiber diameter as the spotsize). Subsequently, the radiant exposure is calculated by dividing the pulse energy by the spotsize (area). This, however, yields the average radiant exposure over the spot. If this approach is used for a Gaussian beam, the resulting radiant exposure represents the average radiant exposure over the area (based on the $1/e^2$ radius of the beam). In reality the radiant exposure in the center of the beam (i.e. at $r = 0$) is twice as high as the average radiant exposure over the laser spot. Thus when using the radiant exposure, H_0, in other calculations that yield crater depth or ablation efficiency, one has to take care to consider the spatial distribution of the laser energy over the irradiated spot.

15.2.7 Fluid Layer

The ablation of hard and soft tissue can be enhanced by placing a thin layer of fluid over the tissue. Typically water is suitable for wavelengths below 2.0 μm, and infrared transmitting fluids such as perfluorocarbon work well for wavelengths from 2 to 7 μm [62]. Kang et al. [63, 64] have compared the importance of a fluid layer upon the ablation of bovine bone tibia using a long pulsed (150 μs) Er,Cr:YSGG ($\lambda = 2.79$ μm) laser. Ablation results for direct irradiation of the bone were compared to results when the bone was covered with 500 μm layer of perfluorocarbon or water or when using a simultaneous water spray at a flow rate of 8 ml/min. The crater for a single laser pulse and for osteotomy cuts are illustrated in Fig. 15.9. Direct irradiation (no fluid layer) and irradiation with a layer of fluid ablated approximately the same volume of tissue for radiant exposures of 45 J/cm^2 per pulse. Direct ablation produced severe carbonization while slight carbonization was associated with the perfluorocarbon layer; whereas no carbonization occurred when the bone was covered with a thin layer (500 μm) of water. Best results in terms of ablation volume and minimizing carbonization occurred with water spray assisted laser ablation. The attenuated laser irradiation reaching the bone was approximately equal for the water layer and spray conditions. In both cases a portion of the laser pulse was needed to vaporize the water that has a room temperature absorption coefficient of 530 mm^{-1} which is in sharp contrast to $\mu_a = 0.05$ mm^{-1} for the perfluorocarbon at $\lambda = 2.79$ μm. The low thermal conductivity [$k = 0.067$ W/m K] of the perfluorocarbon was likely responsible for the slight carbonization noted in Fig. 15.9c. During water assisted ablation the fluid provides cooling, and thermo-mechanical coupling enhances ablation performance.

The dynamics of water assisted ablation is more pronounced when using ns pulses. The results of Q-switched Nd:YAG (30 ns, $\lambda = 1064$ nm) and 1 ns, 800 nm Ti:Sapphire laser irradiation upon bone and aluminum, respectively, covered with 500 μm water clearly demonstrates enhanced ablation relative to irradiation of a dry surface [63, 64]. Irradiances above 10^8 W/cm^2 led to the formation of a plasma at the target-water interface. Photomechanical effects in association with explosive liquid vaporization and plasma confinement accounted for lowered ablation

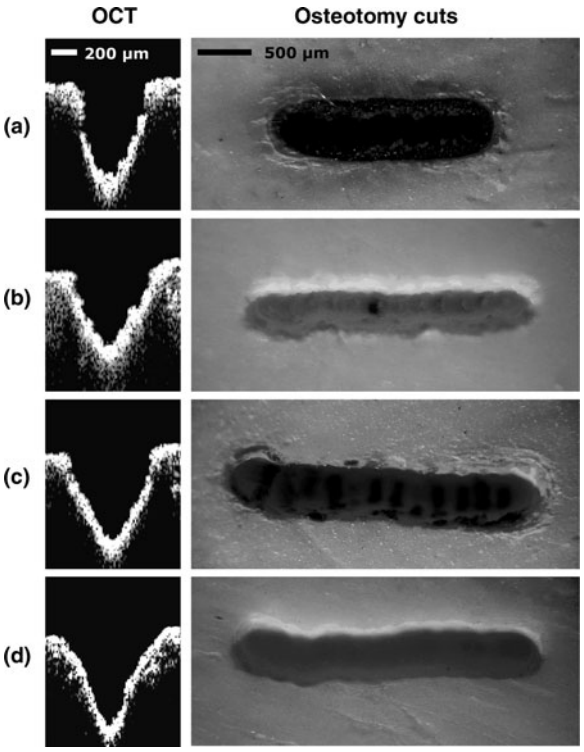


Fig. 15.9 Cross-sectional OCT (*left column*) and *top view* (*right column*) showing osteotomy cuts with multiple sequences of five pulses at $H = 47$ J.cm^2. Four different conditions were tested: (**a**) dry, (**b**) wet (500 μm water layer), (**c**) wet (500 μm perfluorocarbon layer), and (**d**) water spray ablation (flow rate of 8 ml/min)

threshold and facilitated ablation performance. Bubble formation and collapse was considered responsible for the pronounced mechanical effect that enhanced ablation with a fluid layer [64].

15.3 The Effect of Dynamic Changes in Optical Absorption

All previous models for ablation have assumed either a static absorption coefficient during the entire laser pulse, i.e. no tissue is removed during the pulse and the absorption coefficient is constant during the pulse, or the absorption coefficient goes to zero once the heat of ablation is delivered to a particular tissue volume, i.e. the tissue is immediately removed once the heat of ablation is deposited in the tissue [65]. Yet, data indicate that in the IR [66–72] and UV [73, 74] the absorption coefficient is not static during a laser pulse. The dynamic optical properties in the IR were recognized long ago [75] and are well understood and elaborated upon below. The dynamic optical properties in the UV have been less intensely investigated. Nevertheless, although an explanation for the changes is largely lacking, the data available to date are compelling and discussed below.

Mid-infrared laser-tissue interactions have long been predicated upon the low-intensity water absorption spectrum [39, 51, 60, 76, 77] (see Fig. 15.10) [39].

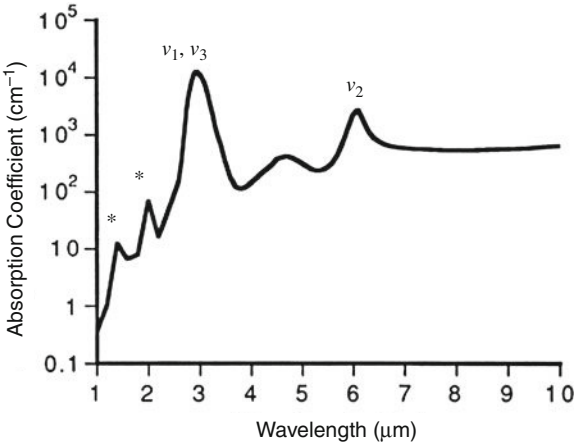


Fig. 15.10 The low-irradiance absorption spectrum of water in the infrared. The symmetric and asymmetric stretch, ν_1 and ν_3, the bending mode, ν_2, and the combination and overtones bands, $*$, are noted. (Reprinted with permission from [39])

Considerable effort by engineers, scientists, clinicians, and industrial personnel has gone into the development of mid-IR emitting lasers for medical and surgical use. All development efforts have been based on the water absorption spectrum shown in Figs. 15.10 and 14.2a. However, the low-intensity absorption spectrum of water does not describe the absorption of radiation during high-power laser irradiation of tissue as discussed in Chapter 9. In this section, the results of mid-infrared laser ablation of tissue will be reviewed first. Particular emphasis will be placed upon results that are disparate with theoretical expectations. Subsequently, the infrared spectroscopy of water, presumably the tissue constituent that controls the optical properties of tissue in the mid-infrared, will be discussed. Given the spectroscopy of water, a model for tissue ablation based upon a dynamically changing absorption coefficient is presented. This ablation model has lead to predictions that have been tested using Er:YAG and Er:YSGG lasers as well as Ho:YAG and Tm:YAG lasers. The results indicate that consideration of dynamic changes in the absorption coefficient of water will drastically alter the understanding and interpretation of laser-tissue interactions.

15.3.1 The Moses Effect

Perhaps the most obvious example of the optical properties not being constant is when water turns into water vapor, e.g. during laser cavitation (bubble formation) and it becomes largely transparent to the infrared radiation (see Fig. 15.2). The reason for this is in all likelihood two-fold: first the reduced density of water molecules contributes to this; and second, as will be discussed in more detail later, the absorption of the water molecules changes with temperature. When the laser pulse duration is on the same order as the bubble lifetime, the first part of the laser pulse is used to vaporize the (tissue) water, while energy delivered later in the laser pulse interacts

not with (tissue) water in the liquid state but rather with water vapor with a significantly reduced absorption coefficient. Hence this laser light traverses the vapor cavity with relative ease and is deposited at the vapor-liquid interface. This so-called "Moses-effect" [78] is responsible for the much deeper than expected effective penetration of laser radiation when using free running pulsed IR lasers such as the Ho:YAG or Er:YAG. This effect is also harnassed in situations where the delivery fiber tip is not or cannot be in contact with the target tissue; in this case the Moses effect allows the creation of a vapor channel between the fiber and the target tissue that allows transmission of energy to the target tissue. As an example, the use of the Ho:YAG laser for cartilage reshaping in the knee joint filled with synovial fluid is effective largely because the 350 μs long laser pulse facilitates the creation of a vapor channel in the synovial fluid such that the latter part of the laser pulse reaches the cartilage even though the fiber may be more than 1 mm from the target tissue and the penetration depth is approximately 350 μm (which would otherwise result in over 99% of the light being attenuated) [84].

The deposition of additional energy inside the vapor bubble gives rise to the typical elongated pear shape of these vapor cavities (see Fig. 15.2). The asymmetry in bubble formation results in an asymmetric collapse (or multiple smaller collapse sites) and thus reduces the collapse induced pressure transients that are so prominent during the collapse of perfectly spherical bubbles. The dynamic changes in the optical properties associated with phase changes are expected and obvious. However, even in the liquid state the optical properties of water are not constant, a topic to which we now turn.

15.3.2 IR Spectroscopy of Water

For more than three decades the laser-medicine community has embarked upon a substantial effort to measure and model the propagation of light in tissue. In the infrared region of the spectrum all data taken to date indicates that the optical properties of tissue are dominated by water absorption. Scattering in the infrared, in particular the mid-infrared, is assumed negligible although at certain wavelengths that show moderately high absorption coefficients, such as the Ho:YAG, this assumption may not be entirely accurate. The data taken by a number of researchers and carefully compiled in a review article by Hale and Querry [39] have been quoted often and are shown graphically in Chapter 14, Fig. 14.2a. The Hale and Querry paper tabulated the optical absorption coefficients of room temperature, atmospheric pressure water as a function of wavelength.

As early as 1909, however, it was recognized that the infrared absorption properties of water were temperature dependent [75]. Figure 15.11 shows a graphical presentation of the temperature dependence of the absorption in the IR [79].

Over the past decades, numerous researchers have documented the effect of temperature, pressure, and solute-solvent interactions on the optical properties of water. The following is a brief summary of the relevant spectroscopy of water.

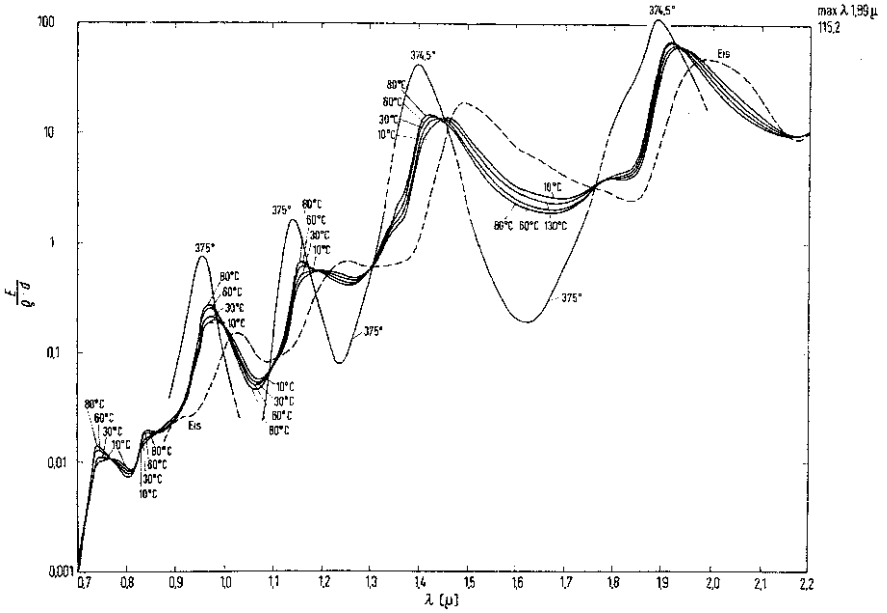


Fig. 15.11 The absorption spectrum of water in the infrared as a function of temperature at constant volume. (Reprinted with permission from [79])

Water has three IR-active vibrational modes: (v_1 is the symmetric stretch, v_2 is the symmetric bend, and v_3 is the asymmetric stretch). The assignment of the fundamental modes as well as the overtones (e.g. $2v_2$) and combinations (e.g. $v_1 + v_2$) is illustrated in Fig. 15.10. The location of each band in Fig. 15.11 is dependent upon the fundamental vibrational frequency of the bend and/or stretch (i.e. the oscillation) of the OH bonds in the water molecule. The intensity of a band is based upon the dipole moment of the oscillation. Hydrogen bonding complicates the harmonic oscillation of the OH bonds. Hydrogen bonding exists because water is a polar molecule with a permanent dipole moment – that is, the oxygen is electro-negative and the hydrogens are electro-positive. Thus, the hydrogen atoms of one molecule "bond" with the oxygen atoms of another molecule. Water is IR-active because it has a permanent dipole moment, p, that changes when the bond length, r, changes. The frequency of dp/dr is in the IR region of the electromagnetic spectrum and causes water to absorb or emit in the IR.

Any process that affects hydrogen bonding will affect the dipole moment and thus the absorption spectrum of water. Hydrogen bonding in water is affected by pressure, temperature, and the addition of "impurities" such as salts. Figure 15.11 shows the effect of temperature on the infrared spectrum of water. Note that as the temperature increases, i.e. as hydrogen bonds are weakened, the absorption bands in this part of the spectrum shift toward lower wavelengths (higher wavenumbers) and change in intensity.

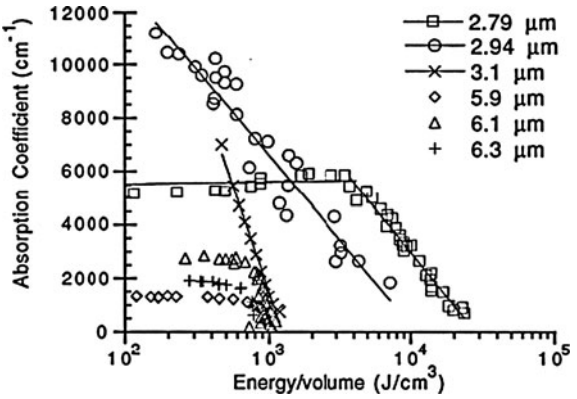


Fig. 15.12 The absorption coefficient of water as a function of energy deposited per unit volume for various infrared wavelengths. (The data near the 6 μm bending-mode peak of water are unpublished from Cummings and Walsh; the data near the 3 μm stretching-modes peak of water are from [72])

In a series of experiments, Vodopyanov [71, 72] showed that the 2.94 μm absorption coefficient of water decreases monotonically with increasing energy density. The 2.79 μm absorption coefficient, however, increased slightly for small rises in energy density then decreased as the energy density rises further. Figure 15.12 summarizes Vodopyanov's results. Note that for energy density greater than 2000 J/cm^3 the absorption coefficient at 2.79 μm is greater than at 2.94 μm. Given that more than 2000 J/cm^3 must be deposited to achieve ablation, the spectroscopic data indicate that during the ablative pulse the effective absorption coefficient for Er:YSGG ($\lambda = 2.79$ μm) laser radiation will be greater than that for Er:YAG ($\lambda = 2.94$ μm) laser radiation – an opposite conclusion would be reached by considering the low-intensity water spectrum.

In summary, the data indicate that the 2.94 μm band amplitude decreases and the peak shifts toward lower wavelengths with increasing energy deposition. Thus, whereas at the 2.94 μm peak of the band the absorption decreases monotonically with increasing energy density; at wavelengths shorter than 2.94 μm the absorption coefficient may rise with rising energy density as the peak of the band moves toward the visible (i.e. a blue shift). When the energy deposited per unit volume gets high enough, the integrated intensity decrease dominates over the peak shift and the absorption coefficient decreases at all wavelengths in the 2.94 μm band.

Thus far the discussion of water spectroscopy has concentrated on the 2.9 μm band. There are considerable data on the effect of temperature on the other water absorption bands. In brief, the 1.4 and 1.9 μm bands shift to lower wavelengths as the temperature is increased. Some data indicate that absorption near the 1.4 and 1.9 μm bands *increase* with temperature [79]. The 6 μm band, which is a fundamental due to the symmetrical bend of water, appears to shift to slightly longer wavelengths with increasing temperature or energy density. The absorption near the 6 μm band may increase with temperature, but good data in this spectral region are sparse [67, 80]. It is *vital* to note that the spectroscopic data are taken at known temperature and pressure, and no data are available above the critical temperature of water, $T_{\text{critical, water}} = 647$ K. During ablation, however, the critical temperature is undoubtedly exceeded ($T_{\text{estimated}}$ can be greater than 7000 K)

and the pressures can be tremendous ($P_{\text{estimated}}$ can be greater than 10^3 atm). Thus, although measurements of the water spectra as a function of temperature and pressure are vital to the understanding of the properties of water, the measurement of the spectra as a function of energy deposited per unit volume are more pertinent to the understanding of laser-tissue interactions.

15.3.3 Modeling Dynamic Shifts in Absorption

In medical applications, a change in absorption coefficient during the laser pulse may lead to changes in the energy distribution and thus in the heat source term. Consequently, the generated temperature at a particular location will be affected, which, in turn, may also have an impact on the ablation process and on the rate reactions that govern the damage processes. Several models have been developed to calculate the laser induced temperature fields using iterative approaches that allow for dynamic changes in absorption. In order to predict the effects of a changing absorption coefficient on parameters like ablation threshold and thermal damage, a model is needed that incorporates these dynamic properties.

Jansen et al. showed the results of a model that assumed a linear relationship between the absorption coefficient and the temperature increase [69]:

$$\mu_a\,(T,z) \;=\; \mu_{a,0}\,(1 \;+\; \beta\rho c\Delta T(z)) \tag{15.14}$$

where $\mu_a(T,z)$ is the temperature dependent absorption coefficient (mm^{-1}), $\mu_{a,0}$ is the absorption coefficient at the starting temperature (i.e. room temperature) (mm^{-1}), β is a constant (mm^3/mJ), ρ is the density (mg/mm^3), c is the heat capacity (mJ/mg °C), and $\Delta T(z)$(°C) represents the one-dimensional temperature distribution:

$$\Delta T\,(z) \;=\; \frac{\mu_a Q_0 e^{-\mu_a z}}{\rho c A} \tag{15.15}$$

where μ_a is the absorption coefficient (mm^{-1}), and z is the depth (mm), A is the area (mm^2), Q_0 is the incident energy (mJ). The one-dimensional model assumed that during the delivery of a pulse of energy there was no heat conduction and bubble formation did not occur (i.e. the pulse is sufficiently short <20 μs). It was also assumed that the radial beam profile was flat and no divergence of the laser beam took place in the water. An iterative technique was used, where a small fraction of the energy was deposited and the temperature distribution along the z-axis was calculated. Using the experimentally determined values in Table 15.2 and Eq. (15.14), the values for μ_a along the z-axis were calculated and the next fraction of energy was delivered. The temperature rise due to this was calculated with the new $\mu_a(T)$. This temperature rise was added to the previous temperature rise and a new $\mu_a(T)$ was calculated. The model uses time steps of 5 ns, assuming a total pulse length of 1 μs and the distance steps were 1 μm. The one-dimensional temperature distribution calculated assuming a constant absorption coefficient and a temperature dependent

Table 15.2 Values for $\mu_{a,o}$ and β at 2.12, 2.09, and 2.01 μm wavelength, derived form the linear curve fit of spectrophotometer data

	2.12 μm	2.09 μm	2.01 μm
$\mu_{a,o}(\text{mm}^{-1})$	2.982	3.703	7.507
$\beta\,(\text{mm}^3/\text{mJ})$	−0.0005278	−0.0008391	−0.0010172
R^2	0.97	0.99	0.99

Note that the $\mu_{a,o}$ values are given for room temperature (22°C). For the model it is assumed that at room temperature the temperature rise $\Delta T = 0°\text{C}$
Adapted from [69]

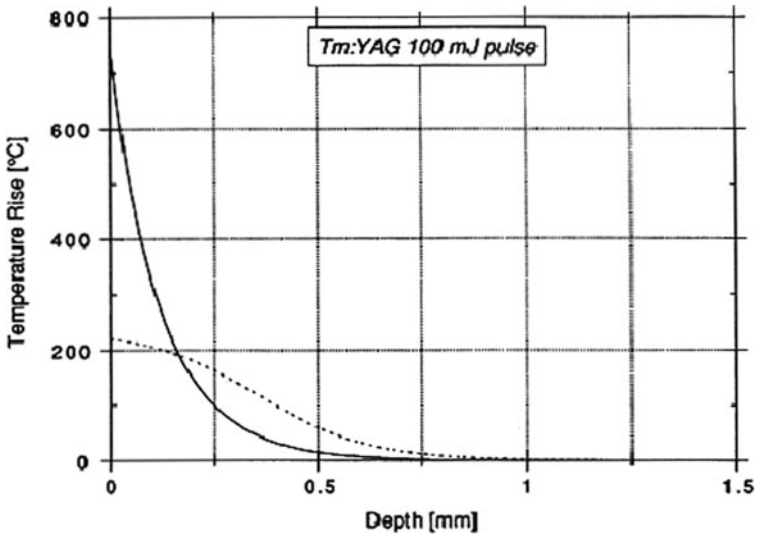


Fig. 15.13 Temperature distribution in water along the z axis with constant (*solid line*) and dynamic μ_a (*dotted line*) used in the numerical model for a 100 mJ/pulse (353 mJ/mm^2). Thulium:YAG pulse (2.01 μm). It is assumed that there is no heat conduction, no bubble formation, no beam divergence, and a block pulse of laser energy going in. The model also assumes a 600 μm diameter spot size. (Adapted from [69])

absorption coefficient for a 100 mJ/pulse (353 mJ/mm^2) Tm:YAG laser pulse is shown in Fig. 15.13.

A similar approach was used by Walsh et al. to model the effect of dymanic changes in absorption around the 2.9 μm absorption peak. Their comparison of the depth of energy deposition for Er:YAG and Er:YSGG laser radiation in tissue (water) is shown in Fig. 15.14. Note that Er:YAG laser radiation penetrates more deeply within tissue than Er:YSGG laser radiation. The model leads to two predictions that can, and have been, critically tested: (1) Er:YAG laser radiation should cut more deeply than Er:YSGG laser radiation; and (2) Er:YSGG laser radiation should induce a shock wave with a higher velocity than Er:YAG laser radiation. Opposite

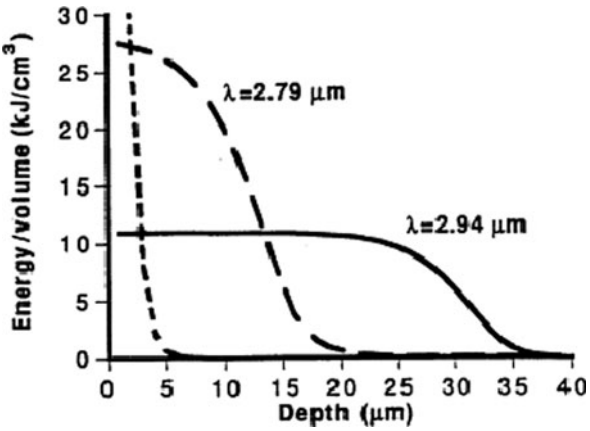


Fig. 15.14 Results of dynamic absorption coefficient model for 2.79 and 2.94 μm large-dashed line and solid line, respectively. For comparison, the low-irradiance, constant-absorption coefficient of water a 2.94 μm would lead to energy deposition as shown by the short-dashed line. Incident radiant exposure = 30 J/cm^2

predictions would be made if only the low-irradiance water absorption curve was considered [81].

15.3.4 The Dynamic UV Optical Properties of Tissue

In contrast to mid-IR laser ablation, 193 nm ArF excimer laser cuts are typically shallower in depth than the predicted 1/*e* absorption depth of the incident optical radiation [73]. Similarly, the zone of thermal damage produced during 193 nm excimer laser ablation of tissue is less than the 1/*e* penetration depth of the 193 nm radiation [40]. These results can exist either if the low-irradiance measurements of the optical absorption coefficient of tissue at 193 nm is incorrectly low (which is a possibility given that only one quantitative measure of this property exists in the literature [40]) or if the optical attenuation increases during the excimer laser pulse.

The veracity of the low-irradiance absorption coefficient must await corroboration. Data exist, however, indicating that the attenuation of 193 nm radiation through thin collagen films and corneal sections increases during the excimer laser pulse (see Fig. 15.15). In brief, the temporal profile of the incident and transmitted laser pulse have been recorded simultaneously. The results indicate that initially the temporal profile of the incident and transmitted pulses are identical; but about 10 ns after the beginning of the pulse, the transmitted pulse is attenuated compared with the incident pulse. The data indicate the attenuation of the incident 193 nm radiation increases during the laser pulse. The authors report that this attenuation is dependent upon the hydration of the tissue and speculate that the mechanism for the attenuation is either ablation induced scattering [74], photochemical bleaching [73] or photolytic annihilation [73]. Independent of the mechanisms of the induced optical property changes, it is apparent that a complete model and understanding of UV-laser ablation will need to include consideration of dynamic optical properties of either the tissue or the ablation plume (Fig. 15.15).

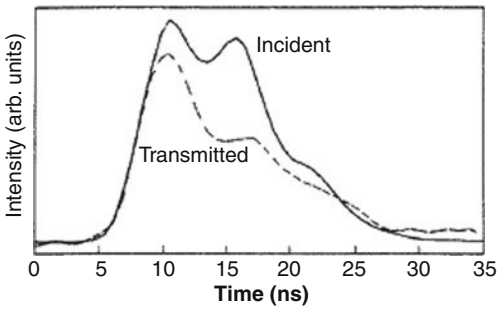


Fig. 15.15 Transmission of 193-nm radiation through a thin section of bovine cornea. Note the transmitted pulse is truncated indicating that either the absorption or scattering of the incident radiation increases during the ≈ 20-ns-long laser pulse. (Reprinted with permission from [81])

15.3.5 Implications of UV and IR Dynamic Optical Properties

In summary, the heat source term in all laser tissue interactions is critically dependent upon knowledge of the optical properties of the tissue. Considerable effort has been expended to quantify tissue optical properties using low irradiance, spectrophotometer-based techniques. Such techniques, however, can yield data that do not accurately indicate the optical properties of the tissue during high-irradiance. In brief, the high-irradiance optical properties should be measured so that the source term in the laser tissue interactions problem can be determined accurately. Dynamic changes in tissue optical properties should always be considered if the results of a laser-tissue interaction do not correlate well with a reasonably conceived model of the interaction.

15.4 Summary

The process of ultraviolet and mid-infrared pulsed laser ablation of soft tissue causes the formation of vapor bubbles (see Chapter 14) and dynamic changes in optical properties (see Chapter 9). Fast expanding and imploding water vapor bubbles produced by the transfer of absorbed laser energy to the surrounding (tissue) water is likely the mechanical cause of dissections observed in the crater wall after ablation. The extent of the dissections is much larger than of the thermally induced necrosis.

For the laser parameters of the XeCl excimer and mid-infrared holmium:YAG lasers, thermal confinement but no stress confinement is expected. Consequently, we assume the damage effects of laser induced stress waves to be minimal with respect to the damage induced by the expanding bubble. Thus, during pulsed laser ablation mechanical damage will be produced by similar mechanisms: (a) bubble expansion within the target tissue, and (b) bubble expansion and implosion within the immersion fluid. These two mechanisms plus the mechanical (confinement) properties of tissue will greatly influence both the ablation process and the mechanical collateral damage.

Furthermore, the threshold radiant exposure for bubble formation agrees with a temperature rise up to 100°C. Due to the virtual absence of the latent heat of vaporization, the threshold radiant exposure is much smaller than theoretically expected. This is due to the formation of small bubbles at nucleation sites. Thus, at threshold radiant exposure, only a fraction of the water layer in front of the fiber tip is actually vaporized. The partial vaporization model provides an accurate description of the ablation process that takes this into account.

The mid-infrared radiation absorption by water was shown to change due to absorbed energy. Before vaporization takes place, the absorption of Holmium, Thulium and Erbium:YAG laser light by liquid water decreases as the temperature of the water increases. For the relatively long mid-infrared (holmium and thulium) laser pulses, the bubble formation (i.e. transition from liquid water into water vapor) results in an up to 8-fold increased penetration depth compared to the theoretically expected penetration depth. As a water vapor bubble is formed, the mid-infrared radiation is transmitted through the water vapor which is virtually transparent to these wavelengths. These phenomena are important for modeling and dosimetry calculations of pulsed-laser tissue interactions.

References

1. Welch AJ, Motamedi M, Rastager S, LeCarpentier GL, and Jansen D. Laser thermal ablation. *Photochem. Photobiol.*, 53:815–823 (1991).
2. Walsh JT Jr, Flotte TJ, and Deutsch TF. Er:YAG laser ablation of tissue: Effect of pulse duration and tissue type on thermal damage. *Lasers Surg. Med.*, 9:314–326 (1989).
3. Nelson JS, Yow L, Liaw LH, Macleay L, Zavar RB, Orenstein A, Wright WH, Andres JJ, and Berns MW. Ablation of bone and methacrylate by a prototype mid-infrared erbium:YAG laser. *Lasers Surg. Med.*, 8:494–500 (1988).
4. Watanabe S, Flotte TJ, McAuliffe DJ, and Jacques SL. Putative photoacoustic damage in skin induced by pulsed ArF excimer laser. *J. Invest. Dermatol.*, 90:761–766 (1988).
5. Jacques SL. Laser-tissue interactions: Photochemical, photothermal and photomechanical. *Surg. Clin. North Am.*, 72:531–559 (1992).
6. McKenzie AL. Physics of thermal processes in laser-tissue interactions. *Phys. Med. Biol.*, 35:1175–1209 (1990).
7. van Gemert MJC and Welch AJ. Time constants in thermal laser medicine. *Lasers Surg. Med.*, 9:405–421 (1989).
8. Vogel A, Schweiger P, Frieser A, Asiyo MN, and Birngruber R. Intraocular Nd:YAG laser surgery: Light-tissue interaction, damage range, and reduction of collateral effects. *IEEE J. Quant. Electron.*, 26:2240–2260 (1990).
9. Yashima Y, McAulife DJ, Jacques SL, and Flotte TJ. Laser-induced photoacoustic injury of skin: Effect of inertial confinement. *Lasers Surg. Med.*, 11:62–68 (1991).
10. Paltauf G, Reichel E, and Schmidt-Kloiber H. Study of different ablation models by use of high-speed sampling photography. In: Jacques SL and Katzir A (eds) *Laser-Tissue Interaction III*, Vol. 1646. SPIE, Bellingham, pp. 343–352 (1992).
11. LeCarpentier GL, Motamedi M, McMath LP, Rastegar S, and Welch AJ. Continuous wave laser ablation of tissue: Analysis of thermal and mechanical events. *IEEE Trans. Biomed. Eng.*, 40:188–200 (1993).
12. Verdaasdonk RM, Borst C, and van Gemert MJC. Explosive onset of continuous wave tissue ablation. *Phys. Med. Biol.*, 35:1129–1144 (1990).

13. Eisberg R and Resnick R. Molecules. In: *Quantum Physics of Atoms, Molecules, Solids, Nuclei and Particles*. Wiley, New York, pp. 451–479 (1974).
14. Morcos NC, Berns M, and Henry WL. Phycocyanin: Laser activation, cytotoxic effects, and uptake in human atherosclerotic plaque. *Lasers Surg. Med.*, 8:10–17 (1988).
15. Ortu P, LaMuraglia GM, Roberts WG, Flotte TJ, and Hasan T. Photodynamic therapy of arteries: A novel approach for treatment of experimental intimal hyperplasia. *Circulation*, 85:1189–1196 (1992)
16. Linsker R, Srinivasan R, Wynne JJ, and Alonso DR. Far-ultraviolet laser ablation of atherosclerotic lesions. *Lasers Surg. Med.*, 4:201–206 (1984).
17. Srinivasan R and Leigh WJ. Ablative photodecomposition: Action of far-ultraviolet (193 nm) laser radiation on poly(ethylene therethalate) films. *J. Am. Chem. Soc.*, 104:6784–6785 (1982).
18. Izatt JA, Albagli D, Itzkan I, and Feld MS. Pulsed laser ablation of calcified tissue: Physical mechanisms and fundamental parameters. In: Jacques SL, Katzir A (eds) *Laser-Tissue Interaction I*, Vol. 1202. SPIE, Bellingham, pp. 133–140 (1990).
19. Flotte TJ, Yashima Y, Watanabe S, McAuliffe DJ, and Jacques SL. Morphological studies of laser-induced photoacoustic damage. In: Jacques SL, Katzir A (eds) *Laser-Tissue Interaction I*, Vol. 1202. SPIE, Bellingham, pp. 71–77 (1990).
20. Dyer PE and Al-Dhahir RK. Transient photoacoustic studies of laser tissue ablation. In: Jacques SL and Katzir A (eds) *Laser-Tissue Interaction I*, Vol. 1202. SPIE, Bellingham, pp. 46–60 (1990).
21. Dingus RS and Shafer BP. Laser-induced shock wave effects in materials. In: Jacques SL and Katzir A (eds) *Laser-Tissue Interaction I*, Vol. 1202. SPIE, Bellingham, pp. 36–45 (1990).
22. Doukas AG, Birngruber R, and Deutsch TF. Determination of the shock wave pressures generated by laser-induced breakdown in water. In: Jacques SL and Katzir A (eds) *Laser-Tissue Interaction I*, Vol. 1202. SPIE, Bellingham, pp. 61–70 (1990).
23. Thomsen SL, Pearce JA, and Cheong WF. Changes in birefringence as markers of thermal damage in tissues. *IEEE Trans. Biomed. Eng.*, 36:1174–1179 (1989).
24. Thomsen S. Pathologic analysis of photothermal and photomechanical effects of laser-tissue interactions. *Photochem. Photobiol.*, 53:825–835 (1991).
25. Lord Rayleigh OM. On the pressure developed in a liquid during the collapse of a spherical cavity. *Phil. Mag.*, 34:94–98 (1917).
26. Izatt JA, Albagli D, Britton M, Jubas JM, Itzkan I, and Feld MS. Wavelength dependence of pulsed laser ablation of calcified tissue. *Lasers Surg. Med.*, 11:238–249 (1991).
27. Zysset B, Fujimoto JG, and Deutsch TF. Time-resolved measurements of picosecond optical breakdown. *Appl. Phys. B*, 48:139–147 (1989).
28. van Leeuwen TG, Veen JM van der, Verdaasdonk RM, and Borst C. Non-contact tissue ablation by holmium: YSGG laser pulses in blood. *Lasers Surg. Med.*, 11:26–34 (1991).
29. Vogel A, Hentschel W, Holzfuss J, and Lauterborn W. Cavitation bubble dynamics and acoustic transient generation in ocular surgery with pulsed neodymium:YAG lasers. *Ophthalmology*, 93:1259–1269 (1986).
30. van Leeuwen TG, van Erven L, Meertens JH, Motamedi M, Post JM, and Borst C. Origin of arterial wall dissections induced by pulsed excimer and mid-infrared laser ablation in the pig. *J. Am Coll. Cardiol*., 19:1610–1618 (1992)
31. Cleary SF. Laser pulses and the generation of acoustic transients in biological material. In: Wolbarsht ML (ed) *Lasers Applications in Medicine and Biology*, Vol. 3. Plenum, New York, pp. 175–219 (1977).
32. de la Torre R and Gregory KW. Cavitation bubbles and acoustic transients may produce dissections during laser angioplasty. *J. Am. Coll. Cardiol.*, 19:48A (1992).
33. Bonner RF, Smith PD, Prevosti LG, Bartorellil A, Almagor Y, and Leon MB. Laser sources for angioplasty. In: Abela GS (ed) *Lasers in Cardiovascular Medicine and Surgery: Fundamentals and Techniques*. Kluwer Academic, Boston, Dordrecht, London, pp. 31–44 (1990).

34. Tomaru T, Geschwind HJ, Boussignac G, Lange F, and Tahk SJ. Comparison of ablation efficacy of excimer, pulsed-dye, and holmium-YAG lasers relevant to shock waves. *Am. Heart J.*, 123:886–895 (1992).
35. Tomaru T, Geschwind HJ, Boussignac G, Lange F, and Tahk SJ. Characteristics of shock waves induced by pulsed lasers and their effects on arterial tissue: Comparison of excimer, pulse dye, and holmium YAG lasers. *Am. Heart J.*, 123:896–904 (1992).
36. Ward B and Emmony DC. Conservation of energy in the oscillations of laser-induced cavitation bubbles. *J. Acoust. Soc. Am.*, 88:434–441 (1990).
37. Vogel A, Busch S, Jungnickel K, and Birngruber R. Mechanism of intraocular photodisruption with picosecond and nanosecond laser pulses. *Lasers Surg. Med.*, 15:32–43 (1994).
38. Flotte TJ, Frisoli JK, Boetschkes M, and Doukas AG. Laser-induced shock wave effects on red blood cells. In: Jacques AL and Katzir A (eds) *Laser-Tissue Interaction II*, Vol. 1427. SPIE, Bellingham, pp. 36–44 (1991).
39. Hale GM and Querry MR. Optical constants of water in the 200-nm to 200-μm wavelength region. *Appl. Opt.*, 12:555–563 (1973).
40. Puliafito CA, Steinert RF, Deutsch TF, Hillenkamp F, Dehm EJ, and Adler CM. Excimer laser ablation of the cornea and lens. *Ophthalmology*, 92:741–748 (1985).
41. Walsh JT, Flotte TJ, Anderson RR, and Deutsch TF. Pulsed CO_2 laser ablation: Effect of tissue type and pulse duration on thermal damage. *Lasers Surg. Med.*, 8:108–118 (1988).
42. Srinivasan R, Dyler RE, and Braren B. Far-UV laser ablation of cornea; photoacoustic studies. *Lasers Surg. Med.*, 6:524–527 (1984).
43. Srinivasan R. Ablation of polymers and biological tissue by ultraviolet lasers. *Science* 234:559–565 (1986).
44. Walsh JT and Deutsch TF. Measurement of Er:YAG laser ablation plume dynamics. *Appl. Phys. B*, 52:217–224 (1991).
45. Domankevitz Y and Nishioka NS. Reflection during pulsed holmium laser irradiation of tissue. *Lasers Surg. Med.*, 9:375–381 (1989).
46. Walsh JT and Deutsch TF. Pulsed CO_2 laser tissue ablation: Measurement of the ablation rate. *Lasers Surg. Med.*, 8:264–275 (1988).
47. Jansen ED, van Leeuwen TG, Motamedi M, Borst C, and Welch AJ. Temperature dependence of the absorption coefficient of water for midinfrared laser radiation. *Lasers Surg. Med.*, 14:258–268 (1994).
48. van Leeuwen TG, Jansen ED, Welch AJ, and Borst C. Excimer laser induced bubble: Dimensions, theory and implications for laser angioplasty. *Lasers Surg. Med.*, 18:381–390 (1996).
49. Jansen ED, van Leeuwen TG, Motamedi M, Borst C, and Welch AJ. Partial vaporization model for pulsed mid-infrared laser ablation of water. *J. Appl. Phys.*, 78:564–571 (1995)
50. Cummings JP and Walsh JT. Tissue tearing caused by pulsed laser induced ablation pressure. *Appl. Opt.*, 32:493–503 (1993).
51. Zweig AD, Meierhofer B, Müller OM, Mischler C. Romano V, Frenz M, and Weber HP. Lateral thermal damage along laser incisions. *Lasers Surg. Med.*, 10: 262–274 (1990).
52. Brugmans MJP, Kemper J, Gijbers GHM, van der Meulen FW, and van Gemert MJC. Temperature response of biological materials to pulsed non-ablative CO_2 laser irradiation. *Lasers Surg. Med.*, 11:587–594 (1991).
53. Carslaw HS and Jaeger JC. *Conduction of Heat in Solids.* Oxford University, Oxford, 2nd edition, pp. 58–62 (1959).
54. Walsh JT and Cummings JP. Tissue tearing caused by pulsed laser induced pressure. In Jacques SL (ed) *Laser-Tissue Interactions I*, Vol. 1202. SPIE, Bellingham, pp. 12–21 (1990).
55. Long FH and Deutsch T. Pulsed photothermal radiometry of human artery. *IEEE J. Quant. Electron.*, QE-23:1821–1826 (1987).
56. Meijering LJT, van Gemert MJC, Gijsbers GHM, and Welch AJ. Limits of radial time constants to approximate thermal response of tissue. *Lasers Surg. Med.*, 13:685–687 (1993).

57. van Gemert MJC, Lucassen GW, and Welch AJ. Time constants in thermal laser medicine II: Distributions of time constants and thermal relaxation of tissue. *Phys. Med. Biol.*, 41:1381–1399 (1996).
58. Zweig AD, Frenz M, Romano V, and Weber HP. A comparative study of laser tissue interaction at 2.94 μm and 10.6 μm. *Appl. Phys. B*, 47:254–265 (1988).
59. Walsh JT and Cummings JP. The effect of dynamic changes in the water absorption coefficient on mid-infrared laser ablation. *Lasers Surg. Med.*, 15:295–305 (1994).
60. Walsh JT, Flotte TJ, and Deutsch TF. Er:YAG laser ablation of tissue: Effect of pulse duration and tissue type on thermal damage. *Lasers Surg. Med.*, 9:314–326 (1989).
61. Frenz M, Pratisto H, Könz F, Jansen ED, Welch AJ, and Weber HP. Comparison of the effects of absorption coefficient and pulse duration of 2.12 μm and 2.79 μm radiation of laser ablation of tissue. *IEEE J. Quant. Electron.*, 32(12):2025–2036 (1996).
62. Mackanos MA, Jansen ED, Shaw BL, Sanghere JS, Aggarwal I, and Katzir A. Delivery of midinfrared (6 to 7-micron) laser radiation in a liquid environment using infrared-transmitting optical fibers. *J. Bioimed. Opt.*, 4:583–593 (2003).
63. Kang HW and Welch AJ. Effect of liquid thickness on laser ablation efficiency. *J. Appl. Phys.*, 101:083101 (2007).
64. Kang HW, Lee H, and Welch AJ. Laser ablation in a liquid-confined environment using a nanosecond laser pulse. *J. Appl. Phys.*, 103:083101 (2008).
65. Dabby FW and Paek UC. High-intensity laser-induced vaporization and explosion of solid material. *IEEE J. Quant. Electron.*, QE-8:106–111 (1972).
66. Cummings JP and Walsh JT. Erbium laser ablation: The effect of dynamic optical properties. *Appl. Phys. Lett.*, 62:1988–1990 (1993).
67. Cummings JP and Walsh JT. Thermal changes in the absorption spectrum of water near 6.1 μm. *Lasers Surg. Med.*, Suppl. 5:2–3 (1993).
68. Visuri SR, Cummings, JP, and Walsh JT. Dynamics of the absorption coefficient of water using 2.1-μm laser radiation. *Lasers Surg. Med.*, Suppl. 5:3 (1993).
69. Jansen ED, Leeuwen TG van, Motamedi M, Borst C, and Welch AJ. Temperature dependence of the absorption coefficient of water for mid infrared laser radiation. *Lasers Surg. Med.*, 14:258–268 (1994).
70. Graener H, Seifert G, and Laubereau A. New spectroscopy of water using tunable picosecond pulses in the infrared. *Phys. Rev. Lett.*, 66:2092–2095 (1991).
71. Vodopyanov KL. Bleaching of water by intense light at the maximum of the $\lambda = 3$ μm absorption band. *Sov. Phys. JETP*, 70:114–121 (1990).
72. Vodopyanov KL. Saturation studies of H_2O and HDO near 3400 cm^{-1} using intense picosecond laser pulses. *J. Chem. Phys.*, 94:5389–5393 (1991).
73. Ediger MN, Pettit GH, Weiblinger RP, and Chen Ch. Transmission of corneal collagen during ArF excimer laser ablation. *Lasers Surg. Med.*, 13:204–2190 (1993).
74. Pettit GH and Ediger MN. Pump/probe transmission measurements of corneal tissue during excimer laser ablation. *Lasers Surg. Med.*, 13:363–367 (1993).
75. Bode G. Das Verhalten des Eises im ultraroten Spectrum. *Ann. Phys.* (Leipzig), 30:326–336 (1909).
76. Hibst R and Keller U. Experimental studies of the application of the Er:YAG laser on dental hard substances: I. Measurement of the ablation rate. *Lasers Surg. Med.*, 9:338–344 (1989).
77. Wolbarsht ML. Laser surgery: CO_2 or HF. *IEEE J. Quant. Electron.*, 20:1427–1432 (1984).
78. Isner JM, Clark RH, Katzir A, Gold D, DeJesus ST, and Lalaburka KR. Transmission characteristics of individual wavelengths in blood do not predict ability to accomplish laser ablation in a blood field: Inferential evidence of the 'Moses Effect'. *Circulation*, 74:II–361 (1986).
79. Luck WAP. Zur Assoziation des Wassers III. Die Temperaturabhängigkeit der Wasserbanden bis zum kritischen Punkt. *Ber. Bunsenges. Phys. Chem.*, 69:626–637 (1965).
80. Falk M and Ford TA. Infrared spectrum and structure of liquid water. *Can. J. Chem.*, 44:1699–17071 (1966).
81. Walsh JT Jr. Pulsed laser angioplasty. In: Welch AJ and van Gemert MJC (eds), *Optical-Thermal Response of Laser-Irradiated Tissue*. Plenum, New York , 1st edition (1995).

82. van Erven L, van Leeuwen TG, Post MJ, vander Veen MJ, Velema E, and Borst C. Mid-infrared pulsed laser ablation of the arterial wall. Mechanical origin of 'acoustic' wall damage and its effect on wall healing. *J. Thoracic Cardiovasc. Surg.*, 104:1053–1059 (1992).
83. van Leeuwen TG, Jansen ED, Motamedi M, Welch AJ, and Borst C. Excimer last ablation of soft tissue: A study of the content of fast expanding and collapsing bubbles. *IEEE J. Quant. Electton.*, 30(5):1339–1345 (1994).
84. Pratisto H, Frenz M, Koenz F, Altermatt HJ, and Weber HP. Combination of erbium and holmium laser radiation for tissue ablation. *Proc. SPIE*, 2681:201–206 (1996).

Part III
Medical Applications

Chapter 16
Introduction to Medical Applications

Ashley J. Welch and Martin J.C. van Gemert

16.1 Introduction

The first two parts of this book describe various theories associated with light propagation in tissue and the resulting response. If coherence or polarization information is not needed, we assume that the transport equation governs the optical interaction of light with tissue and the heat conduction equation provides the basis for estimating the thermal response of tissue to laser radiation. In part III of this book, the theory for optical and thermal interactions of laser light with tissue are used to analyze medical applications. In particular, the concepts of parts I and II

a. can describe the optical and thermal interaction of therapeutic applications of lasers;
b. can estimate dosimetry requirements for photochemical and thermal clinical treatments;
c. can provide predictive information for design of medical systems; and
d. can provide a framework for interpreting optical and thermal measurements, especially for diagnostic applications.

In the following chapters a variety of examples are presented that either directly use the theory of the first two parts, describe nonlinear effects of lasers, or discuss the delivery of laser light. In this chapter we summarize some of the optical and thermal interactions that are of relevance to these applications.

A.J. Welch (✉)
Department of Biomedical Engineering, The University of Texas, Austin, TX, USA
e-mail: welch@mail.utexas.edu

A.J. Welch, M.J.C. van Gemert (eds.), *Optical-Thermal Response of Laser-Irradiated Tissue*, 2nd ed., DOI 10.1007/978-90-481-8831-4_16,

16.2 Light Propagation in Tissue

16.2.1 Treatment

Modeling light scattering is essential for establishing dosimetry for photochemical reactions in tissue such as in photodynamic theory (PDT). Typically, excitation wavelengths for PDT are between 600 and 800 mm to maximize penetration of laser light into the tissue. In this limited spectrum, absorption is typically much less than scattering. The key dosimetry parameter for PDT is the fluence rate $\phi(r)$. Typically, the laser spot size for PDT of skin is large, to cover as much of the lesions as possible and reduce the number of individual spots that must be treated. In these cases the spot size is much larger than the effective penetration depth, δ_{eff}, The large area of irradiation maximizes the penetration depth of the light per unit of irradiance since most of the scattered light remains within the cylinder of collimated light in tissue (see Chapter 3). Only at the edges is light lost in a radial direction. Thus the fluence rate along the center of the beam can be estimated with a radial independent solution of the transport equation. Often the statement is made that for highly scattering tissue the diffuse light in tissue decreases exponentially with depth, according to

$$\phi(z) = Ae^{-\mu_{eff} z} \tag{16.1}$$

where A is a constant and μ_{eff} is the effective attenuation coefficient, $\sqrt{3\mu_a\left[\mu_a + \mu_s\,(1-g)\right]}$. It is tempting to assume that the constant A represents the magnitude of irradiance, E, at the surface. Then at a depth of $\delta_{\text{eff}} = 1/\mu_{\text{eff}}$ the fluence rate is $0.37E$ [W/m^2] as shown in Fig. 16.1. However, from the material in Chapters 3, 5 and 6 it is apparent that A may be several times larger than E. For example suppose A=3E. Then the depth at which $\phi(z)$ is approximately equal to $0.37E$ is equal to 2.1 δ_{eff} as illustrated in Fig. 16.1. Thus, the effective treatment depth is actually twice the depth predicted by an improper interpretation of the fluence rate of light in a scattering medium.

The selection of dosimetry parameters is not as obvious as it seems. In a closed cavity such as the bladder, the fluence rate is further increased by an "integrating sphere" effect; van Gemert et al. have suggested that the maximum fluence rate

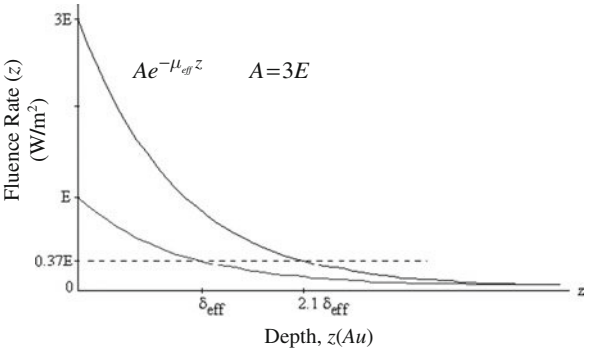


Fig. 16.1 Beer's Laws attenuation of diffuse light for an irradiance, E, and attenuation according to Eq. (16.1) when $\mu_a \ll \mu_s$

in the bladder wall at 633 nm may be 7–9 times the primary irradiance [1]. This effect has been measured in a bladder phantom and in human bladders by Star and colleagues [2]. Fluence rates higher than expected produce a greater than expected photochemical reaction rate or heat production. Thus, in addition to photochemical reactions, the dosimetry for thermal applications are affected by light scattering.

Do values of fluence rate larger than irradiance suggest that we get something for nothing? Not at all. The laws for conservation of energy are not violated. Keep in mind this discussion of fluence rate is a steady state analysis and at any instant scattered photons have been in the tissue for different periods of time. Nevertheless, the steady state energy balance can be formulated. Assume a uniform irradiance of a semi-infinite medium with rate of energy delivery E. For matched boundary conditions, the delivered energy must be balanced by light energy absorbed and light remitted from the medium (see Fig. 16.2). Obviously the remitted light must be less than E. The remaining light is absorbed and converted to heat.

The total rate of heat generation per unit area [W/m^2] is computed by integrating the volumetric heat source term $S(z) = \mu_a \phi(z)$ over z. Using Eq. (16.1) as a one-dimensional estimate for $\phi(z)$ for constant irradiance E yields

$$\int_0^\infty S(z)dz = \int_0^\infty \mu_a A e^{-\mu_{eff} z} dz = \frac{\mu_a A}{\mu_{eff}} \tag{16.2}$$

which should be equal to E–RE. However, Eq. (16.1) was only an approximation in our example, especially near the surface.

Conservation of energy, here in the case of matched boundaries, now implies that the sum of total rate of heat generation per unit area and the total back scattered photon power per unit area equals the irradiance

$$\int_0^\infty \mu_a \phi(z)dz + ER = E \tag{16.3}$$

It is impossible to work this out for the example of Fig. 16.2 because we do not know the expression for R, although it has to be $(1 - 3\mu_a/\mu_{\text{eff}})$, based on Eqs. (16.2) and (16.3). A simple example where Eq. (16.3) can be tested is with strictly 1-D transport theory, described in the first volume of the book but not in

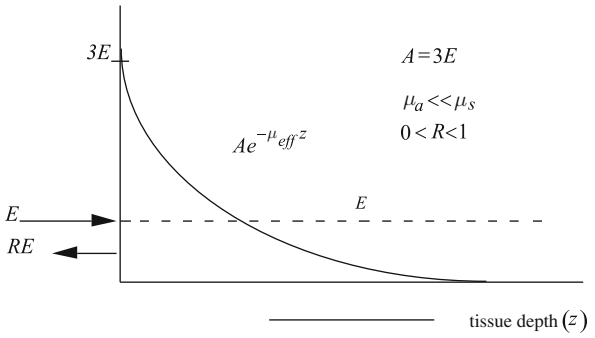


Fig. 16.2 Geometry for performing a volumetric energy balance, Eq. (16.2)

this volume. Then, for an infinite tissue and matched boundaries, $A = E(1 + R)$ and $R = \left[(\mu_a/\sigma) + 1 - \sqrt{(\mu_a/\sigma + 1)^2 - 1}\right]$, whereas $\mu_{\text{eff}} = \sqrt{\mu_a(\mu_a + 2\sigma)}$, where σ denotes here the back-scattering coefficient. It is straightforward to show Eq. (16.3) to be true.

16.2.2 Diagnosis

An important use of lasers is their application to diagnostic problems. This class of problems includes the use of light to either identify the status of tissue or to classify tissue. For example, can the extent of thermal damage during photocoagulation be monitored using either reflection, OCT or fluorescence measurements? For tissue that whitens during thermal coagulation, reflection is an excellent indicator of the progress of the response. Examples are the enhanced scattering of clear albumin or the retina during photocoagulation. In these cases, a white lesion depicts the region of thermal damage. Thus, the diameter of a thermal lesion is clearly delineated by the white coagulation boundary and can easily be monitored with a ccd array. The depth of coagulation is more difficult to determine. As albumin or the retina coagulates there is a dramatic increase in the scattering coefficient as evidenced by the whitening of the tissue. As the thickness of coagulated tissue (thickness of highly scattering material) increases, remittance increases as described in Chapters 3, 5 and 6. Thus, the thickness of coagulation is related to the magnitude of the reflectance image [3].

There are a number of examples of the use of tissue optics for analyzing fluorescence for the purpose of classifying tissue. As stated in Chapters 3 and 20 the measured fluorescence signal is dependent upon

a. the fluence rate distribution of the excitation light;
b. the quantum yield of the fluorochromophore; and
c. the attenuation of the fluorescence light by the tissue.

Because of the wavelength dependence of the optical properties of the tissue, the measured fluorescence line shape is not equal to the intrinsic line shape of the fluorochromophore. By using Monte Carlo models, it is possible to simulate the effects of tissue attenuation, tissue sample geometry and detector field of view and alignment. The optical modeling results are obtained by combining two Monte Carlo simulations: a single photon, variable step size model for the excitation ray and a weighted photon, variable step size model for the fluorescence. A simple algorithm for this simulation is depicted in Fig. 16.3.

Although this model may require a longer simulation time than other models, parameter values, geometry, and output format are easily modified. As shown in Fig. 16.3, one fluorescent photon is generated for each absorbed excitation photon. This scaling has many advantages during the development of the computer program and when analyzing results. At any time the results can be multiplied by the (real) quantum yield.

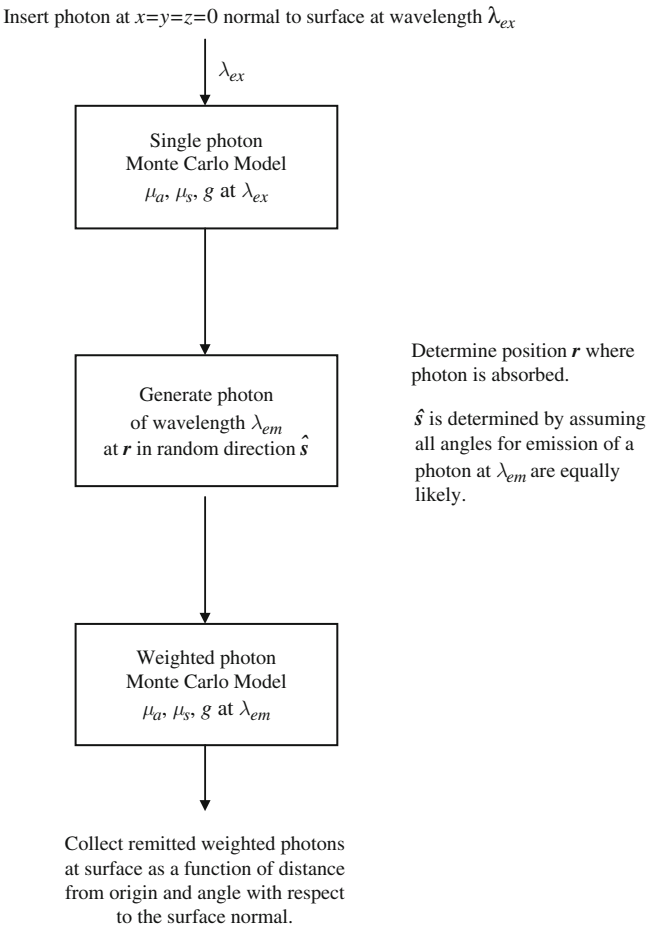


Fig. 16.3 Block diagram for simulation of fluorescence

16.3 Optical-Thermal Response

The optical-thermal analysis of laser irradiated tissue has three distinct parts:

a. The diffusion approximation or Monte Carlo simulation is used to calculate the fluence rate and rate of heat generation in the tissue during laser radiation.
b. The heat conduction equation is employed to calculate the transient temperature field.
c. The rate process damage model is used to estimate the extent of damage.

If the physical properties of the tissue are known (optical, thermal, and rate process), the model indicates the general effect of irradiation wavelength, power and exposure duration in terms of temperature rise and extent of damage upon the tissue.

An example of optical-thermal analysis is the treatment of Port Wine Stains (PWS) by laser (Chapter 23). In this case, it is generally assumed that the ectatic dermal blood vessels are to be selectively injured through a two-step process: first laser heating of the red blood cells in the vessel, followed by injury of the vessel wall through conduction of heat.

Selective injury of the vessel wall, by conduction of heat from the heated red blood cells, without (too much) injury of the dermal tissue around the vessel, determines in principle the optimal laser pulse duration. In Chapter 23 it is argued that such optimal pulse duration is most likely around a few ms; a more precise value cannot be given as neither the temperatures of red blood cells in the lumen of the vessel is known in response to laser irradiation, nor the "critical temperature" of irreversible injury of the ectatic wall of a vessel. In any case, the pulse duration should be longer than 0.1 ms as to prevent tearing of the vessel followed by complete healing rather than by thrombosing.

Through the years, the choice of "best" laser wavelength has always been, and in fact still is, subject for controversy among the workers in the field. At least it has been established that the CO_2 laser is unsuitable for PWS treatment: the radiation is absorbed in the epidermis; coagulation of a vessel in the dermis produces a full thickness skin injury to the level of the target vessel [6].

The "best" wavelength for treatment of PWS is obtained by examining the rate of change of temperature (dT/dt) in the red blood cells of target vessels relative to the epidermis and bloodless dermis. Wavelengths that maximize this difference in dT/dt between blood and the other skin constituents are considered optimal as a function of the depth of a target vessel. For short irradiation times (but long enough for conducting heat through the vessel wall), this is equivalent to saying that the volumetric production of heat in the red blood cells is as large as possible, or

$$\frac{dT}{dt} \cong \frac{1}{\rho c}\,\mu_{a,b}\,(\lambda)\,\phi\,(\lambda,\, z_v) \tag{16.4}$$

is as large as possible; $\mu_{a,b}(\lambda)$ is the absorption coefficient of blood at wavelength λ and $\phi\,(\lambda,\, z_v)$ is the fluence rate at the top surface of a target vessel at a depth z_v.

Previously, it was assumed that a PWS could be modeled as one simple ectatic blood vessel at a dermal depth z_v in a bloodless dermis. Maximizing dT/dt at z_v, required maximizing $\mu_{a,b}(\lambda)$ as a function of wavelength since fluence rate $\phi\,(\lambda,\, z_v)$ is not dependent on wavelength from 575 to 600 nm. Hence, $\lambda = 577$ μm, at the maximum of oxyhemoglobin absorption, was chosen as the best wavelength [4, 5]. However, it was overlooked that a PWS is not one ectatic vessel in a bloodless dermis but, instead, a conglomerate of many ectatic vessels which, when the skin was irradiated by a 3–10 mm diameter laser beam at 577 nm, absorbed quite a number of scattered laser photons before they reached and injured a deep lying target vessel (Fig. 16.4). In other words, the $\phi\,(\lambda,\, z_v)$ was severely reduced when 577 nm was used with maximal absorption for red blood cells. An increase in dT/dt, or in $\mu_{a,b}(\lambda)\ \phi\,(\lambda,\, z_v)$, can then be achieved by *reducing* $\mu_{a,b}(\lambda)$, thereby strongly *increasing* $\phi\,(\lambda,\, z_v)$. It is shown in Chapter 23 that, depending upon the dermal blood concentration, if the wavelength is increased to a few nm longer that 577 nm,

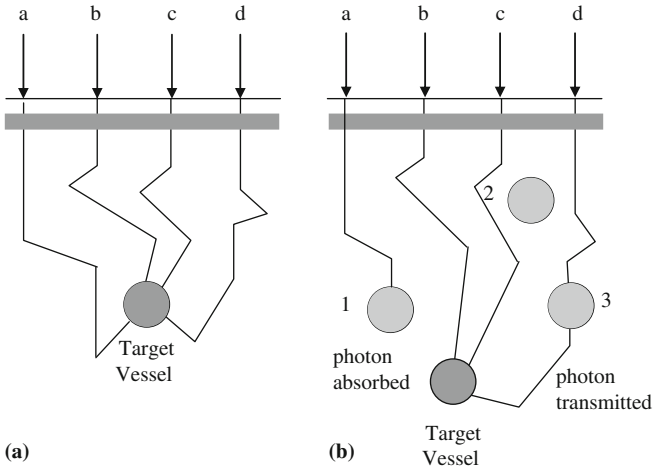


Fig. 16.4 (**a**) Depiction of photons propagating to target vessel that is in the dermis of the skin. (**b**) Alteration of photons' migration by vessels anterior to the target vessel

the rate of heat generation in blood at depth z_v is increased. Longer wavelengths than 577 nm are primarily considered because melanin absorption and epidermal and dermal scattering increase as λ is decreased. Increased absorption reduces the penetration depth of the laser light.

Clinically, 585 nm has been proposed as the "best" wavelength. However, this analysis predicts that the "best" wavelength depends on the PWS anatomy itself! An unknown factor in the analysis is whether indeed 577 nm is the best wavelength for selective photothermolysis of *one single* vessel, irrespective of its size, as is tacitly assumed so far. The consequence is that, according to this assumption, the "best" wavelength depends also on the laser beam diameter: for very small diameters that irradiate only one or two vessels ($\sim$ 0.1–0.2 mm), 577 nm is going to be the "best" wavelength again. This prediction holds for a telangiectatic lesion because it has only one irradiated vessel.

References

1. van Gemert MJC, Cheong W-F, Welch AJ, and Star WM. Light delivery for whole-bladder photodynamic therapy. *Lasers Med. Sci.*, 2:273–283 (1987).
2. Marijnissen JPA, Star WM, In't Zandt HJA, D'Hallewin MA, and Baert L. In situ dosimetry during whole bladder wall photodynamic therapy: Clinical results and experimental verification. *Phys. Med. Biol.*, 37:1–16 (1993).
3. Jerath MR, Kaisig D, Rylander HG III, and Welch AJ. Calibrated real time control of lesion size based on reflectance images. *Appl. Opt.*, 32(7):1200–1209 (1993).
4. Anderson RR and Parrish JA. Microvasculature can be selectively damaged using dye lasers: A basic theory and experimental evidence in human skin. *Lasers Surg. Med.*, 1:263–270 (1981).
5. van Gemert MJC and Hulsbergen Henning JP. A model approach to laser coagulation of dermal vascular lesions. *Arch. Dermatol. Res.*, 270:429–439 (1981).
6. van Gemert MJC, Welch AJ, Tan OT, and Parrish JA. Limitations of CO_2 lasers for treatment of portwine stains. *Arch. Dermatol.*, 123:71–73 (1987).

Chapter 17
Optical Fiber Sensors for Biomedical Applications

Lee C.L. Chin, William M. Whelan, and I. Alex Vitkin

17.1 Introduction

Optical fiber technology offers a convenient, affordable, safe and effective approach for the delivery and collection of light to and from the tissue region of interest, and has been employed clinically since the 1960s [1]. This chapter discusses and reviews the recent developments in optical fiber sensor technology in the field of biomedicine.

Before proceeding, we distinguish between the previous edition of this chapter (Optics of Fiber and Fiber Probes, 1995 edition, Chapter 19), which primarily examined the role of optical fibers for *tissue modification* (e.g., photo-therapeutics such as laser heating), and the current treatise. Here, we focus on the clinical application of optical fiber technology for *tissue assessment*, in the contexts of *diagnosis* and *therapeutic monitoring*. While the term *probe* can describe either therapeutic or diagnostic intent, we classify probes for therapy as *applicators* and probes for diagnosis as *sensors*, and limit our discussion to the latter.

The use of optical fiber technology offers numerous advantages that are well-suited for clinical use. The sensors can be made biologically compatible (non-toxic and bio-chemically inert) and are immune from electromagnetic interference. They can be placed *non-invasively* in contact with external organs such as the skin or surgically exposed surfaces. In addition, due to their flexibility and thin outer diameter, they can also be placed into bodily cavities (endoscopic approach), inserted interstitially via minimally invasive trocars, (e.g., hollow bore needles), or positioned intravascularly. As such, measurements can be performed in difficult-to-access parts of the human body with greater "local" sensitivity. Finally, it is technologically possible to bundle multiple sensors with different measurement capabilities into a single probe as a packaged instrument, thus potentially increasing useful information content.

L.C.L. Chin (✉)
Medical Physics Department, Odette Cancer Centre, Toronto, ON, Canada
e-mail: Lee.Chin@sunnybrook.ca

A.J. Welch, M.J.C. van Gemert (eds.), *Optical-Thermal Response of Laser-Irradiated Tissue*, 2nd ed., DOI 10.1007/978-90-481-8831-4_17,

The choice and design of an optical fiber sensor is generally dictated by technological and clinical issues such as access to the target site, spatial resolution, spatial localization, desired sampling volume, (non)invasiveness, accuracy and overall clinical intent. The combination of these factors forms a unique biomedical engineering problem that dictates the various system parameters such as the choice of light source and detector characteristics, source-detector arrangement, and fiber tip modifications, among others. In addition, appropriate signal analysis must be chosen for specific applications (e.g., background fluorescence subtraction in fiber-based Raman spectroscopy, model fits in fiber-based spatially dependent diffuse reflectance, corrections of fluorescence spectra for tissue attenuation distortions, etc.).

In this chapter, the fundamentals of optical fiber sensors will be reviewed, along with basic definitions, classifications and applications. Selected examples are chosen to illustrate the various design concepts. While this chapter by no means constitutes an exhaustive review, we hope to provide the reader with a sampling of the type of approaches and innovations that can be used for solving clinical problems.

17.2 Optical Fiber Sensors

17.2.1 Basic System Design

A typical optical fiber sensing system, shown in Fig. 17.1, is comprised of four major components: a light source, optical fiber(s), optical coupling/filters and a detector.

1. The light source is typically a laser (pulsed, modulated, or steady-state) or white light (xenon or mercury lamp, with or without filters) used to probe or excite

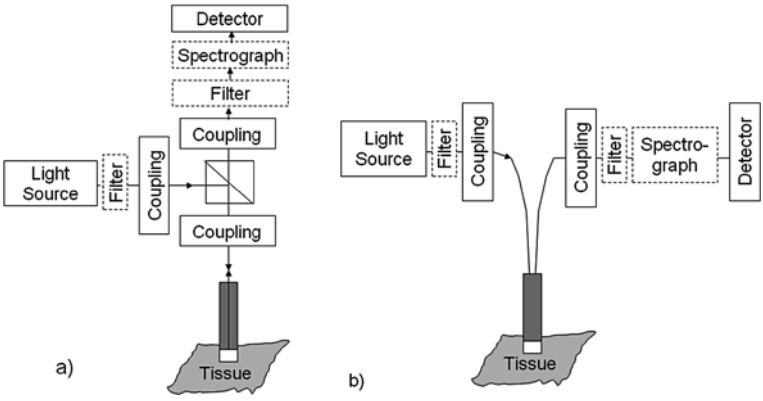


Fig. 17.1 Typical optical fiber sensors employing (**a**) a single fiber for both delivery and detection or (**b**) separate source-detector fibers (after Utzinger and Richards-Kortum [2]). *Solid boxes* represent essential components always present in fiber-based sensor systems; *dotted boxes* are optical components that may be optimized for a specific application

the tissue. The light source characteristics provide a wide range of properties for optical interrogation of tissues including intensity, phase, polarization state, and wavelength content.

2. Optical fibers deliver light from a light source to a sensing mechanism (see Section 17.2.2 below) located at the fiber end, which is coupled to the interrogated medium (e.g. tissue). Here, the light experiences a *modulation* by the interrogated tissue, and is returned via the same (Fig. 17.1a) or different fiber(s) (Fig. 17.1b) to a light-measurement device where it can be detected and analyzed. Often, a portion of the interrogating light can be used as a reference (the ratio method), to correct for optical fluctuation and to provide improved noise rejection.

   Fibers can be bundled to increase the interrogation volume and improve signal-to-noise by capturing a larger signal. This typically requires stripping the outer buffer (jacket) off the fiber to reduce overall diameter. A hexagonal packing configuration provides the optimal arrangement for minimizing the dead space (inactive area between fibers) due to the cladding. The total number of packed fibers, T, in a circular cross section for such an arrangement is

$$T = 1 + \sum_{i=0}^{k} 6i \tag{17.1}$$

   where k is the total number of rings. With the jacket removed, a fiber bundle can achieve a total active area of almost 65%.
3. Optical components such as connectors, lenses, mirrors, fiber couplers, circulators, polarizers, phase modulators, and beam-splitters can be configured to optimize the signal in the detector fiber. Single source-detector geometry results in a small sensor diameter with a small interrogation volume and high light collection efficiency. However, compared to multi-distance sensors and/or fiber bundle geometries, single-fiber sensors suffer from high unwanted background signals generated within the fiber. These can stem from fiber autofluorescence generated by the excitation light or from backscattered illumination light. As such, filters (bandpass, monochromatic) can be used to remove background light or reduce source intensities to appropriate safety levels for both the tissue and the light detection system. Filters can also be applied to the source beam to improve spectral purity and/or to enable spectroscopy.
4. For single wavelength detection and/or spectrally-unresolved total intensity, a photomultiplier tube or photodiode is used, whereas for spectral detection, a charged coupled device (CCD) or spectrograph/monochrometers is used. Conversion of detected light to electrical signal generally requires standard analog-to-digital circuitry.

17.2.2 Classification of Fiber Optic Sensors for Biomedicine

Optical fiber sensors can be divided into two main categories based on the sensing mechanism: *direct* and *indirect*. Direct sensors (e.g., photometric sensors)

utilize the tissue itself to modulate the illuminating light, whereby the collected light is the result of backscattering directly from the interrogated tissue or tissue fluorescence/Raman induced by an optical source. Indirect sensors employ an intermediary in response to the tissue property of interest (e.g., temperature, enzyme presence). Indirect sensors can be subdivided into *intrinsic*, which utilize the fiber itself (core and/or cladding) as the sensing element, and *extrinsic*, which incorporate an additional sensing element at the fiber end (e.g., transducer or substrate). Examples include physical sensors that employ miniaturized transducers that modulate the light in response to such physical parameters as temperature, pressure and radiation dose. Chemical or biosensors evaluate the change in a molecular reagent attached to the end of the fiber via spectroscopic or fluorometric measurement. Such sensors have been utilized for the measurement of pH, glucose, and other intrinsic metabolites. Figure 17.2 presents a diagrammatic overview of the classification schemes for optical fiber sensors in biomedicine.

The primary discussion in this chapter focuses on photonic sensors, since this is the most relevant sensor type in the biophotonics community. Depending on the clinical purpose (diagnostic or dosimetric), *photonic* sensors offer a range of potential modalities that provide unique physical measurements including white light spectroscopy, fluorescence, Raman scattering, optical coherence tomography (OCT), and polarimetry. Of particular importance for sensor design is the choice of the appropriate modality for the given application. Table 17.1 lists the various modalities discussed in this chapter, including an overview of the relevant physical measurements, their advantages and disadvantages, and their potential applications.

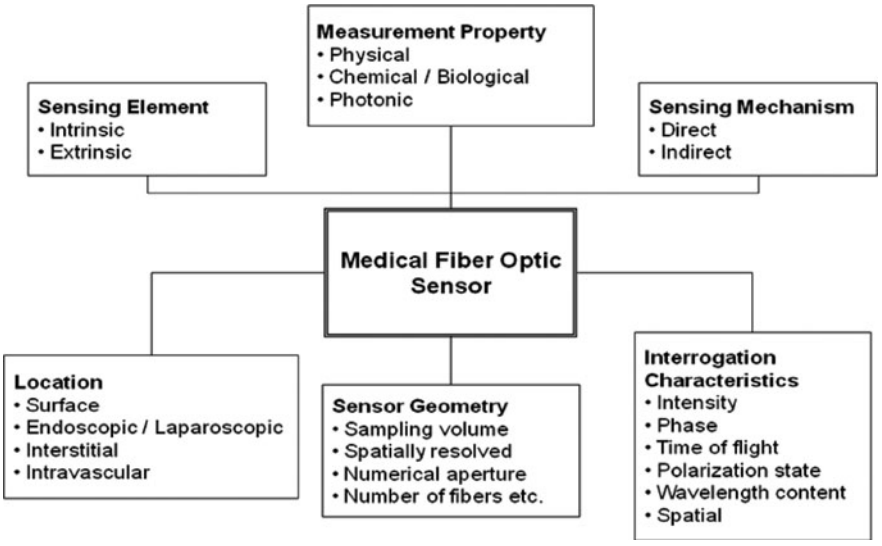


Fig. 17.2 Classification schemes for optical fiber sensors in biomedicine

Table 17.1 Optical fiber sensor modalities to be discussed in this chapter, listing the physical measurements, applications, advantages and disadvantages

Modality	Detection/ determination	Applications	Advantages	Disadvantages
White light spectroscopy	• absolute light intensity • optical properties • chromophore concentration • scatterer spacing and density • oxygen saturation	• diagnosis • oximetry • light intensity dosimetry • chromophore/ drug concentration	• range of sampling volumes • relatively inexpensive • potentially simple data processing	• non-imaging • many contributions to the measured signal
Fluorescence	• fluorescence intensity optical properties fluorophore concentration	• diagnosis • treatment evaluation • fluorophore/ drug concentration	• direct measurement • wide array of possible contrast agents	• uncertain sampling volume • auto fluorescence may interfere
Raman	• metabolite concentration • optical properties • tissue compositional information • oxygen saturation	• diagnosis • treatment monitoring/ evaluation	• high specificity	• weak signals • long(er) integration times • need careful analysis to remove confounding signals
OCT	• tissue microstructure • blood flow/ Doppler • birefringence	• diagnosis • treatment monitoring/ evaluation • tissue functional status (Doppler blood flow maps)	• 3-D imaging	• limited penetration depth

The choice of modality, in turn, guides a specific probe design that must consider the accessibility of the target site. Superficial (exposed surface) tissue can typically be examined using surface reflectance or endoscopic sensors, while internal structures may require sensors suitable for interstitial insertion. Furthermore, in the case of superficial targets, layered structures may require additional design considerations for spatial-depth discrimination. Such factors can be adjusted for using appropriate source-detector geometry and fiber-tip design.

While the specifics of actual sensors will be reviewed, we also examine the physics behind sensor design parameters and their relation to light collection efficiency and sampling volume in turbid media. By focusing on the physics of sensor designs, it is hoped that the reader can extend the information of this chapter to a specific biomedical application.

17.3 Fiber Fundamentals

17.3.1 Light Transmission

A typical optical fiber consists of (1) a central core of refractive index n_1; (2) a cladding that encases the core with index of refraction n_2 (with $n_2 < n_1$); and (3) a buffer coating – also known as a "jacket." The role of the buffer is to improve the mechanical robustness of the fiber and minimize structural compromise from bending. A step-index fiber is shown in Fig. 17.3 which is characterized by a discrete step in the index of refraction at the boundary of the core (n_1) and cladding. Another common design, the graded-index fiber, where the index of refraction decreases gradually from the center of the core to the cladding, and will be discussed later.

When a light ray is directed at the entrance of a fiber with core index of refraction n_1, a portion of the incoming beam is reflected back into the surrounding medium (index of refraction n_0), while a portion of the beam is transmitted into the fiber according to Snell's law:

$$\frac{\sin(\alpha_1)}{\sin(\alpha_0)} = \frac{n_0}{n_1} \tag{17.2}$$

where α_0 is the incidence angle and α_1 is the transmission angle into the fiber. Light transport inside optical fibers is based on the concept of total internal reflection, which requires that $n_2 < n_1$ and that the incidence angle at the core/cladding boundary is greater than a critical value, θ_c according to:

$$\sin(\theta_c) = \frac{n_2}{n_1} \tag{17.3}$$

Light rays incident at angles less than θ_c such as at point B in Fig. 17.3 (i.e., more "normal" to the interface) are refracted from the core into the cladding and, hence, do not propagate along the fiber. Now, it is clear from Fig. 17.3 that $\alpha_c = 90° - \theta_c$. Hence, in order to propagate light along an optical fiber, it is necessary to direct light

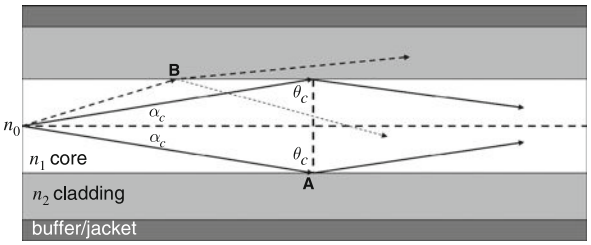


Fig. 17.3 Optical fiber design. n_0, n_1 and n_2 are the refractive indices of the surrounding medium, fiber core and fiber cladding, respectively. Light rays incident at point A with angles greater than θ_c (Eq.(17.3)) undergo total internal reflection and constitute the propagating light modes; those light rays incident at angles smaller than θ_c (point B) are refracted from the core into the cladding and do not propagate down the fiber. A step-index fiber is shown for illustration (sharp n_1/n_2 boundary)

into the fiber such that it refracts at or below the critical propagation angle, α_c. The maximum angle of incidence that a light ray can enter a fiber and be transported along the fiber (propagate at or below α_c) is known as the acceptance angle, θ_a. For fiber characterization, the acceptance angle is typically reported in terms of a numerical aperture, *NA*:

$$NA = n_0 \sin(\theta_a) = \sqrt{n_1^2 - n_2^2} \tag{17.4}$$

Hence, for a given acceptance angle, light enters (or leaves) the fiber within an acceptance cone of $2\theta_a$.

From Eq. (17.4), it follows that the NA can be expressed in terms of the refractive indices of the fiber material (core n_1, and cladding n_2), but is also dependent on the refractive index of the surrounding environment n_0. As such, compared to an air environment (n_0=1), the acceptance angle will be correspondingly smaller in water (n_0=1.33). The *NA* of an optical fiber characterizes not only its ability to collect light from a source, but also preserve the light inside the fiber.

The choice of wavelength range plays an important role in the selection of appropriate core and cladding materials that allow for suitable transmission of light. Conventional glass and plastic core fibers, highly transparent to wavelengths in the visible spectrum, have absorption losses on the order of 0.1% per meter or less. However, optical transmission in the ultraviolet (UV) or infrared (IR) wavelength ranges requires the use of high-grade fused silica cores. In these wavelength ranges, the presence of hydroxyl (OH) radicals in the silica core strongly affects the absorption and transmission characteristics of the fiber. For wavelengths in the UV region (200 nm < λ < 360 nm), high-OH fibers are preferred, whereas in the near infrared (NIR) and up to ~2400 nm, low-OH concentration fibers provide the most favorable transmission. In the infrared region beyond ~2400 nm, absorption by silica begins to dominate. In this regime, sapphire fibers provide superior transmission but can suffer additional signal decreases due to high reflection losses at the fiber interface (because of sapphire's high refractive index).

17.3.2 Fiber Classifications

Fibers fall into two broad categories, single mode or multi-mode. A mode is defined as a set of similar paths that light rays travel along the fiber. The term single mode is used to describe a fiber that supports one transmission mode, whereas multi-mode describes a fiber that can support more than one transmission mode. Multi-mode fibers are commonly employed in biomedical applications and come in two basic configurations: step-index and gradient index.

The step-index multi-mode fiber is the simplest design, and operates in the fundamental manner described in Section 17.3.1. Its name comes from the sharp change in refractive index in the fiber material at the core-cladding boundary. In contrast, with gradient-index fibers, the refractive index decreases gradually from the center of the core to the cladding.

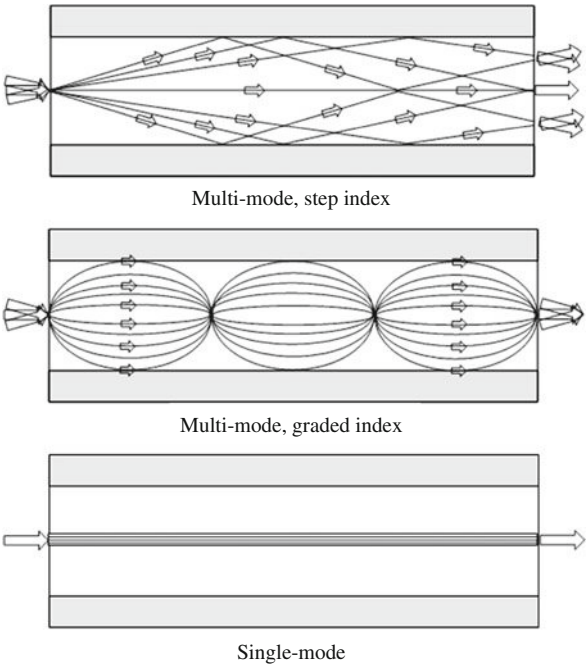


Fig. 17.4 Multi- mode (high *NA*, large core diameter) and single-mode (lower *NA*, smaller core diameter) fibers

The concept of modes is illustrated in Fig. 17.4. Fibers with a high *NA* allow light at a large range of incident angles to enter the fiber. The result is that incoming rays with smaller angles of incidence will travel shorter path-lengths closer to the central (optic) axis of the fiber (low order modes), while rays entering at larger incident angles will travel longer path-lengths (higher order modes). Further, impurities in the fiber material of both multi-mode and single-mode fibers can cause leaking at the core/cladding interface. However, due to the longer distance traveled by higher order modes, the effect is enhanced with multi-mode fibers leading to greater transmission losses compared to single-mode fibers.

The number of modes that are transmitted along a fiber is determined by the normalized frequency parameter, V

$$V = \frac{\pi d}{\lambda} NA \tag{17.5}$$

where d is the core diameter and λ is the input wavelength. For step-index and graded-index fibers the number of modes, N_m can be calculated via:

$$N_m = \frac{V^2}{2} \text{ (step-index)} \tag{17.6}$$

$$N_m = \frac{V^2}{4} \text{ (graded-index)} \tag{17.7}$$

Not surprisingly, single-mode fibers typically consist of a small core diameter (<10 μm) and small numerical aperture, while multi-mode fibers have bigger core diameters (>50 μm) and a corresponding larger numerical aperture (0.2–0.4 for silica-core fibers).

When a high coupling efficiency is desired, the optical fiber should have a large *NA* and the beam spot diameter, *d*, should be focused such that the focal length, *F*, of the coupling lens satisfies the condition:

$$F > \frac{d}{2NA} \tag{17.8}$$

However, a large *NA* results in a larger spectral dispersion of optical pathlengths (since *n* is typically a function of wavelength) that leads to a degradation of optical information at the distal end through spectral dispersion.

The propagating modes (and hence *NA*) of an optical fiber are important for both optical coupling efficiency and transmission characteristics. For example, following an initial short (picosecond range) injection of light into a step index multi-mode fiber, the differences in path-lengths among the various modes causes the light to spread out in time as it propagates along the fiber. This temporal broadening of the light pulse is known as modal dispersion, and is an undesirable effect in pulsed light/time-domain optical studies.

To minimize temporal dispersion in multi-mode fibers, a gradient-index design can be employed. Graded-index multi-mode fibers have smaller core diameters (50–90 μm) compared to step-index multi-mode fibers (> 100 μm). With graded-index fibers the index of refraction of the core decreases continuously between the central axis and the core-cladding boundary. The index gradient causes light rays to bend smoothly rather than sharply as they approach the core-cladding interface, and allows higher-order modes to travel more in the outer lower-index portion of the core. This effectively reduces the differences in propagation time, thereby reducing temporal modal dispersion.

While graded-index fibers reduce modal dispersion, two other kinds of dispersion still exist: material dispersion and waveguide dispersion. In essence, wavelength-dependent fiber material properties and optical/geometric characteristics of the sensors can cause spatial and/or temporal separation of the different optical wavelengths if a broad-band source is used. Material dispersion arises from the multiple wavelengths present in a laser pulse that result in different refractive indices at each wavelength. This effectively causes each wavelength to travel at different speeds and disperse. Similarly, waveguide dispersion (present in step-index single mode fibers as well) also occurs due to the different wavelengths present in a light pulse. Here, the different traveling speeds are due to core and cladding. Further, the other basic components of the fiber sensor system (Fig. 17.1) also exhibit wavelength-dependent characteristics. These considerations must be taken into account for specific sensor system implementations, especially those employing broad-band sources for tissue spectroscopy with pulsed sources for time-of-flight sensor measurements.

The differences amongst multi-mode and single-mode fibers and step-index and graded-index fibers lead to interesting design considerations when optimizing fiber-optic sensors.

Specific fiber types may be chosen based on dispersion characteristics, transmission capabilities, coupling efficiencies, mechanical flexibility, and biological compatibility, among other criteria.

Since single-mode fibers retain the coherence properties of the input laser light, they are commonly employed for fiber-based interferometers (e.g., OCT sensors). On the other hand, multi-mode fibers with larger *NA* and core diameter are typically simpler to couple and can transport and collect higher light signals that may be necessary for achieving adequate signal to noise. The choice of appropriate fiber materials must also be considered, to minimize transmission losses for probing tissues at specific wavelengths; while generally not a major consideration in the visible range, this becomes more important for UV and IR fiber sensors.

17.3.3 Fiber Bending

Bending of an optical fiber can lead to light leakage at the area of sharp curvature. When a light ray reaches the fiber bend, higher-order modes that hit the cladding beyond the critical angle leak out of the fiber. Such an effect is less likely for lower order modes. However, low order modes can be converted to higher order modes that then may leak out the fiber at subsequent bends. Although large bends lead to greater light loss, a series of small micro-bends can also lead to a significant decrease in transmission.

While light loss through micro-bending is generally an unwanted side-effect that leads to signal degradation, the effect can have a number of useful applications. A fiber may be purposely twisted to homogenize and smooth the output beam – also called mode mixing – by converting lower order modes to higher order modes. Similarly, mode stripping is the application of local pressure to optical fibers to modulate the output light patterns exiting a fiber.

Temperature and pressure fiber optic sensors can be constructed to utilize the light loss from fiber bending. Figure 17.5 illustrates a simple micro bending sensor. Optical fibers are positioned between a set of plates with indented groves such that when pressure is applied to the plates micro-bending occurs leading to a decrease in transmitted optical power. Alternatively, the fibers may be encased in a temperature sensitive material which constricts when cooled or heated, again creating light modulation through the resultant change in fiber shape.

17.3.4 Fiber Tip Geometries for Biomedical Sensors

Light exiting or entering an optical fiber can be redirected by sculpting or polishing the fiber tip into a desired configuration. Modified fiber tips are appealing compared

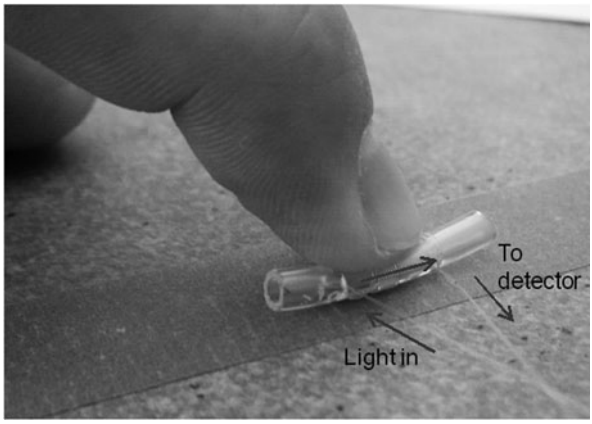


Fig. 17.5 Fiber optic pressure sensor based on light loss from fiber bending. *Arrows* indicate the direction of light travel. As the finger applies pressure to the tubing, fiber bending results in light loss that is manifested as a loss in detected light signal (from www.sensors.com 2008)

to external optical components for a number of reasons. First, they are more compact and reduce overall system cost by reducing the number of total system components. Second, appropriate tip modification improves optical efficiency, reliability and durability by minimizing the number of optical interfaces (coupling/insertion losses, unwanted reflections), compared to the use of external optics coupled to the fiber end.

A variety of modified fiber tip strategies are available. Implementations that can adjust the spot size and divergence of emitted or collected light and/or steer and direct the beam are all possible. In the following section we briefly review the theory, applications and construction of tapered, ball and side-firing fiber tips. We also discuss the rationale for shielding cap use in tissue.

The described fiber tips are often employed in reflectance sensors to adjust their probing and detection properties, while minimizing sensor size by avoiding the addition of external optics (see Section 4.1). Spherical and linear diffusers, typically employed in interstitial applications, will be described in Section 17.4.3.1.

17.3.4.1 Tapered Fiber Tips

Tapered (Fig. 17.6b) and ball-lens terminated (Fig. 17.6d) shaped fiber tips increase the light intensity (fluence rate) close to the fiber tip.

Tapered fibers are constructed by increasing or decreasing the diameter of the fiber at either the distal (emitting/detecting) or proximal (laser incident) end. An "up" tapered fiber increases the diameter of the fiber end while a "down" tapered fiber decreases the fiber diameter (see Fig. 17.7).

Tapered fiber optics is governed by [4]:

$$d_1 \sin a_1 = d_2 \sin a_2 \tag{17.9}$$

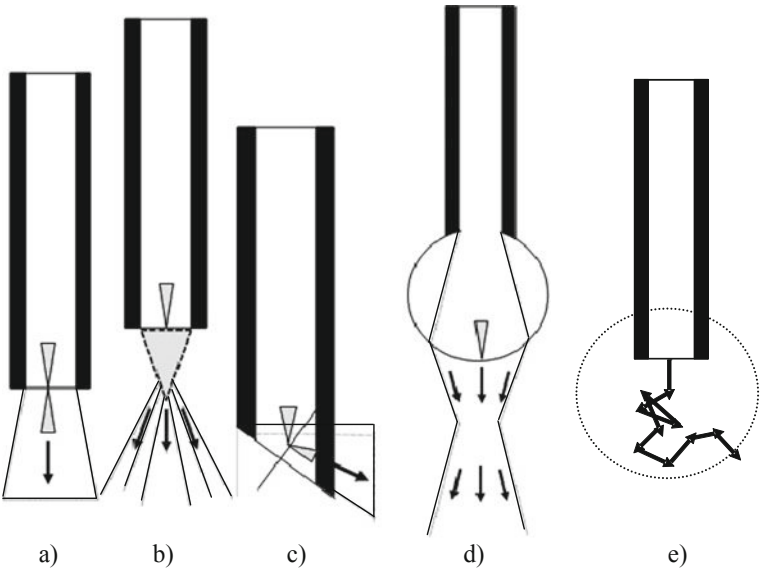


Fig. 17.6 Typical fiber tip modifications (**a**) plane-cut fiber (**b**) tapered fiber (**c**) side-firing (**d**) ball-lens terminated fiber (**e**) spherical diffuser. *Small triangles* and *arrows* are guides for the eye to indicate the direction of light propagation (after Verdaasdonk and Borst [3])

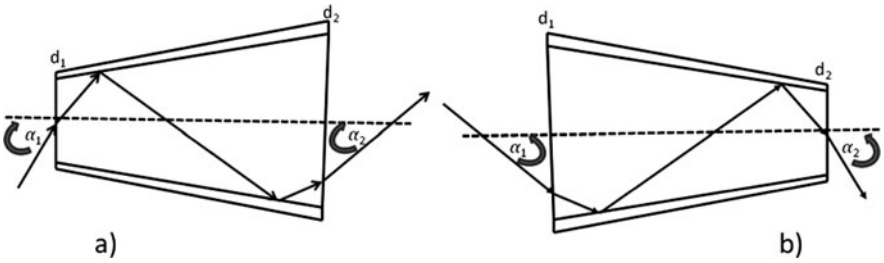


Fig. 17.7 Geometry of incident and reflected rays for tapered fiber tips: (**a**) up (**b**) down. The entrance α_1 and exit α_2 angles of the light rays are governed by Eq. (17.9) (after Polymicro Technologies [4])

Light entering the taper will be reflected back and forth with a decreasing (down tapered) or increasing (up tapered) angle of incidence along the taper-length at the core-cladding interface. This will continue until the incident angle of the traveling photons exceeds the critical angle. At this "leaking" point, the photons are refracted out of the fiber. Hence, a "down" tapered tip can be employed to increase the divergence (numerical aperture) while decreasing the spot size of light exiting near the fiber tip. Conversely, an "up" tapered fiber can be used to reduce the *NA* and enlarge the spot-size of the exiting light. Thus, a down taper tip is a light "spreader" while an up taper is a light "concentrator."

17.3.4.2 Ball Shaped Fiber Tips

Fiber tips can be also constructed to act as parallel focusing lenses. Such tips include hemispherical elements attached to a plane cut fiber, or ball-shaped endings which are sculpted from the fiber core. The resulting curvature focuses the incoming light rays such that a higher density of light is present near (beyond) the fiber tip (Fig. 17.6d).

A lens bends and focuses light by refracting it. The refraction is accomplished by having different rays pass through different optical thicknesses, thus enabling a lens to focus a (collimated) beam to a (focal) point. Typically, varying optical thickness means a transparent optical element made from a material with a given refractive index n_0, shaped to have the desired physical thickness profile (e.g., a lens-shape). However, especially when working in confined spaces, it is possible to fix a shape (e.g., cylinder) and have the refractive index vary in a particular fashion, generating the desired "variable optical thickness" effect. Lenses manufactured with this approach are known as GRINs (GRadient INdex). For example, in the context of fiber optics, cylindrical shaped GRIN lenses with a radially-varying refractive index profile (highest at the centre, dropping off with radial distance, often parabolically, towards the edges) are very popular. Since the end faces of the GRIN lens are flat, relatively straight-forward coupling to a plane-cut termination of an optical fiber is enabled. In this chapter, several optical fiber sensors employ GRIN lenses in their design (see Figs. 17.21 and 17.22).

Alternatively, a fiber tip can be sculpted to a ball shape to adjust beam divergence and spot size and vary the depth of focus of a probing light source. In air, the irradiation properties of spherical tips can be described using paraxial theory [3], which provides an estimation of the position of highest light intensity as a function of the sphere radius.

Paraxial theory predicts that a beam exiting a ball shaped tip in air will be focused to a point at a distance, F_1, given by:

$$F_1 = R_s \frac{n_0}{n_1 - n_0} \tag{17.10}$$

where R_s is the radius of the ball tip.

A beam with finite diameter (i.e. containing multiple rays) will create an irradiation profile with a "waist", which is the convolution of "multiple" focal points displaced laterally from the central axis (all also located at also F_1).

Equation (17.10) predicts that for a ball-shaped tip positioned in water, the focal point and region of highest intensity will be located more distally to the ball tip compared to when the ball tip is in air (as n_0 is now greater). In water, the resulting rays are focused so far from the optical axis that the overall beam diameter at F_1 is larger than the beam diameter at the exit surface of the ball tip. This is despite the fact that the individual rays forming the beam profile are still focused distally at F_1.

In this case, the region of highest intensity is now located between the distal ball surface and F_1.

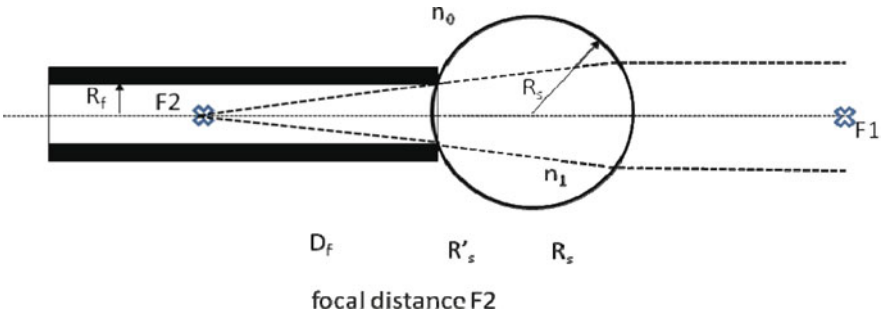


Fig. 17.8 Geometry of ball-lens tipped fiber used in paraxial theory: R_f is the fiber radius, R_s is the ball tip radius, D_f is the distance between the focal point inside the fiber to the end of the fiber, R_f is the radius of the fiber, $F_2 = D_f + R'_s + R_s$ the proximal focal point, while n_1 and n_0 are the index of refraction of the ball tip and air, respectively (after Verdassdonk and Borst [3])

Paraxial theory can also be used to estimate the conditions under which a converging beam profile is converted into a diverging beam profile. As illustrated in Fig. 17.8, consider a set of rays originating from a point F_2 within the fiber, that converge to a distal focal point, F_1. F_2 is also known as the proximal focal point. Using geometry, a relationship can be derived from the radius of the ball tip, R_s, the fiber radius, R_f, the refractive index of the surrounding medium, n_0, the index of refraction of the fiber/ball material, n_1, and the divergence angle, θ.

From Fig. 17.8:

$$F_2 = D_f + R'_s + R_s \tag{17.11}$$

Here $R'_s = \sqrt{R_s^2 - R_f^2}$ is the radius minus the portion of the ball that is within the fiber itself while $D_f = R_f / \tan\theta$ is the distance from F_2 to the end of the fiber.

Substituting for D_f in Eq. (17.11):

$$F_2 = R_s + \sqrt{R_s^2 - R_f^2} + R_f / \tan\theta \tag{17.12}$$

Paraxial theory states the proximal focal point (located within the fiber), F_2, is given by:

$$F_2 = R_s \frac{n_1}{n_1 - n_0} \tag{17.13}$$

Inserting Eqs. (17.12) into (17.13) and expanding, we find that the ratio of the sphere radius and sensor is governed by the following relationship:

$$\frac{R_s}{R_f} = T(N - 1) + \frac{\sqrt{T^2 - N(N-2)}}{N(N-2)} \tag{17.14}$$

where $N = n_1/(n_1 - n_0)$ and $T = 1/\tan\theta$.

The ratio, R_s/R_f, can be used to estimate under what conditions a ball-shaped tip loses its focusing effect. However, it should be noted that the theory is only a first order approximation and ignores the effects of secondary reflections, which can lead to significant deviations. In such cases, ray-tracing provides more reliable predictions of the resulting beam profiles [3].

Both tapered and ball-shaped fiber tips provide a flexible approach for focusing incoming and outgoing light and can be optimized for a variety of applications. These include improving the coupling efficiency with the input laser or other optical components, improving the detection efficiency with a photodiode or photomultiplier tube or, for therapeutic applications, performing high power laser ablation.

17.3.4.3 Side-Firing Tips

Side-firing tips (Fig. 17.9c) redirect the light over a range of angles away from the fiber axis. Side-firing fibers are often employed in therapeutic applications such as photodynamic therapy, thermal therapy, or tissue ablation target scenarios with limited room or maneuverability. Recently, side-firing tips have also been employed as optical sensors for measuring the radiance distribution in turbid media [5, 6].

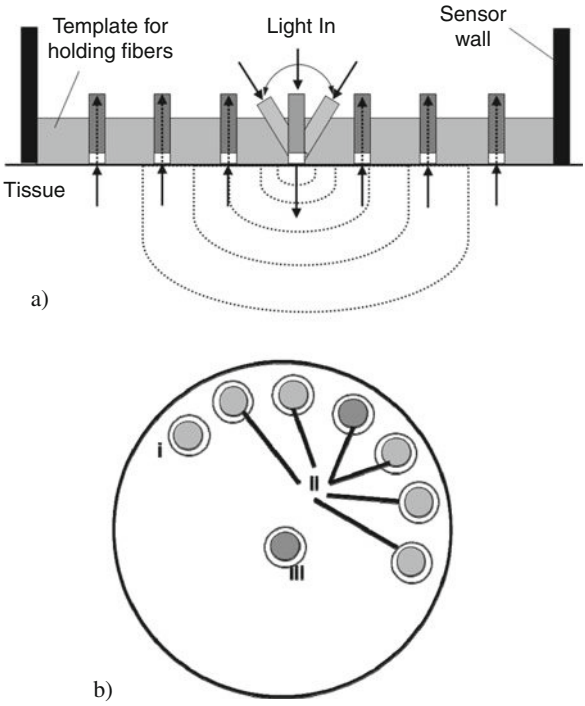


Fig. 17.9 (**a**) Linear arrangement of reflectance sensors (after Wang and Jacques [9], with permission from Applied Optics) (**b**) Circular arrangement of reflectance sensors: (i) source fiber (ii) detection fibers (iii) calibration source (after Ref. [10])

An angled-tip with $\sim 10^{\circ}$ angle can also be used to minimize specular reflection for light-scattering reflectance and/or OCT sensors.

Typically, a side firing fiber with an *NA* of 0.22 should be polished at an angle >40° to ensure total internal reflection. For those light rays reflected at the surface, the exiting beams will be further focused by the curvature of the core, much like a cylindrical lens.

It is important to note that even if the fiber tip is polished at an angle exceeding the angle of total internal reflection, a large portion of rays can still be refracted out the fiber tip. This is because the light traveling within the fiber carries a broad angular distribution such that not all rays can be internally reflected (recall discussion of fiber modes). In addition, when the surrounding environment is water or tissue, the total reflection capabilities of many side-firing tips (calculated with air as the external medium) are lost. In such cases, a transparent shield can be positioned around the tip to allow a surrounding air medium next to the tip, thus preserving its side-firing properties.

Alternatively, right angle micro prisms can be attached, using optical adhesive, to the end of a plain-cut fiber [5, 6]. Here, total internal reflection occurs at the prism's hypotenuse to direct light to/from the fiber. One advantage of this approach is that the prisms can be coated with a thin layer of aluminum to increase the amount of light reflected at the prism hypotenuse (and to capture the off-axis rays that are beyond the total internal reflection condition). Since aluminum is an efficient reflector over a broad range of wavelengths, Al-coated prisms are particularly useful to minimize light leakage in spectroscopic and fluorescence applications. However, the tips can be extremely fragile and typically require a surrounding glass cover to prevent the prisms from being unattached from the fiber during clinical utilization.

17.3.5 Fiber Tip Manufacturing

The majority of the specialized fiber tips described in this section are available via a number of commercial vendors, for example OZ Optics (www.ozoptics.com), Polymicro Technologies, (www.polymicro.com), and LaseOptics (www.laseoptics.com) among others (note that these company names are included for illustrative purposes only, and in no way imply commercial endorsement by the authors) These vendors employ state of the art laser micro-machining equipment for sculpting their fiber tips. However, in-house approaches often offer a more economical and quicker alternative to commercial vendors.

Verdaasdonk and Borst [3] described a simple method for manufacturing ball-shaped fiber tips using a high-temperature burner. Following stripping of the buffer and cladding, a standard silica fiber is exposed to a flame with local temperatures exceeding 2000°C. As a result of the surface tension of the liquid silica, the melted core forms a near spherical droplet that eventually solidifies into a ball-shaped tip.

The authors found that the best approach was to position the flame horizontally with the fiber held beneath the flame at 20–45° from the vertical axis.

A similar approach can be employed to manufacture tapered fibers. Here, the last 1–1.5 cm of the fiber is stripped down to the buffer. The fiber is then positioned in an upside down fashion with a small weight (several hundred grams) attached to the tip. By heating the fiber at the midway point of the stripped region, gravity then draws the molten silica into a taper via the downward pulling weight. The taper angle can be refined by adjusting the fiber diameter, melted volume, and amount of pulling weight. Finally, the taper end is shaped to a point by melting or polishing.

Right-angle side-viewing fiber tips can be created by polishing a plane-cut fiber at an angle equal to or slightly exceeding the critical angle in the medium. In addition, a sputtering machine can be employed to coat the polished surface with a thin, reflecting metal layer (i.e. gold, silver, aluminum) to enhance and supplement the total internal reflection effect.

Finally, shielded fibers can be constructed by encasing the sculpted fiber tips using a clear glass sheath. The glass sheath can be constructed using a glass capillary tube of appropriate diameter by melting the end of the tubing with a flame. The sheath is then attached to the fiber at a location downstream from the tip using a clear UV-cured or chemically cured epoxy, thereby creating an air bubble inside to preserve the tip/air interface.

17.3.5.1 Biomedical Deployment

Combinations of plane-cut, GRIN/ball lens/scattering-sphere terminated, tapered, and side-firing geometries have been employed in the construction of various optical fiber sensors. The different combinations allow for flexibility in desired probing depths and/or collection efficiency of detected light. The relative merits of the selected probe designs will be discussed in subsequent sections of this chapter in the context of a particular modality/application.

17.4 Fiber-Based White Light Spectroscopy

17.4.1 Reflectance Geometry

17.4.1.1 Conventional Reflectance Sensor Design

Light that enters tissue will be scattered and absorbed. The portion of scattered light that escapes the tissue in the backwards hemisphere is termed the diffuse reflectance (e.g., semi-infinite planar geometry, surface irradiation conditions). The diffuse reflectance contains information regarding the scattering and absorption properties of the tissue. The non-invasive nature of surface-probe-based reflectance measurements can furnish a well sampled data set close to and far from the source, yielding

a spatial (and perhaps spectral) reflectance profile well suited for fitting to a theoretical model. As summarized in Table 17.1, the measured reflectance data can be employed to retrieve tissue optical properties, chromophore concentrations, and tissue scattering parameters.

As shown in Fig. 17.9, conventional reflectance sensors employ a single excitation source that injects light into the tissue surface, surrounded by several detection fibers (5–10) at varying distances. To extract the tissue optical properties, the measured reflectance must be fit to an analytic equation, which allows for separation of the absorption and scattering properties [7–9]. The diffusion approximation for a homogeneous, semi-infinite slab is often employed as a model for the purposes of data fitting. Since diffusion theory is valid at positions greater than several mean-free paths (mfp) from the source, this implies source-sensor positions between 0.2 and 2 cm for most tissues. A key advantage of reflectance geometry is that only *relative* data is necessary for determining the optical properties.

Two approaches have been devised for fitting reflectance data. The first is to employ a normal-incidence source and fit the spatially resolved reflectance profile. To enable quantitative data fitting, the entire reflectance profile is normalized to the measured signal at a selected source position. Accurate optical property separation and determination requires that at least one detector fiber be positioned close to the source (∼0.75–1 mfp). This is because measurements close to the sources are more indicative of the reduced scattering coefficient, while optical attenuation far from the source represents a combination of both scattering and absorption [10], which is embedded in the effective attenuation coefficient, μ_{eff}.

An alternative configuration, suggested by Wang et al. [9], is to employ an oblique incident source. Here the relative shift in the center of diffuse reflectance from the entry point is employed as additional information to extract the tissue optical properties.

The detection fibers can be packaged in a linear [9] (Fig. 17.9a) or in a circular (Fig. 17.9b) arrangement [10]. The advantage of the circular geometry is that a single source, centered with respect to the detection fibers, can be employed for convenient calibration. The large range of source-detector separations results in a range of reflectance measurements that can span ∼ four orders of signal magnitude, complicating simultaneous reflectance measurements at all detector positions. To reduce the dynamic range requirements, neutral density filters are often utilized to attenuate the signals at the closer sensor distances [10]. An alternative approach, suggested by Fuchs et al. [11], is to employ variable integration times (hundreds of milliseconds to a few seconds range) combined with an increased number of detection fibers at larger source-sensor separations, to enhance the detection of low signals.

17.4.2 Depth Sensitivity

Depth-sensitive detection is important for fiber-optic assessment of layered structures. Specific clinical applications include evaluating the dyplastic state of a

particular tissue layer or determining the local uptake of light-activated drugs. Depending on the specific application, a reflectance sensor can be designed to preferentially probe the superficial layer (e.g., 200–400 μm below tissue surface) or furnish information about deeper tissue layers (e.g., 1–2 cm below the surface).

In Section 17.4.1.1, we described the conventional design for typical reflectance sensors. These sensors are not optimized for depth localization. In the following, we focus on design approaches that can be employed to refine the depth selectivity of a reflectance sensor.

In general, two approaches can be employed to separate single or superficially scattered photons from bulk photons that undergo multiple scattering events. The first approach is to maintain the conventional reflectance sensor geometry but employ polarization, time-domain, or frequency domain *gating* techniques by appropriate selection of source and detector characteristics. In this case, multi-layer light modeling may also be required to properly extract the appropriate depth information. The second approach is to adjust the *geometric* attributes of the source and detection fibers employed in the sensor. Here, geometrical parameters of interest are the source-detector (S-D) separation, fiber numerical aperture, fiber diameter, and probe-tissue spacing.

The principles behind these approaches, as well as specific selected examples from the literature, are provided in the following sections.

17.4.2.1 Steady-State Spectroscopy: Geometric Approaches

Source-Detector Spacing

Patterson et al. [12] were able to obtain the average depth from which reflectance information is acquired. From diffusion theory, the steady-state, semi-infinite geometry solution for the fluence rate at a point P(κ,z) in a turbid medium (e.g., tissue) is given by:

$$\phi(\kappa, z) = \frac{1}{4\pi D}\left(\frac{\exp\left\{-\mu_{\text{eff}}\left[(z-z_o)^2+\kappa^2\right]^{1/2}\right\}}{\left[(z-z_o)^2+\kappa^2\right]^{1/2}} - \frac{\exp\left\{-\mu_{\text{eff}}\left[(z-z_p)^2+\kappa^2\right]^{1/2}\right\}}{\left[(z-z_p)^2+\kappa^2\right]^{1/2}}\right) \tag{17.15}$$

where κ is a radial distance coordinate, z is a depth coordinate (i.e. below the tissue surface), $D = 1/3\left(\mu_a + \mu_s'\right)$ is the diffusion coefficient, $\mu_{\text{eff}} = \sqrt{\frac{\mu_a}{D}}$ is the effective attenuation coefficient, $z_o = \left(\mu_s'\right)^{-1}$ is the distance of initial scatter from an isotropic point source, $z_e = (5.91)D$ is derived from the extrapolated boundary condition and $z_p = -z_o - 2z_e$ is the location of a negative image source required to meet a zero-fluence-rate boundary condition [12]. Physically, the extrapolated boundary condition assumes that the fluence goes to zero at a certain distance above the tissue surface [13].

The probability that a photon would be re-emitted from the tissue slab and detected by the fiber sensor at location, r, located at the surface, can then be

described by an escape function, E, which is equivalent to the photon current across the boundary at $z = 0$, is given by Fick's Law as:

$$\left(-D\frac{d}{dz}\phi\right)_{z=0} \tag{17.16}$$

Combining Eqs. (17.15) into (17.16) gives the escape function $E(\kappa, \theta, z, r)$:

$$E(\kappa,\ \theta,\ z,\ r) = \frac{1}{4\pi}\left[\frac{z\ \exp(-\mu_{\text{eff}}k)}{k^2}\left(\mu_{\text{eff}} + \frac{1}{k}\right) - \frac{z_p \exp(\mu_{\text{eff}}l)}{l^2}\left(\mu_{\text{eff}} + \frac{1}{l}\right)\right] \tag{17.17}$$

where k and l are the positive roots of:

$$k^2 = (r - \kappa\cos\theta)^2 + \kappa^2\sin^2\theta + z^2 \tag{17.18a}$$

$$l^2 = (r - \kappa\cos\theta)^2 + \kappa^2\sin^2\theta + z_p^2 \tag{17.18b}$$

Using Eqs. (17.15) and (17.17), the change in reflectance, $\Delta R(r)$, due to a small absorbing perturbation, $\Delta\mu_a$ in a small incremental volume, dV, is then:

$$\Delta R(r) = \Delta\mu_a\phi\left(r'\right)\ dV\,E\left(r',\ r\right) \tag{17.19}$$

The mean photon-visit depth, $\langle z\rangle_r$, for a source-detector separation of r is then given by:

$$\langle z\rangle_r = \frac{\int_V \phi\left(r'\right)\ E\left(r',\ r\right) z\,dV}{\int_V \phi\left(r'\right)\ dV\,E\left(r',\ r\right)\ dV} \tag{17.20}$$

Using Eq. (17.20), Patterson et al. [12] examined the change in $\langle z\rangle_r$ as a function of $\delta^{1/2} = \sqrt{1/\mu_{\text{eff}}}$ (Fig. 17.10). They demonstrated that, using a simple empirical relationship, a reasonable estimation of $\langle z\rangle_r$ could be determined from knowledge of r and δ, alone:

$$\langle z\rangle_r = \frac{(r\delta)^{1/2}}{2} \tag{17.21}$$

From Fig. 17.10 it is clear that short source-detector separations (lower line) result in shorter penetration depths while long source-detector separations (upper line) result in deeper sampling volumes. Furthermore, the mean penetration depth increases with decreasing optical attenuation.

Multi-distance reflectance probes that rely on diffusion theory for data fitting typically involved source-detector separations of up to ~2 cm. Such a large probe size may pose serious problems when the optical properties vary significantly over the probe sampling volume (e.g., when evaluating layered structures such as epithelial tissue, or sampling heterogeneous tumors with small localized regions of hypoxia).

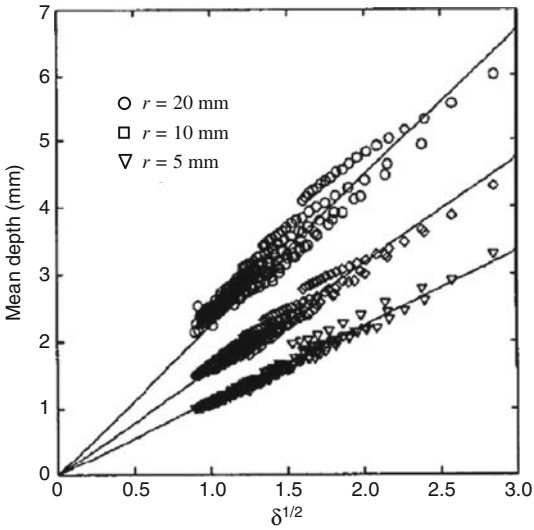


Fig. 17.10 Plot of $\langle z \rangle_r$, the mean photon visit depth vs. $\delta^{1/2}$, the square root of the penetration depth. The *symbols* represent experimental measurements. The *lines* are a linear fit of the resulting data for each source-detector separation, demonstrating an empirical relationship of $\langle z \rangle_r = \frac{(r\delta)^{1/2}}{2}$ as per Eq. (17.21). (Reprinted with permission from [12])

As demonstrated in Section 17.4.2.1, sensors with shorter source-detector separations (2–3 mm) must be employed to confine the volume of interrogation. An example of one such sensor (employed by Finlay and Foster [14]) is shown in Fig. 17.11 and was used to reveal significant heterogeneities in oxygen saturation in small murine tumor models. However, in these cases diffusion theory is no longer valid. As such, when short-detector spacing is required, alternative forward models and inversion algorithms must be employed during fitting for accurate recovery of optical properties and their spectral dependencies. A brief summary of the different modeling approaches employed for short s–d distances is presented in Table 17.2.

While an exhaustive explanation of all models is beyond the scope of this chapter, in general, the approaches can be divided into improved analytical solutions of the radiative transport equation such as the δ-P1 approximation, Born approximation and P3 approximations, or fast inversion algorithms based on Monte Carlo generated

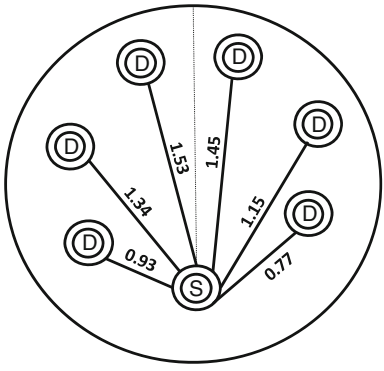


Fig. 17.11 Schematic of a small SD separation reflectance sensor. The source (S) and detector (D) separation distances (mm) are indicated in the figure (after Finlay and Foster[14])

Table 17.2 Analyzing reflectance sensor symbols: models employed for short source detector distances

Model	Application	References
P3 approximation to RTE	Homogeneous medium	Finlay and Foster [14]
δ-P1 approximation to RTE	Homogeneous medium	Seo et al. [15]
Scaled MC	Homogeneous or layered medium	Kienle and Patterson [16]
Born approximation to RTE	Layered medium	Kim et al. [17]
MC-based neural networks	Homogeneous medium	Pfefer et al. [18]
MC-diffusion hybrid	Homogeneous medium	Alexandrakis et al. [19]
Perturbation MC	Layered media	Seo et al. [20]

MC Monte Carlo, *RTE* Radiative Transfer Equation

data (Table 17.2). While analytical solutions carry the advantage of rapid forward calculations, they lack the accuracy of Monte Carlo simulations for many realistic scenarios and geometries. In contrast, Monte Carlo-based strategies are typically accurate but often require longer computational/inversion times.

Specialized Fiber Optic Sensor Geometries

Another approach to superficial depth localization is to employ source-detector separations and geometries that are virtually independent of scattering parameters, such that the path-length is nearly constant across all wavelengths. This is in contrast to the model based approach because it effectively removes the requirement for diffuse light modeling. In these cases, the attenuation effectively reduces to the Beer-Lambert Law, with the change in the measured signal being dictated by the absorption coefficient at each wavelength.

Mourant et al. [21] were the first to demonstrate experimentally that an optimum S-D separation of approximately 1.75 mm resulted in reflectance measurements that were virtually independent of the scattering properties of the measured phantom. These findings were later verified theoretically using diffusion theory [22].

Amelink and Sterenborg [23] developed an approach termed differential-pathlength spectroscopy (DPS) for measuring the optical properties of superficial layers of tissue. The method employs a specialized measurement geometry to reject diffuse light such that the measured signal is relatively insensitive to the surrounding optical properties.

A dual-use delivery-collection fiber (dc) positioned at the tissue surface is used to both illuminate and collect reflected light. A second collection fiber (c) is positioned adjacent to the dc fiber. The c-fiber only detects light that is reflected from the tissue sample. The two fibers are positioned touching side-by-side such that their core-to-core distance is $1.2d_{\text{fiber}}$ where the 1.2 accounts for the presence of the fiber cladding and core. Light collected from the dc fiber contains a combination of both singly backscattered light from *small sample depths* and multiply-scattered, diffuse light from deeper depths.

The principle behind DPS is that when the mean-free-path of photons is much larger than the diameters of the dc- and c- fibers, the probability of collecting *multiply* scattered light from deeper depths is roughly equal for both fibers. Furthermore, the probably of collecting *single* scattered from *deeper* depths is also similar. Subtraction of the collection fiber signal, J, from the delivery-collection fiber signal, I, yields a differential measurement, R, that is representative of single-scattered light that predominates from shallow depths.

In this case, the backscattered signal, R^{single}, is approximated by:

$$R^{\text{single}} = C_{\text{app}} \frac{1}{4\pi} \int_{\Omega NA} d\Omega p\,(\Omega)\, Q_{\text{sca}} \rho A \tag{17.22}$$

where C_{app} is a constant that depends on apparatus design and is a function of probe tip, reference material and other components, $p(\Omega)$ is the phase function of the scattering medium, Q_{sca} is the scattering efficiency, ρ is the scatterer concentration (number/volume), and A is the scattering particle area.

For fibers with small numerical aperture (i.e., $<\sim 0.22$), $p(\Omega) = p(180)$ and Eq. (17.22) simplifies to:

$$R^{\text{single}} \approx C_{\text{app}} \frac{1}{4\pi} \int_0^{2\pi} d\phi \int_{\pi - NA}^{\pi} d\theta \sin(\theta)\, p(180)\, Q_{\text{sca}} \mu_s \tag{17.23}$$

$$R^{\text{single}} = C_{\text{app}} p(180)\, \mu_s \tag{17.24}$$

where the scattering coefficient is $\mu_s = Q_{\text{sca}} \rho A$.

Equations (17.23) and (17.24) have been well described in the literature [24]. However, for most biological tissues scattering is significant and the mfp is typically much smaller than the fiber diameter. In this case, a relatively large number of multiple-scattering events occur at shallow depths close to the fiber tip, and a simple analytical expression is not easily derived. To examine the scattering effects in this regime, Amelink et al. [23] replicated the differential signal, R, using Monte Carlo simulations for a wide range of optical properties (μ_s=15, 25, 50, 80 mm^{-1}, μ_a = 0, 0.2, 0.4, 0.6, 0.8 mm^{-1}, $g = 0.9$) typical of biological tissues between 450 and 1000 nm for a range of fiber diameters ($d_{\text{fiber}} = 200, 400, 600, 800$ μm).

As shown in Fig. 17.12 (for an anisotropy factor of 0.9 and in the absence of absorption) the differential signal R is almost linearly proportional to the reduced scattering coefficient, $\mu_s' = (1-g)\,\mu_s$, such that $R \approx C_1 \mu_s'$ where C_1 is a proportionality constant. Furthermore, the authors demonstrated a Beer-Lambert dependence of the differential signal, R, on the absorption coefficient, μ_a, with a differential path length approximated by $\tau = C_2 d_{\text{fiber}}$. Here, C_2 is a proportionality constant that is ~0.95 for most fiber diameters and optical properties. Provided that $\mu_a d_{\text{fiber}} < 0.6$, this value can be used to approximate the differential signal to better than 15%.

Based on the results from the MC simulations the differential signal can be formally summarized as:

$$R = C_1 \mu_s' \exp(-0.95\, d_{\text{fiber}} \mu_a) \tag{17.25}$$

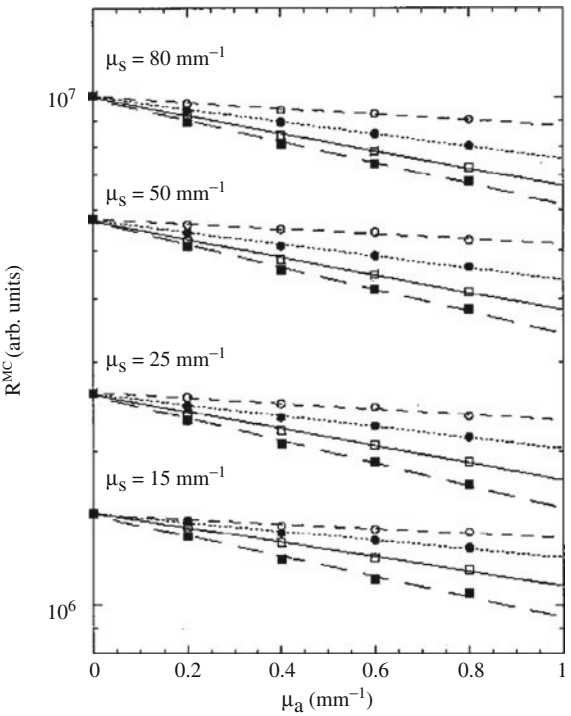


Fig. 17.12 Monte Carlo simulation showing Beer's Law behavior, with an apparent path-length being dependant only on fiber diameter: $d_{fiber} = 200$ μm (*open circles*, *dashed curves*), $d_{fiber} = 400$ μm (*filled circles, dotted curves*), $d_{fiber} = 600$ μm (*open squares, solid curves*), $d_{fiber} = 800$ μm (*filled squares, dashed-dotted curves*). (Reprinted with permission from [23])

Using Eq. (17.25), the optical properties can be extracted using a non-linear multi-parameter fit. However, when the fiber diameter is smaller than the mean free path, single scattering dominates the detected signal. In this regime, the apparent "differential" path-length can be adjusted by changing the fiber diameter. As such, by varying the fiber diameter of the probe, d_{fiber}, the volume of interrogation can by adjusted for the application of interest.

The technique can, therefore, provide highly localized information regarding the absorption and scattering properties of superficial structures with very minimal adjustments in equipment or design. Differential path-length spectroscopy has been employed to perform optical assessment of human breast tissue [25] and oral mucosa [26] as well as localized drug concentrations during PDT [27].

Illumination-Collection Orientation

A possible method to improve depth discrimination of plane-cut fiber reflectance probes is by tilting the orientation of the source or detector fibers at oblique angles to the tissue surface.

Using Monte Carlo simulations, Wang et al. [28] investigated the change in depth sensitivity for reflectance sensors when detection fibers with oblique orientation were employed. In their study, the sensor geometry utilized a single orthogonal

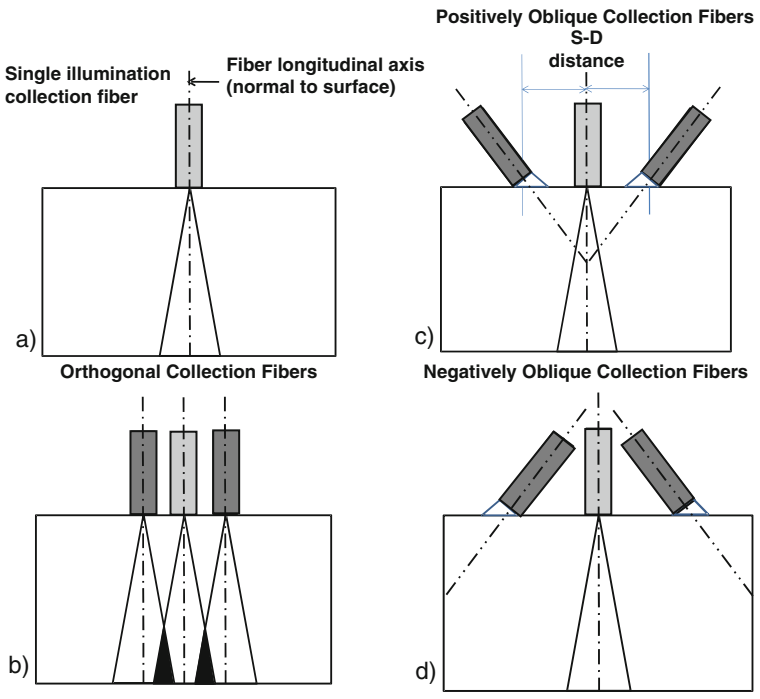


Fig. 17.13 Schematic of the fiber geometries employed for analyzing the effects of collection fiber orientation angle on average depth of penetration of detected photons in reflectance geometry. Grey boxes represent optional index-matching spacers (after Wang et al. [28])

source fiber with detection fibers having varying orientation relative to the surface. The simulated tissue environment mimicked optical properties and geometries representative of human cervical tissues with a thin 450 μm epithelial layer and underlying 10 cm of stromal tissue.

Figure 17.13 shows the various geometries employed in their study: (a) a single orthogonally oriented illumination-collection fiber; (b) conventional orthogonal fiber geometry; (c) positively oblique detection fibers; and (d) negatively oblique detection fiber geometry. An index matching spacer can be used for tilted fibers to minimize reflections at the air tissue interface and improve detection efficiency. Their results demonstrated that, compared to a perpendicular orientation, positive tilting of the detection fibers results in overlapping illumination-collection volume that is shifted closer to the tissue surface.

Using a zero degree (normal incidence) single fiber geometry as reference (Fig. 17.13a), the authors observed that the mean penetration depth decreases (becomes more shallow) and photon spatial distribution becomes more localized (i.e. better defined) as the collection angle becomes more positive. Since the photons exit the tissue obliquely, they require fewer scattering events and smaller scattering angles before escaping to be detected. Furthermore, the slanted orientation of

the detector improves selection of photons diffusely reflected at small exit angles. The result is that one can obtain improved depth selectivity, and greater sensitivity to optical properties of shallower tissue regions [29–31]. Conversely, increasing the collection angle in the negative direction preferentially selects photons that have traversed deeper tissue regions, while also decreasing the overall peak to width ratio.

The primary limitation of this sensor-tilting approach is that, to maintain small probe dimensions and to prevent mechanical damage, the fibers are generally limited to a maximum tilt angle of ∼30°. For superficial layer sensing, however (Fig. 17.13(c)), this upper threshold still results in a sizeable contribution from photons that have sampled the deeper stromal regions.

A potential solution for obtaining a suitable slanted orientation, proposed by the group at the University of Texas, is to focus the incident beams using a half-ball lens. This approach has been described by Schwarz et al. [32].

Figure 17.14a shows the geometry for a ball lens probe. The total angular deviation of the incident ray, θ, is the cumulative results of refraction at the entrance and exit interfaces of the ball lens.

By geometry (Fig. 17.14b), the angle of incidence, θ_i, can be geometrically related to the distance from the centerline of the ball lens, d, and the ball lens radius and index of refraction, by:

$$\sin\theta_i = \frac{d}{R} \tag{17.26a}$$

$$\theta_i = \sin^{-1}\frac{d}{R} \tag{17.26b}$$

From Snell's law, the refracted angle, θ_2, at the entrance interface is then related to the index of refraction of air, n_{air}, and the balls lens, n_L:

$$n_{\text{air}}\frac{d}{R} = n_L \sin\theta_2 \tag{17.27}$$

Since $n_{\text{air}} = 1$, the refraction angle is then:

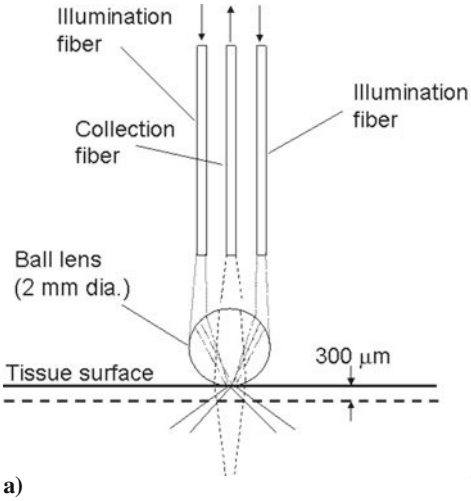
$$\theta_2 = \sin^{-1}\frac{d}{Rn_L} \tag{17.28}$$

Performing a similar calculation for the exit angle of refraction, the total angular deviation of the incident ray, θ, is expressed as:

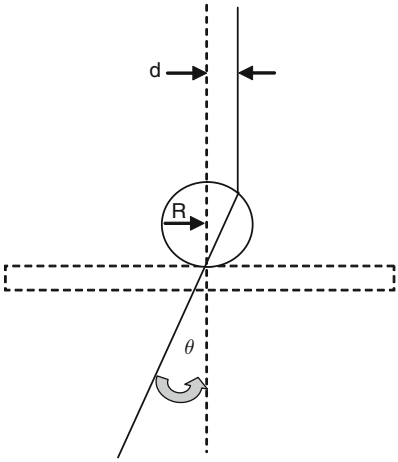
$$\theta = \sin^{-1}\left(\frac{d}{R}\right) - 2\sin^{-1}\left(\frac{d}{Rn_L}\right) + \sin^{-1}\left(\frac{d}{Rn_T}\right) \tag{17.29}$$

where n_T is the tissue refractive index.

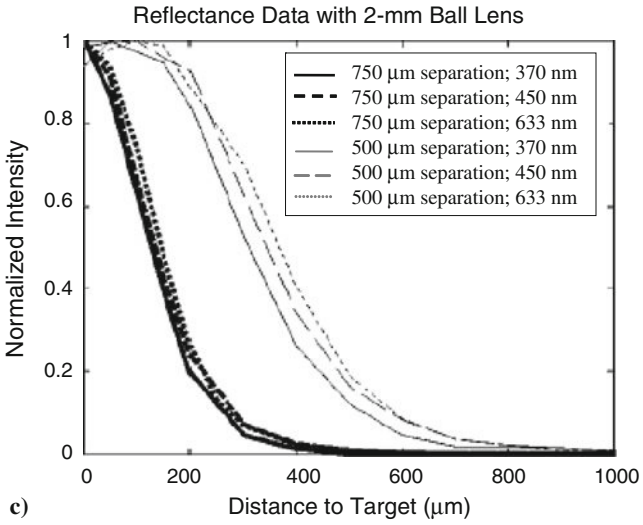
Using the lens, the incident and collected rays are refracted at oblique angles resulting in a similar distribution that can be obtained by physically tilting the fibers. As the fiber separation increases, the region of intersecting cones moves to



a)



b)



c)

Fig. 17.14 (**a**) Ball lens sensor composed of a central collection fiber and two adjacent illumination fibers. The separation distance of the illumination and collection can be used to adjust the probing depth of the ball lens sensor: illumination light rays (*solid lines*), detected light rays (*dotted lines*). As reference, the expected thickness of the epithelium (∼300 μm) is shown. (**b**) Geometry for Eq. (17.26). (**c**) Measured reflectance intensity for a thin scattering layer as a function of ball lens sensor to reflecting target separation, for two different ball lens sensor illumination-collection fiber separations. Note that the depth sensing capability of the sensor is demonstrated by the markedly different reflectance profiles at different illumination-collection fiber separations. (Reprinted with permission from [32])

increasingly superficial depths thereby allowing for preferential selection of shallow penetrating photons (Fig. 17.13b).

Figure 17.13c demonstrates the effects of the ball lens measured experimentally using the geometry of Fig. 17.13a. In these experiments, a thin reflectance target (white business paper) was positioned at different distances from the ball lens sensor, and the reflectance intensity measured as a function of sensor-to-target distance. Two different illumination-collection fiber separations were employed (500 and 750 μm). As expected, the normalized intensity drops faster at shallower depths, as the illumination-collection fiber separation increases.

As noted above, the penetration depth not only depends on physical geometry but on the optical properties of the tissue; an estimate can be obtained via Monte Carlo simulations [29]. In addition to the half-ball lens, a sapphire window can be positioned between the lens and tissue surface to further converge the collection-illumination cones closer to the tissue surface. Arifler et al. [29] have demonstrated that a half-ball lens probe with source-detector fiber separation of 900 and a 300 μm thick sapphire window can result in >98% of collected photons coming from the 300-micron thick epithelial layer (at wavelengths between 450 and 650 nm). The ball-lens sensor has been employed clinically for the diagnosis of cancer of the oral mucosa [33].

Numerical Aperture

The effect of numerical aperture on probing depth is complex and depends on the collection angle of the detection fibers. Using Monte Carlo simulations, Wang et al. [28] have demonstrated that, for collection angles between 0 and 20°, an increase in numerical aperture causes a shift to shallower probing depths. The opposite effect is observed for collection angles between 40 and 75°, whereby the expected probing depth increases with greater *NA*. In addition to depth selectivity, *NA* also plays role in signal to noise. For example, increasing the numerical aperture from 0.22 to 0.34 does not significantly change the probing depth of a reflectance sensor but increases the magnitude of the detected reflectance almost two-fold.

17.4.2.2 Polarization Reflectance Gating: Steady-State Spectroscopy

An alternative to mechanical modification of the source and detector fiber orientation is to employ reflected polarized light to preferentially select the detection depth (optical modification) [34]. Polarized light is quickly depolarized upon multiple scattering. As such, reflected photons that are able to preserve their incident linear polarization are typically representative of light scattering in *upper layer* tissues. In contrast, detection in a *crossed-linear* sensor picks up mostly depolarized light randomized by multiple scattering, indicative of *deeper tissue layers* (and/or any circularly polarized light that may be present). Hence, the crossed polarized reflectance can be subtracted from the parallel component to isolate reflectance originating from superficial depths. The resulting polarization measurement $D(\lambda)$ is given by:

$$D(\lambda) = \frac{I_{\parallel}(\lambda) - I_{\perp}(\lambda)}{I_{\parallel}^{S}(\lambda) + I_{\perp}^{S}(\lambda)} \tag{17.30}$$

Here $I_{\parallel}(\lambda)$ is the parallel polarized light relative to the incident light, $I_{\perp}(\lambda)$ is the perpendicular polarized light relative to the incident light, and λ is the wavelength of the incident light. The denominator is employed for normalization and is the sum of the perpendicular, $I_{\perp}^{S}(\lambda)$, and parallel, $I_{\parallel}^{S}(\lambda)$, components collected from a diffusing calibration medium. By normalizing, the detected signal now accounts for systematic sensor parameters such as the source and spectrometer spectral characteristics.

A simplified fiber polarization reflectance probe utilized by Johnson and Mourant [34] is shown in Fig. 17.15.

The probe is composed of a source (a) and two detection fibers for collection of parallel and cross polarized light. The source and one detection fiber are covered with a linear polarizing film while a different cross polarizing film covers the other detector. Myakov et al. [35] employed a similar sensor configuration to measure the reflectance spectra of oral cavity mucosa in vivo. The resulting spectra can be fit using Mie theory to obtain estimates of nuclear size for cancer detection. In this case, the polarization ratio is related to the average nuclear diameter, $\bar{d}$, and the number of nuclei per unit volume, ρ, by:

$$D(\lambda) = \frac{I_{\parallel}(\lambda) - I_{\perp}(\lambda)}{I_{\parallel}^{S}(\lambda) + I_{\perp}^{S}(\lambda)} a(\rho) B(\lambda) + b(\rho) S(\lambda) F(\lambda) + S(\lambda) \tag{17.31}$$

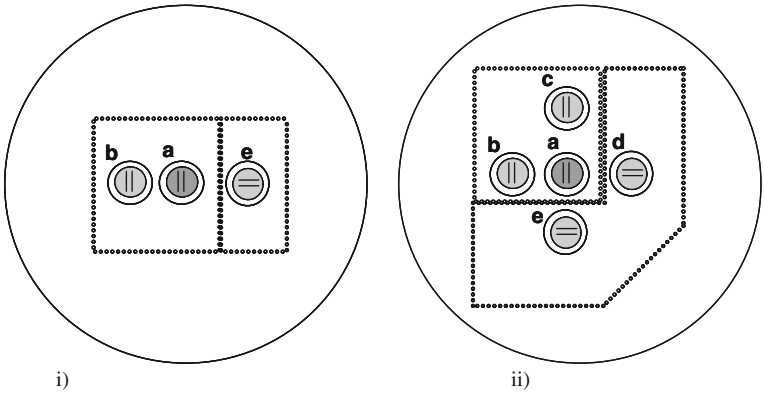


Fig. 17.15 Polarized reflectance sensor: The *dotted lines* within the sensor are representative of polarizing sheets while the *parallel lines* within the circular fibers are representative of the polarization orientation: (i) Polarized reflectance probe (*front view*) composed of an illumination channel (label **a**) with parallel (label **b**) and cross-polarized (label **c**) detection fibers (ii) This probe configuration employs an illumination fiber (label **a**) with two parallel detectors (labels **b** and **c**) and two cross-polarized detectors (labels **d** and **e**) such that all four polarization permutations can be measured (after Johnson and Mourant [34])

with

$$F(\lambda) = \int_{\theta_F} p\left(\theta, \lambda, \bar{d}, \Delta d\right) \sin\theta \, d\theta \tag{17.32}$$

$$B(\lambda) = \int_{\theta_\eta} p\left(\theta, \lambda, \bar{d}, \Delta d\right) \sin\theta \, d\theta \tag{17.33}$$

$$p\left(\theta, \lambda, \bar{d}, \Delta d\right) = \left\langle p\left(\theta, \lambda, \bar{d}, \Delta d\right)\right\rangle_{N(d)} = \int_0^\infty p\left(\theta, \lambda, \bar{d}, \Delta d\right) N(d) \, dd \tag{17.34}$$

where d is the diameter of the nucleus, $p\left(\theta, \lambda, \bar{d}, \Delta d\right)$ is the scattering phase function averaged over a distribution of nuclear diameters, $N(d)$, Δd is the standard deviation of the average nuclear diameter, θ is the angle between the detection and illumination angles, θ_D is the backward scattering angle range, θ_F is the range of scattering angles in the forward direction, $a(\rho)$ and $b(\rho)$ are nuclei density dependent parameters, and $S(\lambda)$ is the depolarization ratio profile of underlying tissue that can be determined experimentally. The term $p\left(\theta, \lambda, \bar{d}, \Delta d\right)$ is based on Mie theory – with the assumption that the nuclei are homogeneous spheres – and contains information regarding the nuclear diameter distribution, as well as, the spectral dependence of the scattering coefficient.

Groner et al. [36] employed a similar approach to reduce confounding specular reflection signals and enhance the contribution of light detected from deep tissue layers.

17.4.2.3 Time-Resolved Reflectance Spectroscopy

Diffusion theory can also be used to obtain the average penetration depth, $\langle z\rangle_r$, as a function of detection time for time-resolved fiber optic reflectance sensors [12].

In the time-domain case, we are interested in the fluence rate as a function of time at source-detector separation, r, and time, t, following an impulse source. Here, the decrease in photons that escape the tissue due to an absorption inhomogeneity in dt'at r' is:

$$\Delta R\left(r, t, t'\right) = \Delta\mu_a \phi\left(r', t'\right) dV \, E\left(r', r, t - t'\right) dt' \tag{17.35}$$

Here, $\phi\left(r', t'\right)$ and $E(r', r, t - t')$ are the time-resolved equivalents to the continuous wave fluence rate and escape function, respectively. They are given by:

$$\phi\left(\kappa, z, t'\right) = c\left(4\pi Dc\right)^{-3/2} t'^{-3/2} \exp\left(-\mu_a c t'\right) \times \left\{\exp\left[\frac{(z - z_o)^2 + \kappa^2}{4Dct'}\right] - \exp\left[\frac{\left(z - z_p\right)^2 + \kappa^2}{4Dct'}\right]\right\} \tag{17.36}$$

$$E\left(\kappa,\,\theta,\,r,\,t-t'\right)\;=\;\frac{1}{2}\left(4\pi Dc\right)^{-3/2}\left|t-t'\right|^{-5/2}\;\times\;\exp\left(-\mu_a c\left|t-t'\right|\right)$$
$$\times\left\{z\exp\left[\frac{\kappa^2}{4Dc\left|t-t'\right|}\right]\;-\;z_p\,\exp\left[\frac{l^2}{4Dc\left|t-t'\right|}\right]\right\} \tag{17.37}$$

where c is the speed of light in tissue.

Integrating overall possible time, t', gives the total reduction in photons, $R(e, t, t')$ as:

$$R\left(r,\,t\right)\;=\;\Delta\mu_a dV\int_{|r'|/c}^{t-|r-r'|/c}\phi\left(r',\,t'\right)E\left(r',\,r,\;t-t'\right)dt' \tag{17.38}$$

The mean penetration depth for a source-detector separation of r and time of detection, t, is described by:

$$\langle z\rangle_r\;=\;\frac{\int_V z dV\int_{|r'|/c}^{t-|r-r'|/c}\phi\left(r',\,t'\right)E\left(r',\,r,\;t-t'\right)dt'}{\int_V dV\int_{|r'|/c}^{t-|r-r'|/c}\phi\left(r',\,t'\right)E\left(r',\,r,\;t-t'\right)dt'} \tag{17.39}$$

Equation (17.39) can be evaluated using numerical integration. An example is shown in Fig. 17.16 for a nominal set of optical properties. As expected, it is seen that increasing the detection time results in deeper sampling volumes. As such, the depth of penetration can be adjusted at a single source-detector distance by varying the detection times of a time-domain reflectance sensor.

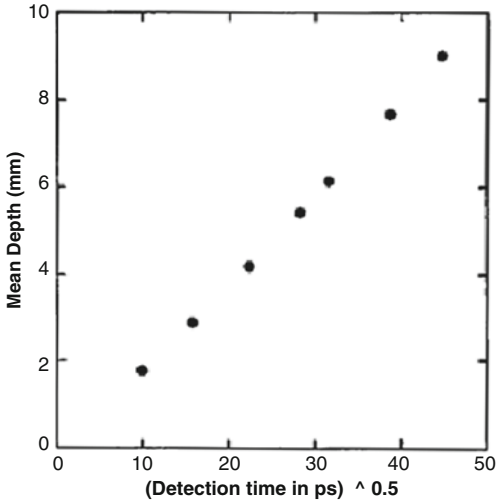


Fig. 17.16 Plot of mean photon visitation depth versus the square root of the detection time using a pulsed source and time-gated detection system. Here $\mu_s' = 10\text{cm}^{-1}$, $\mu_a = 0.1\ \text{cm}^{-1}$, $c = 0.214\ \text{mm ps}^{-1}$, and the source-detector distance of the surface sensor is 10 mm. (Reprinted with permission from [12])

17.4.3 Interstitial Geometry

The advent of optical-based treatments such as photodynamic therapy (PDT) and laser thermal therapy (LTT), often applied interstitially, has increased the necessity for corresponding interstitial sensors that can provide information for pre-treatment planning, and for on-line monitoring of treatment progress. Interstitial sensors measure the internal fluence or radiance (see Chapter 2 of this text for a review of these terms) in a turbid medium to extract information on either the light (directional or integrated) intensity (dosimetry), optical properties (spectroscopy), or extent of treatment effect. The following section reviews the basic design of fluence and radiance probes, followed by clinical applications specific to the interstitial geometry.

17.4.3.1 Fluence Sensors

Spherically Diffusing Tip

Conventional interstitial fluence sensors are constructed by attaching a minimally-absorbing, highly scattering sphere to the end of a plane cut fiber (Fig. 17.17a). Since light entering the diffusing sphere is multiply scattered, the incident light is effectively collected over all polar and azimuthal angles by the plane cut fiber (typically 400–600 μm core diameter) with virtually equal probability over all polar and azimuthal angles. Monte Carlo simulations comparing the fluence with and without a sensor have demonstrated that spherical fluence sensors with diameters as small as 0.5–0.8 mm result in minimal perturbation to the true fluence field in the turbid medium [37].

Two approaches have been reported for constructing the scattering tip. The first employs a UV-curable epoxy polymer mixed with a highly scattering powder (typically titanium dioxide) [38]. A plane cut fiber emitting UV light at microwatt powers

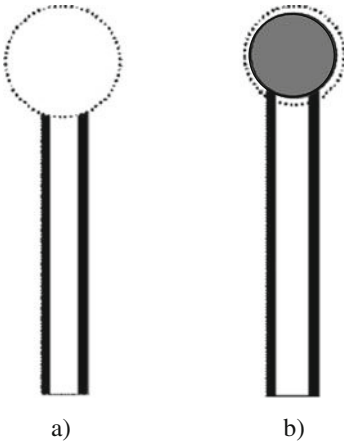


Fig. 17.17 (**a**) Sphere-tipped fluence probe with scattering sphere attached to the end of a plane cut fiber (**b**) Integrating sphere fluence probe with a constructed ball tipped fiber coated with a thin layer of scattering material

is shone into the epoxy mixture. The resulting fluence pattern is then an isotropic spherical distribution that forms symmetrically around the fiber tip. Following the formation of the scattering tip, the epoxy is irradiated at higher powers to fully solidify.

A variation of this geometry is to employ a fiber with a ball-shaped tip instead of a plane cut fiber and grow a thin layer of scattering material around the tip (Fig. 17.17b). The spherically tipped fiber can be purchased through commercial vendors, or constructed in-house by employing a small torch to melt the fiber tip at temperature >2500°C to the desired ball shape. Here the optical fiber tip acts as an integrating sphere and requires less scattering epoxy compared to a bare-tipped fiber to fully randomize the direction of the incoming light. An additional advantage of this configuration is that if the fluence sensor is employed as an isotropic spherical source, higher input powers can be utilized, since the absorbed heat is dissipated over the entire spherical surface instead of being concentrated in the vicinity of the smaller bare-tip area.

A second method is to attach a pre-fabricated sphere, typically made of Arnite, ceramic, or Teflon, to the end of a bare tip, plane cut fiber [39, 40]. A small hole is bored to the center midway point of the sphere, the sphere is then "skewered" by the fiber and fixed using glue to the end of the fiber tip.

The primary weakness of scattering sphere probes is that sizeable diameters (∼1–2 mm) are required to achieve a desired level of collection isotropy and that an inherent "blind-spot" is present at the stem region of the probe. Typically, a 3 mm-diameter probe can achieve an isotropy to within ± 15%, while a ∼1 mm-diameter probe is isotropic to within ∼± 20% over a 320 angular range [39, 40]. Finally, during clinical application, the spherical tip can become unattached from the fiber core if the sensor is pulled back too quickly or awkwardly during its positioning, manipulation, or removal.

Fluorescent Dye-Loaded Tip

To overcome the size limitations of scattering-sphere probes, Lilge et al. [41] developed fluorescent dye-loaded fluence sensors. The use of fluorescent dyes overcomes the limitations of larger scattering-based spherical sensors, because fluorescence emission is naturally isotropic. As such, the isotropy of the sensor is dependent only on the tip shape and not its size [41].

The sensor designs are illustrated in Fig. 17.18. First, a polymethlyl methacrylate (PMMA) dye-free offset is attached to the end of a plane cut fiber (Fig. 17.18a). The offset effectively serves as a spacer between the dye-sensitive region and the fiber. Next, a dye-loaded PMMA tip is added to the offset. The dye-free and dye-loaded PMMA are constructed using methyl-methacrylate (MMA), which is polymerized in an oven at 60 C for ∼48 h and initiated using a primer (Azobis, Polyscience, Milwaukee, Wi, USA). The addition of dimethyl chloride (DMC) changes the solid polymerized PMMA to a viscous form. As such, fluorescent dyes dissolved in DMC can be added to the solid PMMA. Following evaporation of DMC, the dye remains absorbed in the PMMA.

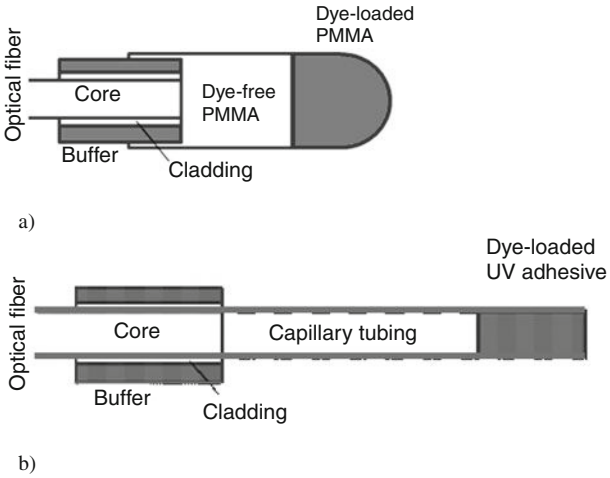


Fig. 17.18 Fluorescent dye-loaded fluence sensors: (**a**) Dye-loaded PMMA design; (**b**) Capillary tube tip design with dye-loaded UV adhesive (after Lilge et al. [41])

Fiber tips with total diameters ranging from ~250 to 435 μm can be constructed using this approach. Smaller tips (<200 μm diameter) can also be produced by attaching a section of dye-filled silica capillary tubing to the end of an optical fiber (Fig. 17.18b). The capillary tube is filled at the end with a dye-filled epoxy, which is cured internally inside the tubing.

The polar isotropy of the sensor is typically ± 10% between 0° and 160° for probe tips <400 μm, although angular variations as large as ± 20% have been noted (caused by difficulties in consistent tip construction). These values are comparable to the response of spherical fluence sensors. The primary disadvantage of dye-doped sensors is their limited lifetime that stems from photodegradation of the fluorescent dye. As such, fluorescent probes should not be considered for long term extended use, especially if employed for dosimetric purposes. Furthermore, the relatively low damage thresholds negates the usage of such probes in some scenarios (e.g., for high-powered (>400 J-cm^{-2}) laser thermal therapy monitoring applications).

Calibration

Fluence sensor calibration is required for: (1) light fluence dosimetry (for example during PDT) and (2) absolute optical property determination. The light collected by a fluence sensor is directed to a photodiode or photomultiplier tube and is converted to a photovoltage (V) that is proportional to the true fluence, ϕ, related to each other by:

$$\phi = A\,(V - B) \tag{17.40}$$

where A is the calibration constant (mW cm^{-2} V^{-1}), and B (volts) is a factor that accounts for the leakage of the detector. Typically, fluence sensors are calibrated in air either using a collimated light source [39] or in an integrating sphere [42]. When applied to tissue, additional calibration factors must be employed to correct for the index of refraction differences between air and tissue, fluence perturbation from the scattering tip, the blind spot from the fiber stem, and non-isotropic responses of the scattering bulb [40].

17.4.3.2 Pre-Clinical and Clinical Applications

Non-invasive surface reflectance geometries allow for a well-sampled data set of radially-resolved surface measurements, which allows for the determination of both μ_t' and μ_{eff}, and thus the unique determination of absorption and scattering. However, interstitial characterization is complicated by: (1) the necessity to limit patient invasiveness and (2) limitations in sensor positioning due to insertion through guidance templates. Clearly, the number of physical allowable measurement positions is limited when inserting sensors directly into tissue. As such, a particular challenge of interstitial geometries is to devise strategies for uniquely separating absorption and scattering properties using measurements far from the source, where optical properties cannot be uniquely separated with relative measurements alone.

Calibrated sensors effectively overcome the limitations of poor spatial sampling by measuring differences in intensity at positions far from the source that are dominated by $\mu_t' = \mu_a + \mu_s'$ and not μ_{eff}. However, accurate calibration can be difficult to accomplish and must be performed for each source-detector pair and for different irradiation wavelengths during spectroscopy. Several interstitial techniques have been devised to avoid absolute calibration when determining tissue optical properties at single wavelengths.

Dimofte et al. [42] have employed multiple fluence sensors and sources attached to a scanning translation stage coupled to a common catheter, for optical property determination during PDT of the prostate. The set-up allows for multiple spatial measurements to be acquired in the *same* catheter, thereby avoiding additional insertions. Optical properties can be assessed throughout the prostate volume in quadrants with different source-detector pairs sampling a different volume of the prostate. A similar approach, reported by Pomerlea-Delacourt and Lilge [43], is to employ fluorescent sensors with multiple detection locations along the same fiber.

Interstitial optical measurements have recently been developed for monitoring of thermal damage during laser thermal therapy (LTT) [5]. The approach relies on the significant difference in scattering between native and thermally coagulated tissue. The optical approach carries two significant advantages compared to conventional temperature sensors used for monitoring LTT. First, due to the speed of light in tissues, optical measurements provide volumetric information in near real-time even at positions distant from the source. This is in contrast to temperature sensors, where charring or coagulation events can occur some tens to hundreds of seconds before manifesting as a temperature change at typical monitoring positions of 0.5–1 cm from the source. Second, temperature sensors are unable to directly monitor thermal

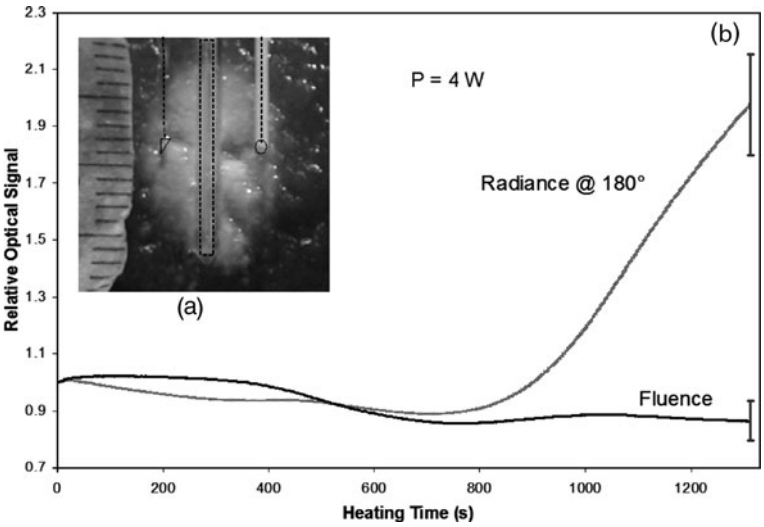


Fig. 17.19 (**a**) Optical monitoring of coagulation boundary using real-time optical fluence measurements (source power (*P*), source-detector separation (*r*)). As the coagulation boundary approaches the fluence sensor, a decrease in light intensity is observed. This is due to a "light trapping" effect from the increased scattering properties of coagulated tissue. The same effect causes an increase in light intensity when the coagulation boundary reaches the location of the fluence sensor (located approximately at the *triangle*). Hence, a fluence sensor can be positioned strategically at the boundaries of a tumour to terminate the LTT treatment. The inset is a photograph of the resulting region of damage following LTT, indicating the locations of the source and detectors employed for optical monitoring. Note that, as expected from the termination of the treatment due to the increase in fluence, the damage boundary is located almost at the location of the fluence sensor (**b**) Comparison of fluence and 180° radiance measurements for optical monitoring of the coagulation boundary in tissue simulating optical phantoms. The 180° radiance enhances the sensitivity to the passing of the coagulation boundary compared to fluence measurements. (Reprinted with permission from [5])

damage and instead rely on thermal dose models or threshold temperatures to estimate the extent of coagulation. However, optical sensors are able to directly measure coagulation changes that are independent of the thermal dose models.

Figure 17.19a illustrates the concept of optical monitoring during LTT of ex vivo bovine liver. As the coagulation boundary approaches the fluence sensor location, a decrease in light intensity is observed. This is due to a "light trapping" effect from the significant increase in optical scattering within the coagulated region. This same scattering increase results in a significant increase in light intensity when the coagulation boundary reaches the sensor location. Hence, a fluence sensor strategically positioned at the margins of a cancerous region could be used to terminate an LTT treatment and spare healthy surrounding tissue. In addition, directional light measurements (radiance) measurements can be employed instead of (isotropic) fluence sensors to offer enhanced sensitivity to the passing of the coagulation boundary. In particular, Chin et al. [5] demonstrated in tissue simulating phantoms that the backwards (or 180° radiance) provides almost a two-fold increase in sensitivity

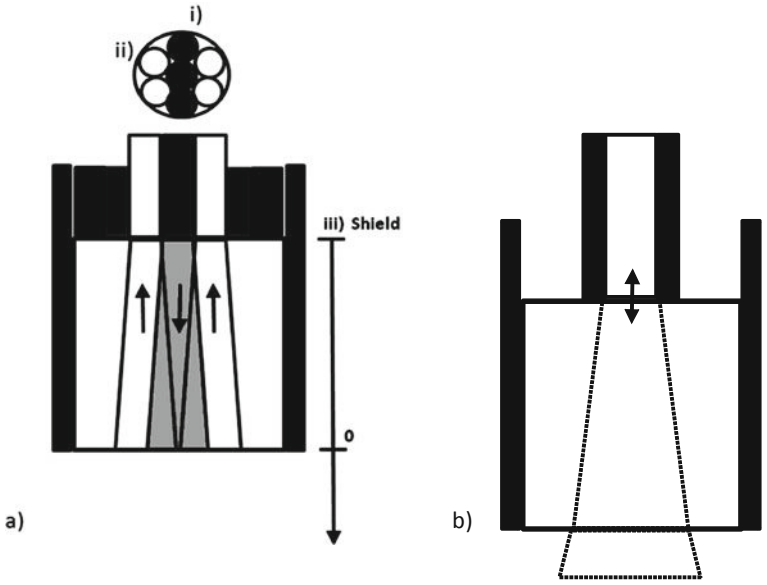


Fig. 17.20 Conventional fluorescent sensors. (**a**) Multi-fiber design (i) excitation fibers, (ii) collection fibers and (iii) quartz shield (**b**) Single fiber design: variable aperture (after Utzinger and Richards-Kortum [2])

compared to fluence measurements (Fig. 17.20b). They demonstrated that the light intensity increase from the passing of the coagulation boundary is primarily due to the increase of backwards traveling light randomized by the increased scattering properties of the coagulated region. Hence unlike fluence measurements, 180° radiance sensors effectively "filter out" the radiance that is insensitive to the passing of the coagulation volume.

The optical approach provides virtually instantaneous information regarding the onset, approach and final extent of thermal damage. Combined with interstitial temperature sensors, optical monitoring can be employed as a feedback technique to adjust laser power levels and determine treatment termination when complete coverage of the target volume is achieved. As such, optical monitoring provides the advantages of volumetric information of thermal damage with the simplicity point optical sensor technology.

17.5 Fluorescence Spectroscopy Fibers

17.5.1 Basic Approach

Fluorescence spectroscopy is widely used for the detection and evaluation of neoplastic tissues. The technique is potentially able to provide quantitative information regarding the biochemistry, morphology and composition of pre-cancerous lesions

in several tissues including cervix, esophagus, bronchus, colon and skin. The primary fluorescent tissue components, or fluorophores, are structural proteins such as pyridine nucleotide, collagen, tryptophan, porphyrins, elastin, and flavoprotein [44]. The presence and relative concentration of these fluorophores and their distribution in tissue provides a means of distinguishing between normal and cancerous tissues. In addition, the technique is able to provide dosimetric information and treatment evaluation of photonic treatments such as PDT and LTT. In the case of PDT, fluorescent probes can be employed to determine the localization and concentration of drugs (exogenous photosensitizers) necessary for treatment effect, as well as photobleaching for treatment monitoring.

Given the range of target sites evaluated by fluorescence spectroscopy, it is important to design fiber sensors that provide customized capabilities for their specific clinical applications. In the following section we describe a typical fluorescent sensor and the specific design characteristics that influence the detected signal.

17.5.2 Fluorescence Sensor Design

Similar to elastically-scattered light sensors of section 4, a conventional fluorescent sensor consists of (a) an excitation fiber, (b) a collection fiber and (c) a quartz shield (employed as a spacer). Carbon-filled or low-fluorescent epoxy can be employed to hold the fibers. Two potential configurations exist. In the first, the excitation and collection fibers can be comprised of separate fibers (Fig. 17.20a). Alternatively, a single excitation-collection fiber (Fig. 17.20b) can be employed for a more localized fluorescence measurement. Multi-fiber sensors employ fiber bundles at different detection locations to achieve an overall larger interrogation volume. This configuration is particularly attractive for characterizing larger tissue, or when the margins or location of a suspected lesion are unknown.

17.5.2.1 Single and Multi-Fiber Sensors: Pros and Cons

It is important to compare the relative merits of single fiber and multi-fiber (i.e. multi-distance) fluorescent sensors. Both simulation and phantom experiments have demonstrated that single fiber probes provide improved signal intensity and greater sensitivity to superficial fluorophores. This is primarily due to the overall shorter pathlengths traveled by collected photons in the single-fiber geometry. In contrast, multi-fiber separation probes detect photons that travel longer pathlengths and, as such, are more sensitive to distortions and signal loss from blood absorption. For shorter separations, multi-distance sensors can still provide reasonable depth sensitivity with peak fluorescent detection sensitivity of ~0.1 mm depth per 0.1 mm separation, up to 0.4 mm. Beyond ~0.4 mm larger separations result in a virtually uniform sensitivity to all depths.

Zhu et al. [45] described a 7-fiber sensor that combines the benefits of both single and multi-distance designs. In their configuration a central fiber is surrounded in a ring by six additional fibers. The central fiber acts as a single fiber (illumination

and collection) sensor providing superficial probing sensitivity. The central fiber and one of the surrounding fibers can also be utilized in parallel for illumination and collection as a larger diameter 'single fiber' sensor to provide intermediate depth sensitivity. Finally, the central fiber and one of the surrounding fibers can be employed as a multi-fiber sensor (one fiber illuminating and the other fiber detecting) as a means to provide information on deeper-seated structures.

17.5.2.2 Effect of Spacer-Thickness

The shield/spacer thickness, typically 1–7 mm, determines the fiber(s)-tissue separation and is a vital sensor design parameter. In the absence of a spacer (i.e., zero fiber-tissue separation, contact sensor) the excitation and detection regions are not co-localized. Therefore, light must travel laterally through the interrogated tissue before being collected, thereby increasing the overall path-length and decreasing the intensity of detected photons due to attenuation. However, by increasing the fiber(s)-tissue distance, the spacer allows the excitation and collection regions to intersect, causing a decrease in the average lateral distance required for fluorescence collection, and correspondingly, an increased collection efficiency and improved sensitivity to fluorophores located in superficial regions. Note that, as described in Section 17.4.2.1, by adjusting the source-detector separation, the mean probing depth of detected photons can also be adjusted for steady-state reflectance sensors.

The collection efficiency, β_t, for a multi-fiber sensor (depicted in Fig. 17.20a) has been described by [46]:

$$\beta_t = \frac{\beta_0}{(z + \text{shield}_t)^2} \tag{17.41}$$

where z is the sensor to tissue distance along the optical axis of the sensor, shield_t is the length of the spacer cap, and β_0 is a constant that incorporates the detector efficiency. Equation (17.41) assumes isotropic fluorescence emission and does not hold for tissues with high forward scattering. For example, in arterial tissue, the collection efficiency has been shown empirically to follow a $1/\text{shield}^n$ dependence, with $n \sim 1.1$ [47].

Equation (17.42) breaks down for very small spacer thicknesses < 0.5 mm [48]. This is due to a shift in the overlap of the excitation and collection volumes of the sensor. When no spacer is employed, the collection fibers do not directly view the region of excitation. Here, photons must travel laterally (hence, be attenuated) before being detected. However, as the spacer size increases slightly, the overall field of view of the collection fibers increases such that the collection fibers directly view the excitation spot. In this case, no lateral propagation is needed for detection and the sensor preferentially detects superficial fluorescence. In this spatial window, an increase in β_t is observed. Beyond this regime, the increasing overlap of the excitation-collection volumes now detects a greater proportion of deeper lateral-traveling photons with longer path-lengths, thereby decreasing the collection

efficiency. Equation (17.42) is representative of this latter regime. Single-fiber sensors, which obviously have the greatest volume of illumination-collection overlap also exhibit such trends.

It is important to note that the larger the shield thickness, the larger the diameter of the shield and, consequently, the overall sensor diameter. To reduce the overall diameter, the quartz spacer may be substituted using a coated glass rod or thick portion of optical fiber. These elements can effectively mix the illumination and collection volumes of the sensor. With this design, illumination and collection fibers are stripped of their outer jacket, randomly fixed with epoxy and packed hexagonally into a thin glass rod held together using shrink tubing. This greatly reduces the overall diameter of the resulting sensor, while maintaining the high collection efficiency of larger diameter spacer-based sensors. To ensure uniform illumination, the length of the rod should be greater than $2R_{\text{fiber}}/NA$, where R_{fiber} is the radius of the illumination and collection fibers.

17.5.2.3 Fiber Diameter and Orientation

Similar to reflectance spectroscopy sensors (Section 17.4.2.1), varying both the source and collection diameter has shown to influence the depth sensitivity of collected fluorescence. Monte Carlo models and phantom studies have been performed to elucidate the influence of sensor design on both the intensity and primary origin of the detected fluorescence signals [45, 48].

For small diameter fibers, sensitivity to superficial depths can be observed. This is a result of the reduced number of possible path-lengths of fluorescence photons for detection. However, with larger diameters the fluorescence sensitivity is averaged over many depths, thereby homogenizing and increasing the overall signal. These changes are due to the same physical effects that occur with varying source-detector separation: an increase in measured photons that travel longer and deeper paths through tissue before being collected.

Previous studies have demonstrated that, in the range of 0.22–0.4, increasing the collection fiber numerical aperture results in an increase in collected intensity and an enhancement in fluorescence sensitivity to superficial layers. These trends are thought to stem from the increased acceptance angle afforded by larger NAs that allow for the collection of higher angle photons that originate under the illumination region. However, beyond an NA of 0.4, the collection efficiency ceases to increase significantly due to the longer path-lengths of high angle photons, resulting in significant photon attenuation [48].

Orienting the excitation fiber to greater than 45° incidence has shown to enhance collection efficiency and increase sensitivity to superficial layers [45]. This is because angled illumination results in a shallower excitation light distribution, which leads to a shallower volume of origin for the generated fluorescence. Recall that a similar principle can be employed for depth discrimination for spectroscopic reflectance sensors (Section “Specialized Fiber Optic Sensor Geometries”). However, for illumination angles less than ∼30°, very little difference in detected fluorescence characteristics is observed compared to normal illumination.

17.5.2.4 Effect of Illumination-Collection Fiber Separation

Compared to single-fiber sensor, multi-fiber sensors also offer the additional flexibility of varying the illumination-collection fiber separation. This provides an additional parameter for depth selectivity [48]. In general, increasing the illumination-collection spacing results in the detection of fluorescent photons from deeper regions. Again, physically this is explained by the longer path-lengths required for collection when larger fiber separations are employed.

17.5.3 Influence of Tissue Optics

While sensor design parameters influence the depth of fluorescence origin and average path length of detected photons, the detected fluorescent intensity and measured spectrum is also "distorted" by optical properties of the interrogated medium. A number of investigators have examined the effect of absorption and scattering on the "intrinsic" fluorescence spectrum [48–50]. These studies have shown that spectral alterations due to optical properties tend to be enhanced when detected photons travel longer path lengths within the tissue. Hence, probe designs with large diameters, large *NA*, shallow illumination angle (i.e., near normal incidence), small source-tissue spacing, and large source-detector spacing will suffer greater changes than their opposite counterparts.

While sensor design can be utilized to minimize the effects of tissue optical properties, an alternative approach is to correct the altered spectra through reflectance data that provide an indication of the degree of fluorescent distortion by combining reflectance-and fluorescence spectroscopy. The most straightforward approach is to utilize the measured reflectance to quantitatively derive information regarding the tissue optical properties (penetration depth, absorption and scattering). Using the measured optical properties, the "intrinsic" fluorescence spectra can be corrected using appropriate photon migration models. While early corrections required simplified geometries such plane-wave irradiation and/or specific source-detector arrangements [51–53], recent improvements in diffuse light models have extended the technique to more complicated three-dimensional geometries [55, 56].

A more simplified approach is to employ an approximate relationship between the measured fluorescence and reflectance to correct for spectral distortions. Finlay et al. [54] recently described an empirical correction method by dividing the measured fluorescence spectrum by the reflectance signal measured at the same wavelengths. The approach, which does not require determination of tissue optical properties, is applicable over a wide range of reflectance measurements typically encountered in vivo. However, their correction is limited to fixed source-detector geometries.

Another alternative is to employ the distorted fluorescence spectrum to directly obtain information regarding the sample's optical properties, and then employ the determined optical properties to eliminate the distortion. Using a forward-adjoint model, Finlay and Foster [56] have demonstrated that it is, in fact, possible to

directly recover the intrinsic fluorescence and information on tissue optical properties from the measured distorted spectrum. Future advancement of this concept will likely allow for fluorescent corrections without the necessity of added reflectance components, thereby greatly simplifying detection technology and computational analysis.

17.6 Optical Coherence Tomography Fibers

Unlike other fiber sensors described in this chapter, OCT forms truly depth-resolved tomographic cross-sectional images, with a depth resolution in the 1–20 μm range (cf. "depth-sensing" discussions of various sensors above). Essentially, an OCT imaging sensor forms the sample arm of an OCT interferometer and typically consists of a single mode optical fiber terminated with a focusing element (e.g., a GRIN or a ball lens) and beam deflecting element (e.g., a retro-reflecting prism or an angle cleaved distal optic) for a side-viewing sensor. The overall principle can be appreciated from the layout of a catheter-based fast OCT imaging system presented by Tearney et al. [57], as shown in Fig. 17.21. Specifics of the distal sensor design, combination of radial scanning and longitudinal fiber pull-back approach, and integration with the rest of the OCT imager can be enabled. Using the system in Fig. 17.21 and its later variants, the authors were able to generate impressive subsurface images of the superficial layers of the gastro-intestinal and broncho-pulmanory tracts, as well as intravascular images of atherosclerotic plaques in coronary arteries [58]. Several other groups have also developed their own versions of intravascular and endoscopic OCT imaging systems, all using their own unique implementations of OCT fiber sensors. For example, Fig. 17.22 shows an endoscopic sensor developed by Yang et al. [59] for esophageal imaging in the gastrointestinal tract.

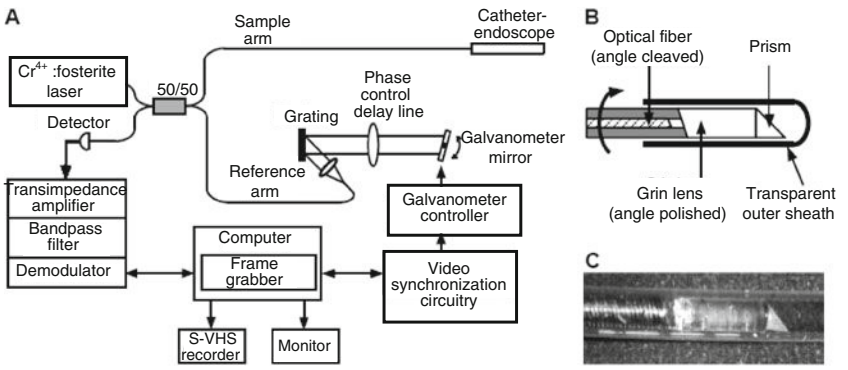


Fig. 17.21 (**a**) layout of the OCT system. (**b**) Second-generation OCT catheter-endoscope. By adjusting the angle cleaving the optical fiber and the angle polishing the GRIN lens, normal incidences are avoided and thus internal reflections can be minimized. (**c**) Photograph of an OCT catheter-endoscope at the distal end of the sensor. (Reprinted with permission from [57])

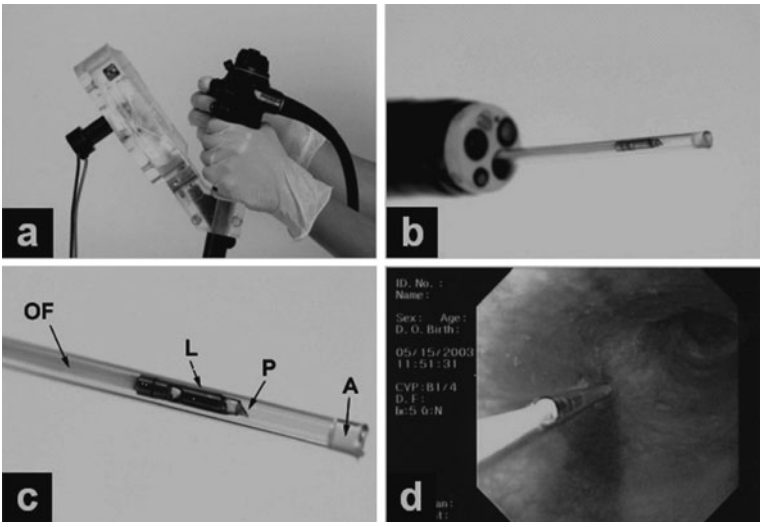


Fig. 17.22 An OCT optical fiber sensor adapted for biomedical imaging during human gastrointestinal endoscopy. (**a**) the proximal end of the endoscope, showing the linear translation motor that pulls the OCT fiber to-fro by ~3 mm to enable longitudinal cross-sectional imaging (**b**) the distal end of the endoscope, with the OCT probe protruding from the instrument channel (**c**) close-up of the OCT fiber sensor within its transparent sheath, with OF = optical fiber, L = focusing (GRIN) lens, P = side-reflecting prism, and A = adhesive sealant (**d**) clinical imaging showing OCT probe in contact with human esophageal tissue. (Reprinted with permission from [59] and [74])

A particularly important design consideration for OCT sensors is the avoidance of normal interfaces and retro-reflective surfaces, which in addition to reducing signal strength by causing photon loss, can inadvertently interfere with the reference beam and generate spurious and artificial lines or streaks in the resultant image. Thus, care must be taken to minimize the number of elements in an OCT sensor, as each one will likely contribute at least one (and often two) normal interfaces whose reflections could fall within the coherence ranging depth of the system and thus cause artifacts in the image. The use of angle-cleaved surfaces to avoid normal incidence, refractive-index matching compounds, and careful optical design to minimize the number of interfaces are thus essential for successful operation of an OCT fiber sensor. Figures 17.21(b, c) and 17.22 illustrate these concepts.

A significant limitation of OCT is its shallow imaging depth, which is typically 1–3 mm in nontransparent mammalian tissues. While a huge variety of exciting biomedical applications are still possible within this depth range (for example, endoscopic and intravascular imaging as described above, where the pathologic changes of interest occur within 1–2 mm of the inner surface), many scenarios would benefit from deep tissue imaging with OCT. Fortunately, owing to the technique's intrinsic compatibility with optical fiber technology, this can indeed be done in the context of an *interstitial* OCT fiber sensor.

Li and co-workers have first demonstrated an "OCT needle" approach suitable for deep tissue imaging in 2006 [61]. Using a single mode optical fiber with a GRIN lens and microprism to focus and deflect the beam, delivered to the desired deep tissue location via a 27-gauge needle (0.41 mm outer diameter), the authors were able to obtain high-quality microstructural images in deep-seated hamster leg muscle tissue. More recently, the Toronto OCT group [60–62] has developed interstitial probes capable of both microstructural OCT and microvascular Doppler OCT imaging in deep tissues. Using both cleaved ball lens and cleaved GRIN lens designs to minimize stray reflections as shown schematically in Fig. 17.23 and utilizing minimum footprint components, sub-mm diameter probes have been produced. X-ray or ultrasound guidance enabled accurate placement of the distal imaging end of the probe at the desired 3D location within, or at the margins, of the deep-seated tumour to be treated. While the tissue microstructure did not show obvious alterations in response to treatment (photodynamic therapy (PDT) in this case), the Doppler OCT signal changed drastically, indicating a significant treatment-induced microvascular response that can be effectively monitored both on OCT optical fiber sensors [62].

17.7 Raman Spectroscopy Fibers

17.7.1 Basic Approach

Raman spectroscopy has proven to be a valuable tool for interrogating biological materials. This technique relies on an inelastic scattering process in which incident photons transfer energy to (Stokes), or gain energy from (anti-Stokes) molecules within a sample. For biological applications, the Raman scattering accounts for approximately 10^{-10} of the incident light and the Raman signal is typically several orders of magnitude weaker than the fluorescence signal. Raman spectroscopic techniques for biomedical applications have seen significant advancement over the past few decades, with improvements in portability of equipment and fiber optic based excitation and collection, which can be used in vivo. Raman spectroscopy has been shown to provide information about the structure of molecular constituents in tissues [63, 64], and has the ability to distinguish between normal and cancerous tissues in some scenarios [65, 66].

17.7.2 Raman Probe Design

The two key parameters for designing a fiberoptic probe for Raman spectroscopy are maximizing the Raman-scattered light collection and minimizing unwanted optical signals (elastically scattered light, fluorescence photons, fiber-generated Raman signals, etc.). The intrinsically weak Raman signals can be masked by strong fluorescence from biomolecules and cells, particularly for excitation in the visible region. Fluorescence decreases rapidly at longer wavelengths, such that this problem

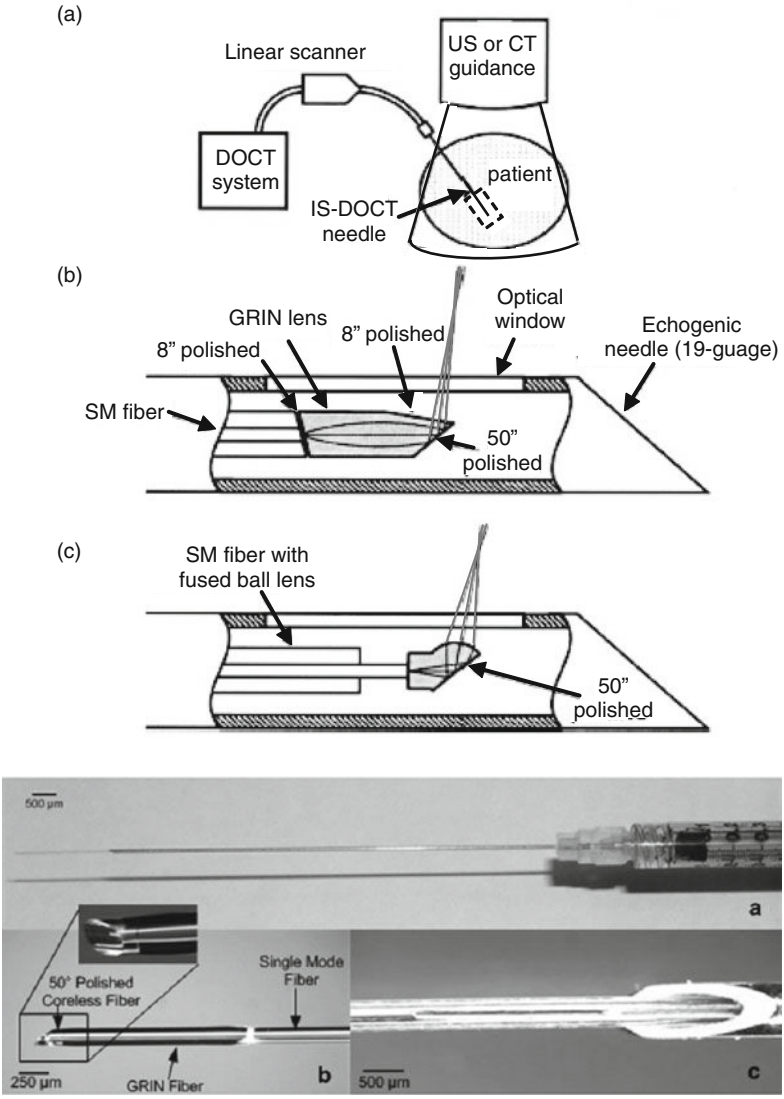


Fig. 17.23 (**a**) Schematic demonstrating interstitial Doppler OCT guided by either ultrasound or computed tomography (**b**) A beam is focused into a single mode fiber and subsequently deflected via a GRIN lens with three polished surfaces. Note the avoidance of normal interfaces in this sensor to minimize unwanted interference effects (**c**) Simplified schematic with a polished ball-lensed fiber. Both designs use an echogenic needle for enhanced ultrasound visualization [60]. The bottom photograph shows (**d**) a 2nd generation Doppler OCT needle probe used for interstitial monitoring of photodynamic therapy treatments [61]

has been largely overcome with excitation in the NIR. Alternatively, fluorescence-free Raman can be achieved with excitation in the UV below ~270 nm, but at the expense of reduced optical penetration (microns). For ex vivo tissue interrogations, integration times of a few seconds yield acceptable signal to noise ratio. For in vivo deployment longer integration times 20–60 s are typically required; however, the acquisition of clinical data in real-time (~1 s integration time) with good signal-to-noise ratios has been demonstrated for margin assessment in patients undergoing partial mastectomy [67]. Fiberoptic bundles and imaging optics can also be used to enhance collection efficiency (signal to noise), and filtering (and/or software correction) can be used to further remove unwanted signals.

Some materials used in Raman systems are themselves Raman active and hence can interfere with the signal coming from the sample under investigation. Hence, filters are required in both the delivery and collection arms of a Raman system. Fused silica fibers used to transmit optical radiation can generate Raman signals, with the intensity depending on the excitation wavelength and wavenumber region investigated. In the fingerprint region (up to ~2000 cm^{-1}), fused silica fibers are characterized by a strong background signal generation. These fiber-induced signals are proportional to the fiber length and can be greater than those from the sample. For high wavenumber investigations (2400 to 3800 cm^{-1}), fused silica has almost no Raman background signal. Santos et al. [68] analyzed several core, cladding and coating materials used in Raman fiber probes for investigations in the high wavenumber region. Overall, low-OH (hydroxyl) fused silica appears to the optimal fiber material for NIR Raman applications.

One must also prevent surface reflected or elastically scattered excitation light from entering the collection fibers. This is typically achieved using a long pass filter (or notched filter) positioned in front of the collection optics. Typically a band pass filter is also used in the excitation line to allow for the transmission of only the excitation light.

A conventional Raman probe consists of a single excitation fiber and a bundle of collection fibers. Short et al. [69] developed a design with a central 200 μm excitation fiber surrounded by twenty seven 100-micron core diameter collection fibers (Fig. 17.24a). The excitation fiber can be coated, for example with gold [69] or aluminum [67], for optical isolation to prevent cross talk between the excitation and collection optics.

Haka et al. [67] developed a fiber bundle comprised of a 200-μm core diameter low OH fused silica excitation fiber (0.22 NA) surrounded by 15 collection fibers (0.27 NA) (Fig. 17.24b). The distal end of the probe contains a 0.55 mm short-pass filter rod, proximal to the excitation fiber, surrounded by a long-pass filter module positioned in front of the collection fibers. Filters are attached to the fibers using an index-matching optical cement. A 2 mm diameter sapphire ball lens (1.77 NA) is used to improve collimation of the excitation light and the collection of the Raman scattered light. The excitation fiber–lens separation determines the spot size and hence the energy density incident on the tissue. Due to the large angular distribution of the remitted Raman light, the ball lens should have a large index of refraction to improve coupling to the collection fibers.

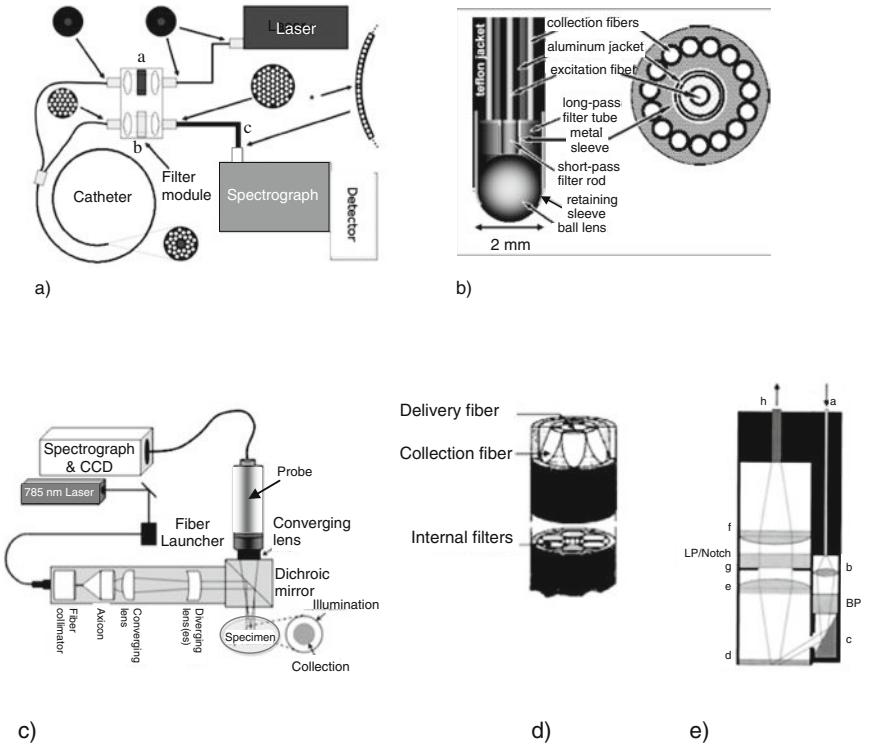


Fig. 17.24 (**a**) Endoscopic Raman probe based on a single illumination fiber surrounded by 27 collection fibers. Band-pass filters (label **a**) are used for the excitation fiber and long-pass filters (label **b**) for the collection fibers. A second fiber bundle consisting of 54 ultralow 100 micron core fibers is coupled to the spectrograph (label **c**) (with permission from [69]). (**b**) Raman probe based on a single illumination fiber surrounded by 15 collection fibers. The excitation fiber is coated with aluminum to prevent crosstalk between the excitation and collection lines. A short pass filter is used to reject background Raman signals generated in the low-OH fused silica and a long pass used on the collection line. A sapphire ball lens is used to collimate the excitation light and improve coupling of Raman light into the collection fibers (with permission from [67]). (**c**) Light is coupled into a 200 micron core fiber, collimated, then directed through an axion. A Galilean telescope design comprised of converging and diverging lenses is used to achieve variable ring diameters for illumination. A dichroic mirror was used to reflect excitation light to the sample and transmit collection light to a 1 mm diameter bundle of fifty 62.5 micron core fibers (with permission from [70]). (**d**) Endoscopic Raman probe comprised on a single central delivery fiber surrounded by seven beveled 300 micron core collection fibers. Band-pass and long-pass filters are used for the excitation and collection lines, respectively (with permission from [71]). (**e**) Raman probe based on excitation light, through a single 200 micron core fiber (label **a**), imaged on the tissue sample via a biconvex lens (label **b**) and mirror (label **c**). A band-pass (BP) filter is used to remove Raman background. The probe terminates in a quartz window (label **d**). Collected light is imaged using two plano convex lenses (labels **e** and **f**) through an aperature stop (label **g**) and notched filter onto a fiber bundle (label h) (with permission from [72])

Single point illumination can limit the amount of laser power that can be delivered due to the potential of inducing thermal damage. To overcome this, Schulmerich et al. [70] collimated light from a 200 μm core fiber directed through a 175° Axicon lens and focused using a positive-negative lens pair to create an illumination ring. Ring diameters of 3.0–14.5 mm were achieved. In this design, light is collected using a bundle of fifty 62.5-μm core diameter fibers (Fig. 17.24c).

Molckovsky et al. [71] used a fiber probe with interchangeable beveled tips (Fig. 17.24d). They demonstrated that beveled tips improved collection efficiency. Moreover, the depth of interrogation is determined by the bevel angle and separation between the central excitation fiber and surrounding seven 300-μm core diameter collection fibers. As the bevel angle is increased and separation between excitation and collection fibers is decreased, the interrogated volume moves closer to the surface. Excellent collection efficiency has been reported using this probe.

Mahadevan-Jansen et al. designed a 12 mm diameter Raman probe for cervical study (Fig. 17.24e) [72]. Light from a diode laser is coupled to a 200-μm core diameter fiber and deflected onto the tissue surface using a gold mirror. The angle of deflection is chosen such that the excitation spot overlaps with the collection region. Remitted light from the sample is collected using two biconvex lenses and imaged onto a fiber bundle consisting of fifty 100 μm core diameter fibers. The probe terminates in a quartz window, which is in contact with the tissue sample. The excitation light will induce both Raman and fluorescence photons in the quartz window, but since both processes are well known for quartz, this unwanted background signal can be identified and removed (in software). More recently, Robichaux-Viehoever et al. [73] utilized beveled collection fibers around a single excitation to achieve overlap of the excitation and collection regions.

For medical applications, Raman fiber probe design must also address a number of clinical considerations. Practical probe dimensions are typically up to 2 mm (diameter) to allow for insertion into needles and endoscopes, and fiber lengths can be on the order of a few meters.

17.8 Summary

Optical fiber sensors offer unique capabilities for accurate assessment of biological tissues, with interesting applications towards early disease diagnosis, therapy guidance and feedback, and treatment response monitoring. Depending on a particular clinical application and intent, they may be advantageous due to their bio-compatibility, freedom from electromagnetic interference, choice of technological design parameters that may be individually optimized, and ability to furnish a variety of useful tissue characterization information that may be difficult/impossible to obtain by other means. The overview presented in this chapter has discussed basic operating principles of fiber sensors, reviewed the types of fibers currently employed in biomedicine, and provided illustrative examples of their biomedical

use. It is hoped that the presented concepts and applications of optical fiber sensors convey the scope of the basic science, enabling technology, and useful approaches and innovations that are occurring in this exciting field.

Acknowledgement The authors would like to thank Dr. Richard Schwarz from the Richards-Kortum lab at The University of Texas at Austin for providing the derivation of the ball-lens equation.

References

1. Kapany NS. *Fiber optics. Principles and applications*. Academic, New York (1967).
2. Utzinger U and Richards-Kortum RR. Fiber optic probes for biomedical optical spectroscopy. *J. Biomed. Opt.*, 8(1):121–147 (2003).
3. Verdaasdonk RM and Borst C. Optics of fiber and fiber probes. In: AJ Welch and MJC van Gemert (eds) *Optical-thermal response of laser-irradiated tissue*. Plenum, New York, pp. 619–666 (1995).
4. Polymicro Technologies Catalog, http://www.polymicro.com/catalog/2_8.htm (2009).
5. Chin LC, Wilson BC, Whelan WM, and Vitkin IA. Radiance-based monitoring of the extent of tissue coagulation during laser interstitial thermal therapy. *Opt. Lett.*, 29(9):959–961 (2004).
6. Dickey DJ, Moore RB, Rayner DC, and Tulip J. Light dosimetry using the P3 approximation. *Phys. Med. Biol.*, 46(9):2359–2370 (2001).
7. Farrell TJ, Patterson MS, and Wilson BC. A diffusion theory model of spatially, resolved, steady-state diffuse reflectance for the non-invasive determination of tissue optical properties in vivo. *Med. Phys.*, 19:879–888 (1992).
8. Hull EL, Nichols MG, and Foster TH. Quantitative broadband near-infrared spectroscopy of tissue-simulating phantoms containing erythrocytes. *Phys. Med. Biol.*, 43(11):3381–3404 (1998).
9. Wang L and Jacques SL. Use of a laser beam with an oblique angle of incidence to measure the reduced scattering coefficient of a medium. *Appl. Opt.*, 34(13):2362–2366 (1995).
10. Nichols MG, Hull EL, and Foster TH. Design and testing of a white-light, steady-state reflectance spectrometer for determination of optical properties of highly scattering systems. *Appl. Opt.*, 36(1):93–104 (1997).
11. Fuchs H, Utzinger U, Zuluaga F, Gillenwater R, Jacob R, Kemp B, and Richards-Kortum R. Combined fluorescence and reflectance spectroscopy: in vivo assessment of oral cavity epithelial neoplasia. *Porc. CLEO*, 6:306–307 (1998).
12. Patterson MS, Andersson-Engels S, Wilson BC, and Osei EK. Absorption spectroscopy in tissue-simulating materials: a theoretical and experimental study of photon paths. *Appl. Opt.*, 34(1):22–30 (1995).
13. Farrell TJ, Patterson MS, and Wilson B. A diffusion theory model of spatially resolved, steady-state diffuse reflectance for the noninvasive determination of tissue optical properties in vivo. *Med. Phys.*, 19(4):879–888 (1992).
14. Finlay JC and Foster TH. Hemoglobin oxygen saturations in phantoms and in vivo from measurements of steady-state diffuse reflectance at a single, short source-detector separation. *Med. Phys.*, 31(7):1949–1959 (2004).
15. Seo I, Hayakawa CK, and Venugopalan V. Radiative transport in the delta-P1 approximation for semi-infinite turbid media. *Med. Phys.*, 35(2):681–693 (2008).
16. Kienle A and Patterson MS. Determination of the optical properties of turbid media from a single Monte Carlo simulation. *Phys. Med. Biol.*, 41(10):2221–2227 (1996).
17. Kim AD, Hayakawa C, and Venugopalan V. Estimating optical properties in layered tissues by use of the Born approximation of the radiative transport equation. *Opt. Lett.*, 31(8):1088–1090 (2006).

18. Pfefer TJ, Matchette LS, Bennett CL, Gall JA, Wilke JN, Durkin AJ, and Ediger MN. Reflectance-based determination of optical properties in highly attenuating tissue. *J. Biomed. Opt*., 8(2):206–215 (2003).
19. Alexandrakis G, Farrell TJ, and Patterson MS. Monte carlo diffusion hybrid model for photon migration in a two-layer turbid medium in the frequency domain. *Appl. Opt.*, 39(13):2235–2244 (2000).
20. Seo I, You JS, Hayakawa CK, and Venugopalan V. Perturbation and differential Monte Carlo methods for measurement of optical properties in a layered epithelial tissue model. *J. Biomed. Opt.*, 12(1):014030 (2007).
21. Mourant JR, Bigio IJ, Jack DA, Johnson TM, and Miller HD. Measuring absorption coefficients in small volumes of highly scattering media: Source-detector separations for which path lengths do not depend on scattering properties. *Appl. Opt.*, 36(22):5655–5661 (1997).
22. Kumar G and Schmitt JM. Optimal probe geometry for near-infrared spectroscopy of biological tissue. *Appl. Opt.*, 36(10):2286–2293 (1997).
23. Amelink A, Kaspers OP, Sterenborg HJ, van der Wal JE, Roodenburg JL, and Witjes MJ. Non-invasive measurement of the morphology and physiology of oral mucosa by use of optical spectroscopy. *Oral. Oncol.*, 44(1):65–71 (2008).
24. Van de Hulst HC. *Light scattering by small particles.* Dover, New York (1957).
25. van Veen RL, Amelink A, Menke-Pluymers M, van der Pol C, and Sterenborg HJ. Optical biopsy of breast tissue using differential path-length spectroscopy. *Phys. Med. Biol.*, 50(11):2573–2581 (2005).
26. Amelink A, OP Kaspers, HJCM Sterenborg, JE van der Wal, JLN Roodenburg, and MJH Witjes. Non-invasive measurement of the morphology and physiology of oral mucosa by use of optical spectroscopy. *Oral Oncol.*, 44:65–71 (2008).
27. Kruijt B, de Bruijn HS, van der Ploeg-van den Heuvel A, de Bruin RW, Sterenborg HJ, Amelink A, and Robinson DJ. Monitoring ALA-induced PpIX photodynamic therapy in the rat esophagus using fluorescence and reflectance spectroscopy. *Photochem. Photobiol.*, 6:1515–1527 (2008).
28. Wang AMJ, Bender JE, Pfefer J, Utzinger U, and Drezek RA. Depth-sensitive reflectance measurements using obliquely oriented fiber probes. *J. Biomed. Opt*., 10(4):044017 (2005).
29. Arifler D, Schwarz RA, Chang SK, and Richards-Kortum R. Reflectance spectroscopy for diagnosis of epithelial precancer: model-based analysis of fiber-optic probe designs to resolve spectral information from epithelium and stroma. *Appl. Opt.*, 44(20):4291–4305 (2005).
30. Nieman L, Myakov A, Aaron J, and Sokolov K. Optical sectioning using a fiber probe with an angled illumination-collection geometry: evaluation in engineered tissue phantom. *Appl. Opt*., 43:1308–1319 (2004).
31. Skala M, Palmer G, Zhu C, Liu Q, Vrotsos K, Marshek-Stone C, Gendron-Fitzpatrick A, and Ramanujam N. Investigation of the fiber-optic probe designs for optical spectroscopic diagnosis of epithelial pre-cancers. *Lasers Surg. Med*., 34:25–38 (2004).
32. Schwarz RA, Arifler D, Chang SK, Pavlova I, Hussain IA, Mack V, Knight B, Richards-Kortum R, and Gillenwater AM. Ball lens coupled fiber-optic probe for depth-resolved spectroscopy of epithelial tissue. *Opt. Lett.*, 30(10):1159–1161 (2005).
33. Schwarz RA, Gao W, Daye D, Williams MD, Richards-Kortum R, and Gillenwater AM. Autofluorescence and diffuse reflectance spectroscopy of oral epithelial tissue using a depth-sensitive fiber-optic probe. *Appl. Opt.*, 47(6):825–834 (2008).
34. Johnson T and Mourant J. Polarized wavelength-dependent measurements of turbid media. *Opt. Express*, 4(6):200–216 (1999).
35. Myakov A, Nieman L, Wicky L, Utzinger U, Richards-Kortum R, and Sokolov K. Fiber optic probe for polarized reflectance spectroscopy in vivo: design and performance. *J. Biomed. Opt*., 7(3):388–397 (2002).
36. Groner W, Winkelman JW, Harris AG, Ince C, Bouma GJ, Messmer K, and Nadeau RG. Orthogonal polarization spectral imaging: a new method for study of the microcirculation. *Nat. Med*., 5(10):1209–1213 (1999).

37. De Jode ML. Monte Carlo simulations of the use of isotropic light dosimetry probes to monitor energy fluence in biological tissues. *Phys. Med. Biol.*, 44:3207–3237 (1999).
38. Henderson B. *An isotropic dosimetry probe for monitoring light in tissue, theoretical and experimental measurement*. Ph.D. Thesis, Heriot Watt University, Edinburgh (1991).
39. Marijnissen JP, Star WM. Calibration of isotropic light dosimetry probes based on scattering bulbs in clear media. *Phys. Med. Biol.*, 41(7):1191–1208 (1996).
40. Marijnissen JP and Star WM. Performance of isotropic light dosimetry probes based on scattering bulbs in turbid media. *Phys. Med. Biol.*, 47(12):2049–2058 (2002).
41. Lilge L, Haw T, and Wilson BC. Miniature isotropic optical fibre probes for quantitative light dosimetry in tissue. *Phys. Med. Biol.*, 38(2):215–230 (1993).
42. Dimofte A, Finlay JC, and Zhu TC. A method for determination of the absorption and scattering properties interstitially in turbid media. *Phys. Med. Biol.*, 50(10):2291–2311 (2005).
43. Pomerleau-Dalcourt N and Lilge L. Development and characterization of multi-sensory fluence rate probes. *Phys. Med. Biol.*, 51(7):1929–1940 (2006).
44. Richards-Kortum R. Fluorescence spectroscopy of turbid media. In: AJ Welch and MJC van Gemert (eds) *Optical thermal response of laser-irradiated tissue*. Plenum, New York, pp. 667–706 (1995).
45. Zhu C, Liu Quan, Ramanujam N. Effect of fiber optic probe geometry on depth-resolved fluorescence measurements from epithelial tissues: a Monte Carlo simulation. *J. Biomed. Opt.*, 8(2), 237–247 (2003).
46. Trujillo EV, Sandison DR, Utzinger U, Ramanujam N, Mitchell MF, and Richard-Kortum R. Method to determine tissue fluorescence efficiency in vivo and predict signal-to-noise for spectrometers. *Appl. Spectrosc.*, 52(7):943–995 (1998).
47. Pope K, Warren S, Yazdi Y, Johnston J, David M, and Richard-Kortum R. Dual imaging of arterial walls: intravascular ultrasound and fluorescence spectroscopy. *Proc. SPIE*, 1878:42–50 (1993).
48. Pfefer TJ, Schomacker KT, Ediger MN, and Nishioka NS. Multiple-fiber probe design for fluorescence spectroscopy in tissue. *Appl. Opt.*, 41(22):4712–4721 (2002).
49. Keijzer M, Richards-Kortum R, Jacques SL, and Feld MS. Fluorescence spectroscopy of turbid media: Autofluorescence of the human aorta. *Appl. Opt.*, 28:4286–4292 (1989).
50. Avrillier S, Tinet D, Ettori D, Tualle JM, and Gelebart B. Influence of the emission reception geometry in laser-induced fluorescence spectra from turbid media. *Appl. Opt.*, 37:2781–2787 (1998).
51. Gardner CM, Jacques SL, Welch AJ. Fluorescence spectroscopy of tissue: Recovery of intrinsic fluorescence from measured fluorescence. *Appl. Opt.*, 35:1780–1792 (1996).
52. Gardner CM, Jacques SL, Welch AJ. Light transport in tissue: accurate expressions for one-dimensional fluence rate and escape function based upon Monte Carlo simulation. *Lasers Surg. Med.*, 18:129–138 (1996).
53. Durkin AJ, Richards-Kortum R. Comparison of methods to determine chromophore concentrations from fluorescence spectra of turbid samples. *Lasers Surg. Med.*, 19:75–89 (1996).
54. Finlay JC, Conover DL, Hull EL, Foster TH. Porphyrin bleaching and PDT-induced spectral changes are irradiance dependent in ALA-sensitized normal rat skin *in vivo*. *Photochem. Photobiol.*, 73:54–63 (2001).
55. Wu J, Feld MS, and Rava R. Analytical model for extracting intrinsic fluorescence in turbid media. *Appl. Opt.*, 32:3583–3595 (1993).
56. Finlay JC and Foster TH. Recovery of hemoglobin oxygen saturation and intrinsic fluorescence with a forward-adjoint model. *Appl. Opt.*, 44(10):1917–1933 (2005).
57. Tearney GJ, Brezinski ME, Bouma BE, Boppart SA, Pitris C, Southern JF, and Fujimoto JG. In vivo endoscopic optical biopsy with optical coherence tomography. *Science*, 276:2037–2039 (1997).

58. Kawasaki M, Bouma B, Bressner J, Houser S, Nadkarni S, MacNeill B, Jang I, Fujiwara H, and Tearney G. Diagnostic accuracy of optical coherence tomography and integrated backscatter intravascular ultrasound images for tissue characterization of human coronary plaques. *J. Am. Coll. Cardiol*., 48:81–88 (2006).
59. Yang SJ, Marcon N, Gardiner G, Qi B, Bisland S, Seng-Yue E, Lo S, Pekar J, Wilson B, and Vitkin I. High speed, wide velocity dynamic range Doppler optical coherence tomography (Part III): in vivo endoscopic imaging of blood flow in the rat and human gastrointestinal tracts. *Opt. Express*, 11(19):2416–2424 (2003).
60. Yang VX, Mao YX, Munce N, Standish B, Kucharczyk W, Marcon NE, Wilson BC, and Vitkin IA. Interstitial Doppler optical coherence tomography. *Opt. Lett.*, 30(14):1791–1793 (2005).
61. Li H, Standish BA, Mariampillai A, Munce NR, Mao Y, Chiu S, Marcon NE, Wilson BC, Vitkin A, and Yang VX. Feasibility of interstitial Doppler optical coherence tomography for in vivo detection of microvascular changes during photodynamic therapy. *Lasers Surg. Med.*, 38(8):754–761 (2006).
62. Standish BA, Lee KK, Jin X, Mariampillai A, Munce NR, Wood MF, Wilson BC, Vitkin IA, and Yang VX. Interstitial Doppler optical coherence tomography as a local tumor necrosis predictor in photodynamic therapy of prostatic carcinoma: An in vivo study. *Cancer Res*., 68(23):9987–9995 (2008).
63. De Jong BWD, Schut TCB, Wolffenbuttel KP, Nijman JM, Kok DJ, and Puppels GJ. Identification of bladder wall layers by Raman spectroscopy. *J. Urol*., 168:1771–1778 (2002).
64. Notingher I, Verrier S, Romanska H, Bishop AE, Polak JM, and Hench LL. *In situ* characterization of living cells by Raman spectroscopy. *Spectroscopy*, 16:43–51 (2002).
65. Mahadevan-Jansen A and Richards-Kortum R. Raman spectroscopy for the detection of cancers and precancers. *J. Biomed. Opt*., 1:31–70 (1996).
66. Haka AS, Shafer-Peltier KE, Fitzmaurice M, Crowe J, Dasari RR, and Feld MS. Diagnosing breast cancer by using Raman spectroscopy. *Proc. Natl. Acad. Sci. U.S.A*., 102:12371–12376 (2005).
67. Haka AVolynskaya, Z, Gardecki J, Nazemi J, Lyons J, Hicks D, Fitzmaurice M, Dasri R, Crowe J, and Felds M. In vivo margin assessment during partial mastectomy breast surgery using Raman spectroscopy. *Cancer Res.*, 66:3317–3322 (2006).
68. Santos LF, Wolthuis R, Koljenović S, Almeida RM, and Puppels GJ. Fiber-optic probes for in vivo Raman spectroscopy in the high-wavenumber region. *Anal. Chem.*, 77(20):6747–6752 (2005).
69. Short M, Lam S, McWilliams A, Zhao J, and Lui Hand Zeng H. Development and prelimninary results of an endoscopic probe for potential in vivo diagnosis of lung cancers. *Opt. Lett.*, 33:711–713 (2008).
70. Schulmerich MV, Dooley KA, Morris MD, Vanasse TM, and Goldstein SA. Transcutaneous fiber optic raman spectroscoopy of bone using annular illumination and a circular array of collection fibers. *J. Biomed. Opt.*, 11(6):060502 (2006).
71. Molckovsky A, Wong Kee Song LM, Shim MG, Marcon NE, and Wilson BC. Diagnostic potetnial of near infrared Raman spectroscopy of the colon: Differentiating adenomatous from hyperplastic polyps. *Gastrointestinal Endosc.*, 57:396–402 (2003).
72. Mahadevan-Jansen A, Mitchell MF, Ramanujam N, Utzinger U, and Richards-Kortum R. Development of a fiber optic probe to measure NIR Raman spectra of cervical tissue *in vivo*. *Photochem. Photobiol*., 683:427–431 (1998).
73. Robichaux-Viehoever A, Kanter E, Shappell H, Billheimer D, Jones H III, and Mahadevan-Jansen A. Characterization of Raman spectra measured in vivo for the detection of cervical dysplasia. *Appl. Spectrosc.*, 61(9):986–993 (2007).
74. Victor XD, Tang YS, Gordon ML, Qi B, Gardiner G, Cirocco M, Kortan P, Haber GB, Kandel G, Vitkin IA, Wilson BC, and Marcon NE, Endoscopic Doppler optical coherence tomography in the human GI tract: Initial experience. *GI Endosc*., 61(7):879–890 (2005).

Chapter 18
Optical Coherence Tomography

Dirk J. Faber and Ton G. van Leeuwen

18.1 Introduction

Seventy percent of our body is made up of water. For that reason, radiation based medical imaging techniques operate in spectral regions where water absorption is low (Fig. 18.1, left panel). Well known modalities are MRI that operates at radio frequencies, and PET/SPECT which work in the high frequency range. Water absorption is also low around the part of the spectrum that is visible to the human eye. In this spectral region, scattering of the light by tissue structures roughly decreases with wavelength. Therefore, most optical imaging techniques such as (confocal) microscopy, optical tomography and Optical Coherence Tomography (OCT) use wavelengths between 650 and 1300 nm to allow reasonable imaging depths.

In addition, a trade-off exists between the depth resolution and probing depth of currently available techniques, as is shown in Fig. 18.1(b). For comparison, non-radiation based ultrasonography is also included in the figure.

Microscopy and confocal microscopy ((C)M in Fig. 18.1b) detect non-scattered light transmitted through excised tissue sections. In conventional microscopy, contrast is provided for by absorption differences in the stained section. Image resolution can be very high (laterally approximately 0.2 micron, in depth 0.7 micron), but imaging depth is very small because multiple scattering blurs the image. Imaging in scattering media is possible with (scanning) confocal microscopy, since light returning from outside the focal volume is sufficiently suppressed. Resolutions are comparable to conventional microscopy, and depth resolved imaging up to 500 μm depth is possible.

In optical tomography, methods such as time-of-flight measurements or frequency modulated optical tomography (upper right side) multiple scattered light are used for image formation and measurement of optical properties. This yields

D.J. Faber (✉)
Department of Biomedical Engineering and Physics, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands
e-mail: d.j.faber@amc.uva.nl

A.J. Welch, M.J.C. van Gemert (eds.), *Optical-Thermal Response of Laser-Irradiated Tissue*, 2nd ed., DOI 10.1007/978-90-481-8831-4_18,

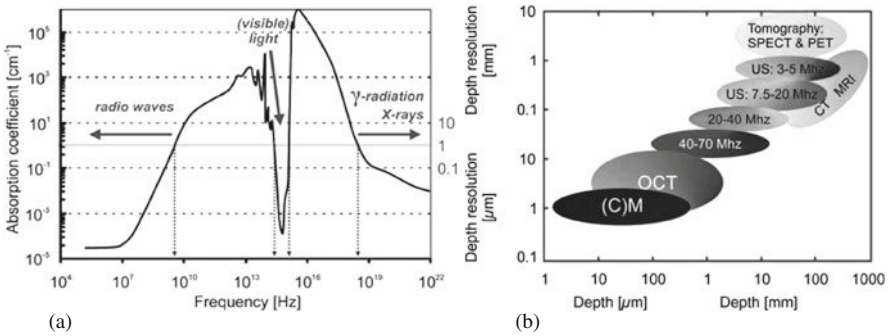


Fig. 18.1 Imaging ranges of medical diagnostic techniques in terms of resolution and imaging depth. (C)M: (confocal) microscopy; OCT: optical coherence tomography; US: ultrasonography; SPECT: single photon emission computed tomography; PET: positron emission tomography; CT: computed tomography; MRI: magnetic resonance imaging

higher imaging depths at the expense of resolution. The underlying theory and models are very complex, and consequently so are reconstruction algorithms. The gap in the spectrum of techniques in Fig. 18.1b is partly filled by Optical Coherence Tomography (OCT) [1], which permits in vivo cross-sectional imaging of superficial tissues. Compared to microscopy, OCT provides morphological details at significantly greater depths in dense tissue because it detects only light that has been scattered once or a few times.

OCT is analogous to ultrasound imaging, yielding cross-sectional images of tissue reflectivity. In ultrasound, the location of reflecting objects is determined by measuring echo delay times; the resolution depends on the frequency of the ultrasound pulses. This approach is not feasible with optical techniques given the high speed of light compared to detector response times. Instead, the location of reflecting objects is determined by using a classical optical measurement technique: white light or low coherence interferometry. In a classic Michelson interferometer (Fig. 18.2), light emitted from a source is divided by a beam splitter in two directions ('arms') towards two mirrors R and S. The back reflected beams recombine at the beam splitter, and are guided to a detector. When R is moved (while S is fixed), an interference pattern is measured by the detector. Interference will only be detected when the difference in path lengths traveled by the light in both arms is less than the so-called coherence length of the light source. A monochromatic light source (for example, a Helium-Neon laser) has a large coherence length so that the interference fringes will be seen over a large range of path length differences. If a low coherent (broad bandwidth, i.e. many wavelengths) light source is used, the coherence length can be very small: for sources used in OCT typically from 20 μm down to 1 μm.

In a clinically used setup, the sample 'replaces' the fixed mirror S, and the mirror R in the reference arm is scanned. Again, interference only occurs if the difference in path lengths traveled by the light reflected from the sample and that returning from

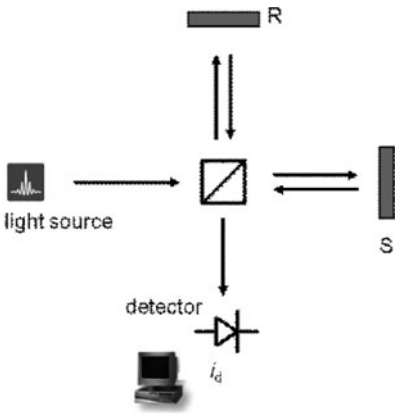


Fig. 18.2 Schematic drawing of an OCT setup. The beam splitter first directs light from the source towards mirrors S and R. After reflection, these beams are recombined and sent to the detector

the reference mirror is within the coherence length. Therefore, when interference is observed at the detector, and the path length in the reference arm is known, the path length in the sample arm and thus the position of the reflecting structure can be calculated within the accuracy of the coherence length. Although not necessarily the same because of alignment challenges, for many OCT systems the coherence length of the light source is reported to specify the axial resolution of the system. Following the terminology of ultrasound imaging, a measurement of reflectivity *vs.* depth is called an A-scan; the OCT image, or B-scan is constructed from adjacent A-scans with reflectivity now plotted on a color scale. For example, Fig. 18.3 shows an (ex vivo) OCT image (left) and corresponding histology section (right) of the aorta of a rat.

The system in Fig. 18.2 is a so-called Time-Domain OCT system because the data is recorded by scanning the optical path length difference in time. State-of-the-art systems acquire data as a function of wave number, either by fast scanning of the (instantaneous, narrowband) wavelength over a large wavelength range

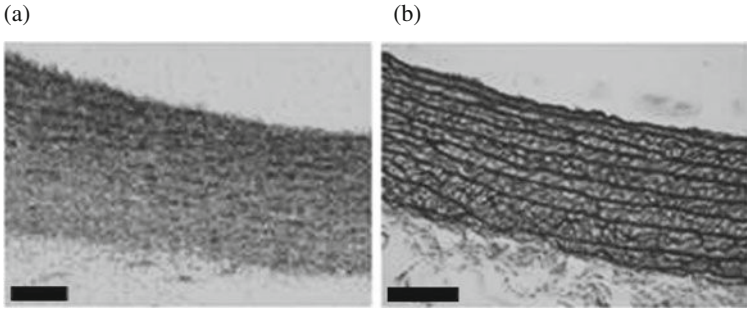


Fig. 18.3 Ex vivo OCT image (**a**) and histology (**b**) of the aorta of a rat. Elastic lamella are clearly visible on both images. Bar size: 100 μm

(Swept-source OCT, or Optical Frequency Domain Interferometry), or by acquiring the broadband spectrum using a spectrometer replacing the point detector in Fig. 18.2 (Spectral Domain OCT). There are several advantages to using these techniques instead of TD-OCT. Because it is not necessary to scan the reference mirror, much higher imaging speeds can be obtained. The sensitivity that can be obtained is also higher. As the names listed here imply, there is a Fourier transform relationship between the signals acquired with Time-domain OCT and the other techniques. The nomenclature is somewhat confusing (probably intentionally introduced to stress the not-so-large differences between the techniques). At least in this chapter, Time-Domain OCT refers to techniques where the reference mirror is *moved*; all techniques where the reference mirror is *stationary* will be called Fourier-Domain OCT.

The lateral resolution in an OCT image is determined by the focusing optics in the sample arm, similar to confocal microscopy [2, 3]. The confocal measurement technique suppresses the detection of light scattered from outside the focus. The difference between high numerical aperture (NA) optics and low NA optics is shown in Fig. 18.4: for high NA optics the lateral resolution is high, but the depth of focus is small (compared to low NA optics, where the situation is reversed).

Ideally, to achieve high axial and lateral resolution, high NA optics should be used, taking care that the measurement position and the position of the focus are matched throughout the depth scan (dynamic focusing). For clinical OCT systems, real time imaging is essential, with typical line scan rates in the order of several kHz. Because it is currently not possible to mechanically move focusing optics at this rate, optics with relatively low numerical aperture are usually used, with a large depth of focus and modest lateral resolution ($\sim$20 μm). Of course, dynamic focusing is only possible for time-domain systems; in FD-OCT the depth scan is acquired instantaneously.

The 'imaging depth' of OCT systems is the maximum depth at which subjectively good images can be obtained. It depends on both the parameters of the system and the optical properties of the tissue that is imaged. Because multiple scattering

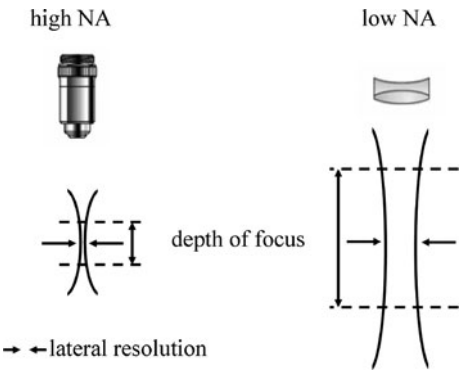


Fig. 18.4 High NA focusing with high lateral resolution but limited depth of focus (*left*) and low NA focusing, with large lateral resolution and large depth of view

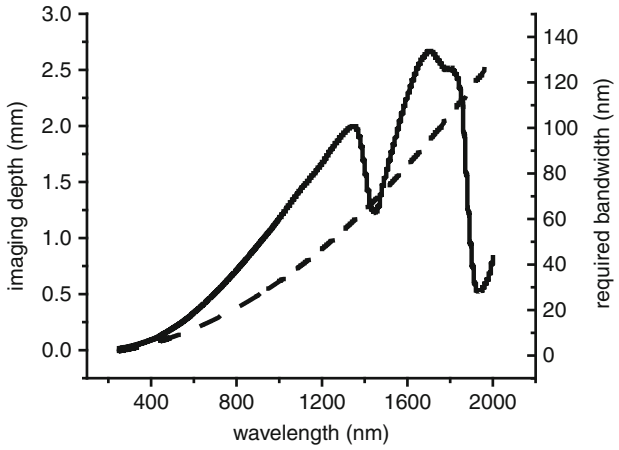


Fig. 18.5 Imaging depth (*solid curve*) and bandwidth required (*dashed curve*) of the OCT source required to maintain a resolution of 10 μm as a function of wavelength (as calculated for typical parameters of an OCT system and μ_s). Imaging depth is defined as the depth at which the contributions of the single and multiple scattering to the OCT signal are equal

blurs the image in conventional microscopy, we can arbitrarily quantify the imaging depth of OCT as the depth at which the contribution of single scattering and multiple scattering to the image is equal.

A typical OCT system used in the cardiovascular clinic, operates at 1300 nm center wavelength and has a resolution of ∼10 μm. Taking into account the refractive index of the medium, this requires ∼50 nm bandwidth of the light source. Using the absorption spectrum of water, and a typical scattering coefficient $\mu_s = 5\ \text{mm}^{-1}$, we find that the imaging depth is approximately 2 mm. Figure 18.5 shows, relative to this situation, the imaging depth (solid curve) and the bandwidth that is required to maintain 10 μm resolution (dashed curve). Calculations are according to the Extended Huygens Fresnel model for the OCT signal (see Section 18.4 for details).

To date, OCT imaging is routinely used in the clinic in ophthalmology [4, 5], but has great potential as an 'optical biopsy' tool in other fields of medicine, e.g. gastro-enterology and cardiology. There are two key factors determining the clinical relevance of OCT. The first is the ability to do 'point-like' measurements, i.e. measure tissue parameters in a highly localized volume ('functional' OCT). Basic OCT measurements record local tissue reflectivity, but localized measurement of flow/perfusion [6, 7], index of refraction [8], birefringence [9], and absorption [10, 11] and scattering coefficient [12, 13] is also feasible. The second key factor is the ability to measure these properties quantitatively. This is possible because (contrary to other optical techniques) the path length of the detected light is controlled (or known). It goes beyond the scope of this chapter to discuss each of these functional modalities and we will therefore limit ourselves to the topic of quantitative

measurements of attenuation coefficients of tissues using OCT. The (bulk) attenuation coefficient describes the attenuation of OCT signal with depth caused by light-tissue interactions, and may possibly be used to discriminate tissue structures. Furthermore, if the attenuation coefficient can be determined within certain spectral regions within the bandwidth of the OCT light source, localized spectroscopy may be feasible, e.g. to locally measure oxygen saturation [14, 15].

This chapter will start by discussing coherence gating, i.e. the OCT signal formation for both time-domain and spectral domain systems. We will then introduce the point spread function of the imaging optics to account for confocal gating. Next, the optical properties of tissues are treated. We conclude by combining the results from these paragraphs into a method for quantitatively measuring attenuation coefficients; and discuss clinical implications.

18.2 Coherence Gating

18.2.1 The Detector Current from a Michelson Interferometer

Most principles of OCT are readily illustrated using a Michelson interferometer as depicted in the left panel of Fig. 18.6. The interferometer is illuminated by an optical field E_0 traveling in the source arm towards a 50/50 beam splitter. The sample arm and reference arm of the interferometer are terminated by the mirrors S and R respectively. The detector arm is terminated by a photo detector.

The lengths of the sample arm and reference arm are l_s and l_r. The field at the detector E_d is the superposition of the fields from the sample arm E_s and reference arm E_r. Define $\Delta l = 2(l_\mathrm{s} - l_\mathrm{r})$ where the factor 2 accounts for the fact that the light traverses the sample and reference arms twice. The photo detector current i_d is proportional to the time averaged intensity on the detector,

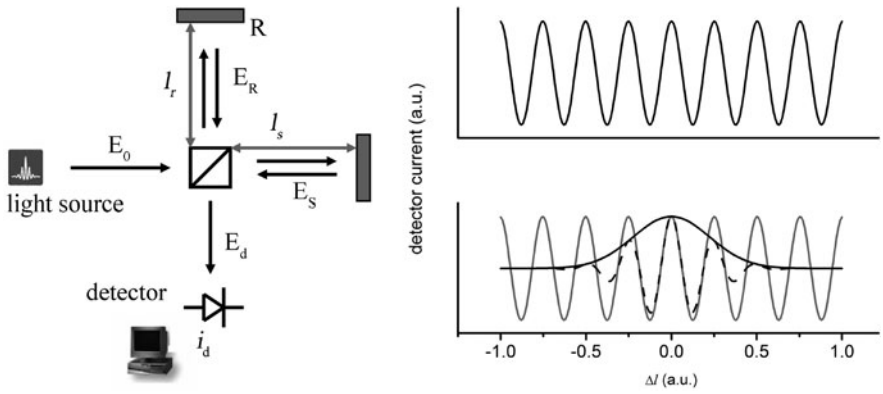


Fig. 18.6 *Left*: A Michelson interferometer. *Right*: detector current vs. path length difference for a monochromatic light source (*top*) and a broad band light source (*bottom*)

$$i_d \propto \left\langle |E_d|^2 \right\rangle \propto \left\langle |E_s + E_r|^2 \right\rangle \tag{18.1}$$

The brackets denote integration over the detector response time. Expanding the quadrature term, i_d can be separated into a term proportional to the total power in sample and reference fields, and an interferometric term of the product of E_s and E_r,

$$i_d \propto |E_s|^2 + |E_r|^2 + \left\langle E_s E_r^* + E_s^* E_r \right\rangle \tag{18.2}$$

where * denotes the complex conjugate. Only this latter part of i_d is of interest because the optical frequencies ($\sim 10^{15}$ Hz) are much higher than the cut-off frequency of the detector (typically $\sim 10^6$ Hz, but up to $\sim 10^{11}$ Hz), so the integration over the detector response time does not need to be carried out explicitly. Now consider a purely monochromatic light source, i.e. emitting only at wave number k. Then the sample and reference field are proportional to

$$E_s \propto E_0\,(k)\, e^{j(k2l_s - \omega t)} \quad E_r \propto E_0\,(k)\, e^{j(k2l_r - \omega t)} \tag{18.3}$$

respectively, where ω is the optical frequency. The interferometric signal is found by inserting Equation (18.3) in Eq. (18.2) and retaining only the last term:

$$i_d\,(k, \Delta l) \propto |E_0|^2 \left(e^{jk\Delta l} + e^{-jk\Delta l} \right) \propto S(k) \cos\,(k\Delta l) \tag{18.4}$$

where $s(k) = |E_0(k)|^2$ is the spectral amplitude of the light source. This equation will be the starting point for deriving expressions for the OCT signal for both Time-Domain and Fourier-Domain systems.

18.2.2 Time-Domain OCT

18.2.2.1 Signal Description

In time-domain OCT, the signal is recorded as a function of the position of the reference mirror in time. The total detector current is obtained by integrating (the middle term of) Eq. (18.4) over the bandwidth of the light source. We take into account that the spectrum S is centered around wavenumber k_0. The term 'cc' denotes the complex conjugate.

$$i_d(\Delta l) \propto \int_{-\infty}^{\infty} S\,(k - k_0) e^{jk\Delta l} dk + cc = R\,(\Delta l) \cos\,(k_0 \Delta l) \tag{18.5}$$

A detailed look at Eq. (18.3) reveals that the reference field is simply a time shifted version ($\Delta t = \Delta l/c$, with c the speed of light) of the sample field and both are proportional to the source field. Therefore the last term on the right hand side of

Eq. (18.2) represents the autocorrelation function of the source field. According to the Wiener-Khinchine theorem, the autocorrelation $R(\Delta l)$ is the Fourier transform of the amplitude spectrum of the light source $S(k)$ so that: $R(\Delta l) = \int S(k) \exp(jk\Delta l)\mathrm{d}k$. Here k is the wave number $k = 2\pi/\lambda$ and λ is wavelength. Note the Fourier transform pair $k \Leftrightarrow \Delta l$ and the equivalence with Eq. (18.5).

The spectrum exists on a finite wavelength (-number) range, and is therefore written as $S(k\text{-}k_0)$ where k_0 is the (center) wave number. Another well known Fourier-transform property says that this shift in the k-domain leads to a modulation in the Δl-domain, which is given by the cosine term in Eq. (18.5) which modulates $R(\Delta l)$.

In Fig. 18.6, right panel, two examples of Eq. (18.5) are depicted, corresponding to a monochromatic (top) and a broadband (bottom) light source in the interferometer. The monochromatic light source has a spectrum that is confined in the k-domain and consequently the autocorrelation function $R(\Delta l)$ in this case extends over a large part of the Δl-domain (top). Conversely, the broad band spectrum covers a large(r) part of the k-domain, and so the autocorrelation is more confined in space, i.e. an interference pattern of detectable amplitude will only exist on a small range of path length differences (bottom).

Light sources commonly used in OCT have (approximately) Gaussian spectra (ω or k-domain); therefore, the envelope of the autocorrelation (t or Δl-domain) is also Gaussian. From Fourier transform theory, both can be characterized by their $1/e$ radii w, which are related through $w_{\Delta l} = 2/w_k$. However, from an experimental point of view, it is more convenient to characterize the Gaussians by their full width at half maximum given by FWHM $= 2w\ln 2$. If the FWHM of the source spectrum $S(k - k_0)$ equals Δk, The FWHM of $R(\Delta l)$ will be $8 \ln 2/\Delta k = 2l_\mathrm{c}$. Here we have implicitly defined the coherence length l_c of the light source.

$$i_d(\Delta l) \propto e^{-4\ln 2\left(\frac{\Delta l}{2l_c}\right)^2} \cos\left(k_0 \Delta l\right) \tag{18.6}$$

Equation (18.6) describes the situation illustrated in Fig. 18.6 right, bottom panel. In practice, only the envelope of the detector signal will be plotted, as a function of $z = {}^1\!/_2\ \Delta l$ because this corresponds to a specific location of the mirror S (or depth inside a sample). The coherence length l_c is commonly expressed in terms of wavelength instead of wave number, i.e.

$$l_c = \frac{2\ln 2}{\pi} \frac{\lambda_0^2}{\Delta\lambda} \tag{18.7}$$

Here λ_0 is the center wavelength of the light source spectrum and $\Delta\lambda$ is its FWHM bandwidth. Equation (18.7) provides a practical definition of the coherence length of the light source and is commonly used to characterize the maximum axial resolution that can be achieved by an OCT system.

18.2.2.2 Sensitivity

To acquire the time-domain OCT signal, the reference mirror R is moved as a function of time [16–18]. As a result of that movement, the field in the reference arm is Doppler shifted, whereas the sample field is not. This causes a beat frequency in the interferometric part of the detector current. Because this Doppler shift is wavelength dependent ($f_{\text{doppler}} = 2V/\lambda$ where V is the velocity of the reference mirror), the detector current will have a certain center frequency and bandwidth. To study this further we need to take a closer look at the phase of the signal, given by $\varphi(k) = kn(k)\ \Delta l$ in Eq. (18.4). The refractive index is accounted for by including $n(k)$ so that the phase terms are written as $\varphi(k) = kn(k)\ \Delta l$ where Δl is (as before) the path length difference measured in air. Because of the broad bandwidth, it is convenient to write phase as series expansions around the center wave number k_0:

$$\varphi(k) = \varphi(k_0) + (k - k_0)\left.\frac{\partial \varphi(k)}{\partial k}\right|_{k=k_0} = k_0 \Delta l_\varphi + (k - k_0)\,\Delta l_g \tag{18.8}$$

where Δl_ϕ is called the phase path length difference; and Δl_g is called the group path length difference. The frequency at which a certain wave number appears in the detector current can be found (by definition) by differentiating the phase of Eq. (18.8) with respect to time.

$$2\pi f(k) = \frac{\partial \varphi}{\partial t} = k_0 \frac{\partial \Delta l_\varphi}{\partial t} + (k - k_0)\frac{\partial \Delta l_g}{\partial t} = k_0 V_\varphi + (k - k_0)\,V_g \tag{18.9}$$

The last term on the right hand side defines the phase velocity V_φ and group velocity V_g of the scanning delay line. Expressing as a function of wavelength rather then wave number,

$$f(\lambda) = \frac{V_\varphi}{\lambda_0} + \left(\frac{1}{\lambda} - \frac{1}{\lambda_0}\right) V_g \tag{18.10}$$

The center electrical frequency f_0 of the detector current is obtained from Eq. (18.10) using $\lambda = \lambda_0$; the electrical bandwidth Δf is obtained from another Taylor expansion around the center wavelength,

$$f_0 = \frac{V_\varphi}{\lambda_0}; \tag{18.11}$$

$$\Delta f = \frac{\Delta\lambda}{\lambda_0^2} V_g = \frac{V_g}{l_c}\left(\frac{2\ln 2}{\pi}\right) \tag{18.12}$$

The center frequency of the detector current depends on the *phase* scan velocity of the scanning delay line and center wavelength λ_0; the bandwidth depends on the *group* scan velocity of the scanning delay line and coherence length l_c (or equivalently, the bandwidth $\Delta\lambda$). The detection electronics have to support this

(broad) bandwidth. So there is an unfortunate trade-off between resolution and signal-to-noise ratio because higher resolution means larger bandwidth.

The Nyquist criterion states that a signal of bandwidth Δf can be completely recovered if the sampling frequency is larger then $2\Delta f$ (assuming demodulation of the signal at f_0, e.g. using a lock-in amplifier). The required sampling time to resolve signals at frequency Δf is $t_s = 1/2\Delta f$. The sampling time is also determined by $t_s = 1/f_{\text{ascan}}n$ where f_{ascan} is the a-scan frequency and n is the number of samples. Combined with Eq. (18.12), and using $V_g \approx l_{s,\max} \cdot f_{\text{ascan}}$, where $l_{s,\max}$ is the total scan range, we can solve for n to get:

$$n = \frac{4 \ln 2}{\pi} \frac{l_{s,\max}}{l_c} \tag{18.13}$$

Equations (18.11) and (18.12) are the starting point for determining the sensitivity of a time domain OCT system, which is conventionally expressed as a signal-to-noise ratio $S/N = \langle i_d^2 \rangle / \sigma_{\text{noise}}^2$ where $\langle i_d^2 \rangle$ is the mean square signal current and σ_{noise}^2 is the noise variance. The expression for the signal is found by examining the photoelectron current which is related to optical power through $i_d = \eta e P / E_v$, where η is the quantum efficiency of the detector, e is the electronic charge, P is the optical power and E_v is the photon energy. Following Eq. (18.2), the signal in units of [A^2] is then given by

$$\left\langle i_d^2 \right\rangle = \frac{2\eta^2 e^2 P_r P_s}{E_\upsilon^2} \tag{18.14}$$

where P_r and P_s are the powers from the reference and sample arms, respectively. In normal imaging conditions, the power returning from the sample will be much smaller than that returning from the reference arm, e.g. $P_s = \alpha P_r$. The two most important sources of noise in OCT are the so-called *shot noise* and *relative intensity noise*. Their noise variance is given by the sum of these two terms [A^2]:

$$\sigma_n^2 = \frac{2\eta e^2 P_r}{E_v} \Delta f + \frac{2\eta^2 e^2 P_r^2}{E_v^2} \frac{l_c}{c} \Delta f \tag{18.15}$$

where Δf is the detection bandwidth (which has to match Eq. (18.12)). Ideally, an OCT system is shot-noise limited and retaining only that noise term yields for the signal-to-noise ratio:

$$S/N = \frac{\eta \alpha P_r}{E_\upsilon \Delta f} \tag{18.16}$$

18.2.2.3 Imaging

Up until now we assumed that the 'sample' was a perfect mirror. A straightforward approximation of a 'real' sample is by describing it as a stack of m imperfect mirrors

with arbitrary spacing, i.e. the (complex) reflection coefficient of the sample will be given by:

$$r\,(k, l_s) = \sum_m h_m\,(k)\,\delta\,(l_s - l_m) \tag{18.17}$$

where δ is the delta function. The resulting signal is then found by combining Eq. (18.17) with Eq. (18.5) to get:

$$i_d(\Delta l) \propto = \sum_m R\,(\Delta l_m)\cos\,(k_0 \Delta l_m) \tag{18.18}$$

i.e. the signal is a summation of autocorrelation functions, centered at the locations $l_s = l_\mathrm{m}$.

The more general case, where the complex reflection coefficient is continuous, is found by explicitly writing out the Fourier transform in Eq. (18.5):

$$\begin{aligned} i_d(\Delta l) \propto \int_{-\infty}^{\infty} r\,(k, l_s)\,S\,(k - k_0) e^{jk\Delta l} dk + cc = \\ h\,(l_s) \otimes R\,(\Delta l)\cos\,(k_0 \Delta l) \end{aligned} \tag{18.19}$$

Here, we made use of the Fourier transform property that a multiplication in one domain leads to a convolution in the other domain. In other words, the OCT signal from an arbitrary sample with complex reflection coefficient $r(k,l_\mathrm{s})$ is given by the convolution of the spatial reflection coefficient $h(l_\mathrm{s})$ with the autocorrelation of the light source.

The refractive index of the sample and its influence on the OCT signal can be included by taking into account its wavelength dependence $n(k)$. The phase terms of Eq. (18.9) are again written as $\varphi(k) = kn(k)\Delta l$ where Δl is (as before) the path length measured in air. From the definitions of phase delay and group delay we see that

$$\Delta l_\varphi = n\,(k_0)\,\Delta l;\; \Delta l_g = n_g\,(k_0)\,\Delta l \tag{18.20}$$

where the definition of the group refractive index is used:

$$n_g = n + k\frac{\partial n}{\partial k} \tag{18.21}$$

This also means we have to alter the previous definition of coherence gating: interference will only be observed if the group path length difference between sample arm and reference arm is less than the coherence length of the light source. By knowing the group path length in the reference arm, the group path length in the sample arm can be calculated with an accuracy given by the coherence length. Because the coherence length only depends on the light source, this means that the accuracy with which the *geometrical* path length in the sample is determined, is increased by a factor $n_g(k_0)$. So, the axial resolution in such a medium will be improved by $n_g(k_0)$.

18.2.3 Spectrometer Based Fourier Domain OCT

18.2.3.1 Signal Description

In this technique, a spectrometer replaces the detector in the interferometer, and the reference mirror is fixed. The raw data now consist of a distribution of photoelectron charges collected by the CCD during the exposure time. The number of photoelectrons, n_e, is related to optical power as $n_e = \eta P\tau/E_v$, where η is the quantum efficiency of the spectrometer, P is the optical power, τ is the exposure time and E_v is the photon energy. The OCT signal is recorded as a function of wave number k (sampled through the pixels of the spectrometer) and is simply given by a re-written Eq. (18.5):

$$i_{\text{spect}}(k) \propto \frac{\eta\tau}{E_v} S(k)\left(1 + \alpha + 2\sqrt{\alpha}\cos(k\Delta l)\right) \tag{18.22}$$

Following the proceedings of the previous paragraph, the Δl-domain signal is obtained by Fourier transforming Eq. (18.22) with respect to k. The spectrum S will transform to the autocorrelation function R; the modulation term $\cos(k\Delta l)$ will cause a shift the autocorrelation to the position Δl. The result is given by Eq. (18.6).

The spectrometer supports only a limited spectral bandwidth Δk, which is divided in n pixels of finite size. From the Nyquist theorem, the maximum path length difference that can be resolved is given by $\Delta l = \pi/\delta k$, or by arbitrarily setting the reference arm length l_r to zero, the maximum probing depth in the sample $l_{s,\max}$ is given by:

$$l_{s,\max} = \frac{\pi}{2\delta k} = \frac{\lambda_0^2}{4\delta\lambda} \approx \frac{n\lambda_0^2}{4\Delta\lambda} \approx \frac{nl_c}{4} \tag{18.23}$$

The last terms of this equation are a rule of thumb that assume the optical bandwidth supported by the spectrometer equals the FWHM of the spectrum of the light source. In general the spectrometer will support a larger bandwidth then that.

Because the spectrometer data is a real signal, the Fourier transform will be an even function. In other words, mirror signals appear on the Δl-axis on the range $[-\pi/\delta k, 0]$ and $[0, -\pi/\delta k]$ centered around zero delay. Consequently (if no other measures are taken to avoid this so called complex ambiguity), one has to make sure that the length of the reference arm corresponds to a position well outside the sample as, otherwise, the image will be folded around zero delay. It also means that per a-scan only $n/2$ pixels are available. Using the rule of thumb in Eq. (18.23), this means that the pixel resolution in the a-scan reduces to $l_c/2$.

18.2.3.2 Sensitivity

To analyze the sensitivity performance of a spectrometer based systems, we consider it a series of n parallel detectors (corresponding to the n pixels of the CCD array) and analyze signal and noise per detector element [19–21]. The signal-to-noise ratio

is defined in the Δl-domain, which means we need to take into account the Fourier transformations of signal and noise variances. Scaling factors (1, n or $1/n$ depending on your Discrete Fourier Transform definition) apply to signal and noise, and are consequently ignored.

The signal power is simply divided by n detector elements in the k-domain and then multiplied by $n/2$ points per depth scan in the Δl-domain to get the total signal power:

$$\left\langle i^2 \right\rangle = \frac{\eta^2 \alpha P_r^2 \tau^2}{E_\upsilon^2} \tag{18.24}$$

Per detector element, the reference arm power is reduced by a factor n. For the relative intensity noise term (Eq. (18.15)), however, because each detector element 'sees' a narrower wavelength spectrum, the coherence length appearing in the relative intensity noise term is increased by a factor n. Eq. (18.15) is then rewritten as:

$$\sigma_{\text{pixel}}^2 = \frac{\eta P_r \tau}{E_\upsilon n} + \frac{\eta^2 P_r^2 \tau^2}{E_\upsilon^2 n^2} \frac{n l_c}{c} \tag{18.25}$$

Considering a shot noise limited system, we retain only the first term of Eq. (18.25). Finally, Eq. (18.16) is modified to:

$$S/N = \frac{\eta \alpha P_r \tau}{E_\upsilon} \tag{18.26}$$

If we compare this to Eq. (18.16), it is clear that the sensitivity differences lie in the differences between τ for spectral domain OCT vs. $1/\Delta f$ for time-domain OCT. We can take a closer look by starting with the expression for Δf which is given by the middle term of Eq. (18.12), and use $V_g = l_{s,\max} \cdot f_{\text{scan}}$, where f_{scan} is the number of a-scans per second. To make a fair comparison, we substitute 18-23 for the maximum imaging depth $l_{s,\max}$. The result is that $\Delta f_{\text{TD}} = n \cdot f_{scan}/4$. Finally, in spectral domain OCT, the ideal bandwidth is given by $\Delta f_{\text{SD}} = 1/2\tau = f_{\text{scan}}/2$. Consequently, the signal-to-noise ratio is increased by a factor $n/2$ for spectral domain OCT compared to time domain OCT.

In spectral domain OCT, the sensitivity (or S/N) is depth dependent for two reasons: first, the pixels of the spectrometer have a finite square size of bandwidth δk which we denote by $\Pi(k, \delta k)$; and second, the resolution of the spectrometer $\Delta k/n$ can be different from δk (i.e. a monochromatic light source could illuminate more then one pixel). This beam profile in the spectrometer is given by $G(k,\Delta k)$. Consequently, $S(k)$ in Eq. (18.22) should be replaced by $S(k) \otimes \Pi(k, \delta k) \otimes G(k, \Delta k)$, because the spectrum is convolved with the pixel bandwidth and beam profile in the spectrometer. The convolutions in the k-domain correspond to multiplications in the Δl-domain so that, eventually, we can describe the decay of sensitivity $\beta(\Delta l)$ as:

$$\beta(\Delta l) = \Im_k \{\Pi(k, \delta k) \otimes G(k, \Delta k)\} = \Im_k \{\Pi(k, \delta k)\} \cdot \Im_k \{G(k, \Delta k)\} \tag{18.27}$$

where the symbol $\Im_k$ denotes the Fourier transform with respect to k. The 'pixel function' $\Pi(k, \delta k)$ is a rectangle function defined as $\Pi(k) = 1$ if $|k| < \delta k/2$ and 0 everywhere else. The Fourier transform of this function is given by $\text{sinc}(\Delta l \cdot \delta k/2) = \text{sinc}(l_s \cdot \delta k)$ where we again set the reference path length l_r to zero and substitute $\Delta l/2$ with the sample arm length l_s.

If we assume that the beam profile in the spectrometer if Gaussian, with FWHM given by the $\Delta k/n$, we have $G(k) = \exp[-4 \ln 2(nk/\Delta k)^2]$. The Fourier transform of $G(k)$ is given by $\exp[-n^2 \Delta l^2 \cdot \Delta k^2/8 \ln 2]$. If we assume that $\Delta k/n = s\delta k$ (i.e. the resolution of the spectrometer is given by a factor s times the pixel resolution) and re-arrange terms, we can write the decay of sensitivity with depth l_s as:

$$\beta(l_s) = \text{sinc}\,(l_s \delta k) \exp\left(-\frac{(Sl_s \delta k)^2}{2\text{In}2}\right) = \text{sinc}\left(\frac{\pi}{2}\frac{l_s}{l_{s,max}}\right) \exp\left(-\frac{\pi^2}{8\text{In}2}\left(\frac{Sl_s}{l_{s,max}}\right)^2\right) \tag{18.28}$$

18.2.3.3 Imaging

If the sample consists of a mirror at certain position l_s, Eq. (18.22) shows that the raw spectrometer output is the source spectrum $S(k)$, cosine modulated with 'frequency' $\Delta l = 2(l_s - l_r)$. As with time-domain OCT, we will start by looking at the sample as a stack of imperfect mirrors as described by Eq. (18.16). The spectrometer output in this case will be given by:

$$i_{\text{spect}}\,(k) \propto \sum_n S(k) \cos\,(k \Delta l_n) \tag{18.29}$$

The Δl domain signal is obtained by Fourier transforming Eq. (18.29) which results in the same expression as Eq. (18.18):

$$i_d(\Delta l) \propto \sum_n R_{ii}\,(\Delta l_n) \cos\,(k_0 \Delta l_n) \tag{18.30}$$

and, consequently, the OCT signal from an arbitrary sample with complex reflection coefficient $r(k,l_s)$ is given by the convolution of the spatial reflection coefficient $h(l_s)$ with the autocorrelation of the light source as described by Eq. (18.19).

18.2.4 Swept Source Based Fourier Domain OCT

18.2.4.1 Signal Description

In swept source OCT (also called Optical Frequency Domain Interferometry), the light source sweeps continuously over a large optical bandwidth Δk, with very narrow instantaneous bandwidth δk. The OCT signal is detected using a PIN detector just as in time-domain OCT.

We assume a sweep-rate of $\kappa = dk/dt$; and a linear sweep of $k(t) = k_0 + K{\cdot}t$ (which is achievable, for example, by using an acousto-optical tunable filter as wavelength selector [22]).The optical frequency is given by $\omega(t) = c{\cdot}k(t)$. Because the light travels different lengths (and thus times) in sample and reference arms, the fields from both arms at the detector correspond to different times of emission from the light source and have different wave numbers.

To obtain the OCT signal, we start out by re-writing Eq. (18.3) by explicitly including this time dependence of k and ω:

$$\begin{aligned} E_s(t) &\propto E_0\left(k\left(t - l_s/c\right)\right) e^{j(k(t-l_s/c)2l_s - \omega(t-l_s/c)t)} \\ E_r(t) &\propto E_0\left(k\left(t - l_r/c\right)\right) e^{j(k(t-l_r/c)2l_r - \omega(t-l_r/c)t)} \end{aligned} \tag{18.31}$$

Working out Eq. (18.1) we end up with:

$$i_{\text{swept}}(k(t)) \propto \frac{\eta e}{E_v} S(k)\left(1 + \alpha + 2\sqrt{\alpha}\cos\left(\left[k_0 + \mathrm{K}(t - t_0)\right]\Delta l\right)\right) \tag{18.32}$$

where $t_0 = (l_\mathrm{r} + l_\mathrm{s})/c$. This is of the same form of Eq. (18.22) for spectrometer based OCT. The OCT signal as function of depth is obtained through the Fourier transform of Eq. (18.32). The maximum imaging depth is given by Eq. (18.23), where it should be noted that the instantaneous bandwidth δk_{swept} is, in general, smaller than the pixel resolution δk_{spect} in spectrometer based OCT.

18.2.4.2 Sensitivity

The derivation of the S/N follows the same lines as that for spectrometer based OCT and will result in an $n/2$ improvement compared to time-domain OCT [21, 23]. As with spectrometer based Fourier domain OCT, the sensitivity is depth dependent. The derivation is analogous to the one leading to (and also results in) Eq. (18.28). We only need to re-define the spectral interval δk (the finite pixel width in spectrometer based OCT) to the k-sampling interval $\delta k_{\text{samp}} = \Delta k/n$ in swept source OCT; and the factor s (the ratio of the spectrometer resolution to the pixel resolution in spectrometer based OCT) as the ratio of the instantaneous line width of the swept laser to the k-sampling interval in swept source OCT $s = \delta k_{\text{swept}}/\delta k_{\text{samp}}$. Again, since, in general, the instantaneous bandwidth δk_{swept} is smaller than the pixel resolution δk_{spect} in spectrometer based OCT, the sensitivity roll-off with depth will be less for swept source OCT compared to spectrometer based OCT.

18.2.4.3 Imaging

Imaging works the same as for spectrometer based OCT.

18.3 Confocal Gating

OCT can be regarded as an extension of confocal microscopy (CM) [24]. The optical arrangement of CM is shown in Fig. 18.7. An image of the source pinhole (sp) is created in the sample overlapping with an image of the detection pinhole (dp). Because of this arrangement, light scattered from outside that location is not imaged onto the detection pinhole; in other words, detection of light scattered from outside the focal volume (symbolized by the 'dot' in Fig. 18.7) is suppressed.

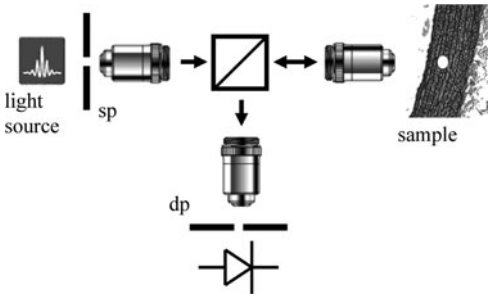


Fig. 18.7 Confocal microscopy. The source pinhole (sp) is imaged in the tissue. Co-localized is an image of the detection pinhole (sp). The location of this focal volume is symbolized by the 'dot' in the sample

The resolution in CM thus depends on the size of the focal volume. High numerical aperture (NA) optics provide a small depth of focus and small lateral spot size. The confocal system can be characterized by the detector signal when a small reflector is imaged. The response to this reflector as a function of its radial and axial position with respect to the focal volume is referred to as the 2-dimensional point spread function (PSF).

OCT is also based on a confocal optical setup. In most clinical systems, single mode fibers (SMF) serve as the confocal pinholes. As with CM, the lateral resolution is determined by the optics, whereas, contrary to CM, the axial resolution is determined by the coherence length of the light source. The combination of coherence and confocal gating provides for the high localization of the probe volume in OCT.

Using high NA optics in OCT, both high axial resolution and high lateral resolution can be obtained. Because the (confocal) depth of focus in such a setup is limited, ideally the positions of the coherence and confocal gates are matched throughout the depth scan. Unfortunately, moving the focusing lens at high velocity is mechanically not possible and therefore clinical OCT systems (which rely on high imaging speed) usually deploy relatively low NA 'fixed' optics. This means that the strength of a reflection in the OCT image depends on its position relative to the focus position. In other words, to quantitatively analyze such OCT data, the PSF has to be known and taken into account.

In this section the PSF for SMF based OCT systems will be derived, using the coupling efficiency η_c of Gaussian intensity profiles, launched from one SMF into another SMF using an arbitrary optical system [25]. A Gaussian beam is characterized by the waist w and its position and the Rayleigh length Z_0 (half depth of focus). The Rayleigh length is given by:

$$Z_0 = \frac{\pi n w^2}{\lambda} \tag{18.33}$$

with wavelength λ (in a vacuum) and the index of refraction n of the medium. In calculating η_c, it has to be taken into account that only an effective beam, equivalent to the fundamental mode of the receiving fiber, can be coupled in. Specifically, the light from one SMF (mode field diameter $2w_f$, Rayleigh length Z_f) is imaged to a new waist w_i and Rayleigh length Z_i. The fundamental mode of the detection optical system and SMF is characterized by mode field diameter $2w_r$, Rayleigh length Z_r. The coupling efficiency can be calculated as the overlap integral of the two Gaussian fields (and not their overlap area, as with multimode fibers). Then, η_c depends on any losses due to offset δ between the waist positions, lateral misalignment x, angular misalignments and mode field diameter mismatch between the two beams (Fig. 18.8).

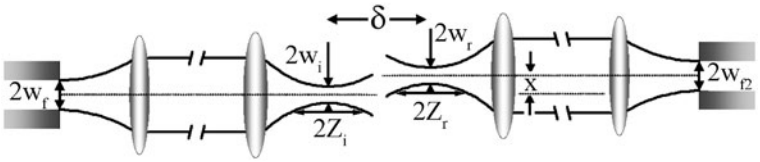


Fig. 18.8 Geometry of coupling light from a single mode fiber, via a lens system into another lens/fiber system taking into account both axial (δ) and lateral (x) misalignment, and mode field diameter mismatch ($2w_i$,$2w_r$)

In the following, angular misalignments will not be considered for the sake of simplicity. The radial distribution of the Gaussian field can be written as

$$E(r) = \sqrt{\frac{2}{\pi}} \frac{1}{w} e^{-\frac{r^2}{w^2}} \tag{18.34}$$

The overlap integral can be calculated at any position in the optical system, in our case the waist position of the receiving optical system. The complex Gaussian field of Eq. (18.34) will diverge from waist (w_i) position over the distance δ to:

$$E_i(r) = \sqrt{\frac{2}{\pi}} \frac{1}{w'_i} \exp\left(\frac{-r^2}{w_i'^2}\right) \exp\left(\frac{-ikr^2}{2R_i'^2}\right) \exp\left(-i\left[k\delta - \theta'\right]\right) \tag{18.35}$$

with

$$w'_i = w'_i(\delta) = w_i\sqrt{1 + \delta^2/Z_i^2}\ R'_i = R'_i(\delta) = \delta\left(1 + Z_i^2/\delta^2\right)\ \theta' = \arctan\left(\delta^2/Z_i^2\right) \tag{18.36}$$

The field of the fundamental mode $E_r(r)$ of the receiving fiber is simply given by Eq. (18.34). For these fields, the coupling coefficient (overlap integral) is calculated as:

$$\eta_c = \int_{-\infty}^{\infty} E_i E_r^* dr \tag{18.37}$$

The power transmission T is then given by $\eta_c \eta_c^*$ (see [25] for details):

$$T = \frac{4\left(\frac{w_r}{w_i}\right)^2}{\left(\frac{\delta}{Z_i}\right)^2 + \left[\left(\frac{w_r}{w_i}\right)^2 + 1\right]^2} \exp\left(\frac{-2\left(w_r^2 + w_i^2\right)\left(x/w_r^2\right)^2}{\left(\frac{\delta}{Z_i}\right)^2 + \left[\left(\frac{w_i}{w_r}\right)^2 + 1\right]^2}\right) \tag{18.38}$$

The Rayleigh length Z_i is defined similar to Eq. (18.33). This expression is a function of the parameters of the Gaussian beams only (and their misalignment). From this general expression, three forms of the point spread function (PSF) will be derived.

18.3.1 PSF for a Point Reflector

Consider a small point reflector, somewhere in the sample arm beam at axial location d and lateral location x. We can think of the light reflected back by this point into the OCT system as a beam with a small waist (corresponding to the size of the point reflector), with waist position at the position of the point reflector. The response to this point reflector (PSF) can then be calculated from Eq. (18.38). Now d, x, are the coordinates of the point reflector with respect to the center of the focal volume of the beam, w_r describes the fundamental mode of the probe beam, and $w_i << w_r$. This yields for the (normalized) point spread function:

$$T(d,x) \propto \frac{1}{\left(\frac{d}{Z_r}\right)^2 + 1} \exp\left(\frac{-2\left(x/w_r\right)^2}{\left(\frac{d}{Z_r}\right)^2 + 1}\right) \tag{18.39}$$

after first re-arranging the term before the exponential, by multiplying the nominator and denominator with $[w_i^4/w_r^4]$, the Rayleigh length Z_r is defined similar to Eq. (18.33). The normalization factor (of course) depends on w_i and is given by $(2w_i/w_r)^2$. Figure 18.9 shows a surface rendering (left) and contour plot of this point spread function, using $w_\mathrm{r} = 4$ μm; $\lambda = 800$ nm; Z_r=65 μm. Reflector size is $w_i = 100$ nm.

18.3.2 PSF for Specular Reflection

Specular reflection occurs at smooth interfaces such as mirrors, air-glass or air-water boundaries and is governed by Fresnel's laws of reflection. The PSF in this case can be measured as the detector signal when moving a mirror through the focus of the

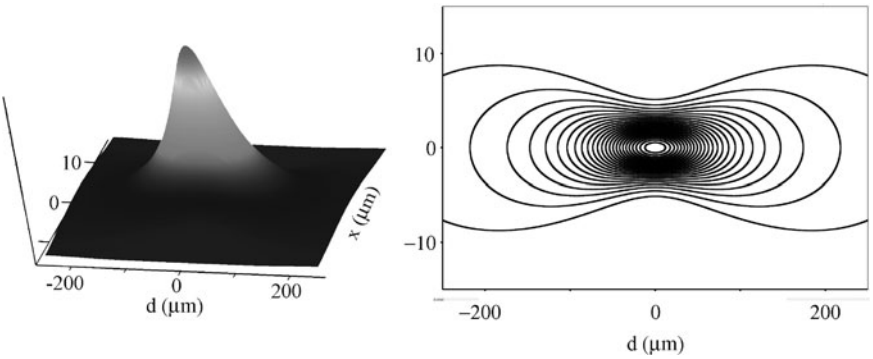


Fig. 18.9 *Left*: Surface map of the PSF of Eq. (18.39) *Right*: contour map representation of the PSF of Eq. (18.39). All spatial units are μm

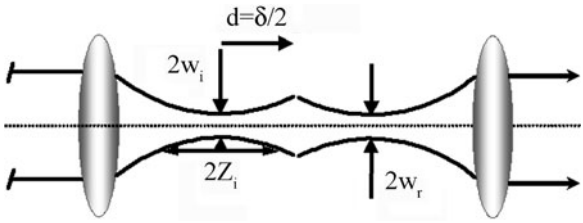


Fig. 18.10 Geometry used in the calculation of Eq. (18.25), PSF of specular reflection

beam. Taking into account that in our OCT systems the same fibers and lenses are used for delivery and collection of the light, i.e. $w_i = w_r$, the geometry can be depicted as in Fig. 18.10 (where the mirror is not drawn; rather, the beam paths are 'folded' around the location of the mirror). Lateral and angular offsets are not present.

When the mirror is positioned at axial distance d from the focus, the actual distance between the waist of incident and receiving beams is given by $\delta = 2d$. This greatly simplifies Eq. (18.38) and yields the PSF for specular reflection:

$$T_s(d) = \frac{1}{\left(\frac{d}{Z_i}\right)^2 + 1} \tag{18.40}$$

with d the distance of the reflecting object to the waist position of the beam.

18.3.3 PSF for Reflection from Backscattering

In the case of reflection caused by backscattering, in analogy of the paper of de Grauw et al. [26], illumination of an object at distance d from the waist position of the incoming beam will form a secondary source. This source will be treated as a new beam with waist ω_r and Rayleigh length Z_r which will then overlap with the incoming beam (Fig. 18.11).

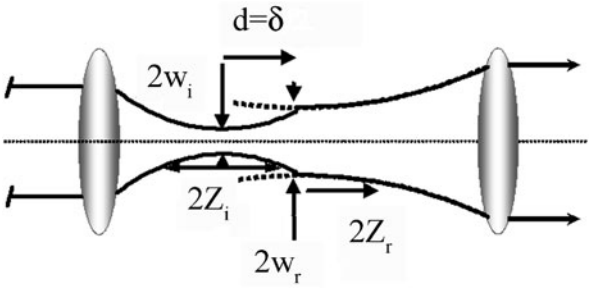


Fig. 18.11 Geometry used in the derivation of Eq. (18.42), the PSF for diffuse reflection

The propagation of the incident beam to the scattering location is calculated in the same way as for Eq. (18.35). The beam propagates to position d according to:

$$w_r(d) = w_i\sqrt{1+\frac{d^2}{Z_i^2}} \quad Z_r(d) = Z_i\left(1+\frac{d^2}{Z_i^2}\right) \tag{18.41}$$

Note that the distance between the waist positions δ equals d. Combining this result with Eq. (18.38), the axial PSF for diffuse reflection can be described as:

$$T_d(d) = \frac{1}{\left(\frac{d}{2Z_i}\right)^2+1} \tag{18.42}$$

i.e., for reflection from backscattering, the depth of focus is doubled compared to the case of specular reflection. In the derivation of Eq. (18.42), we assume the beam is not distorted prior to, or after, backscattering.

18.4 Quantitative Measurements of Optical Properties with OCT

In the derivations in Section 18.2, we considered a single reflector in the sample arm, an artificial sample modeled as a stack of mirrors, and a 'generic sample' described by a complex reflectivity $h(l,k)$. The latter, and thus the recorded OCT signal, will depend on the optical properties of the tissue. As argued before, quantitative measurement of the optical properties can help to differentiate between different tissues. In the previous section, we derived an expression for the point spread function which also influences the OCT signal *vs*. depth. Here, we start by briefly introducing the optical properties of interest. We proceed by giving expressions for the OCT signal in two cases: assuming single backscattering; and when assuming multiple, forward scattering.

18.4.1 The Optical Properties of (Bulk) Tissue

The relevant optical properties of (bulk) tissue are the refractive index, absorption coefficient, scattering coefficient and scattering anisotropy. The refractive index and its influence on the OCT signal can be accounted for by taking into account the wavelength dependence $n(k)$ as before. Absorption by a single particle (or molecule) is governed by its absorption cross section σ_a. A medium containing many absorbers is then characterized by the absorption coefficient $\mu_a = N\sigma_a$, where N is the number of particles per unit volume. The absorption coefficient equals the probability of absorption per unit path length. Likewise, scattering by a single particle is characterized by the scattering cross section σ_s; for a medium the scattering coefficient μ_s (mm^{-1}) is the probability of scattering per unit path length. Their sum is the (total) attenuation coefficient μ_t, its reciprocal is the mean free path, i.e. the expected value of the path length of photon travels *without* interaction. If a photon is scattered, its direction may be changed. The probability function of the scattering angle is called the phase function. Often, the Henyey-Greenstein phase function is used, which is described as a function of one parameter only, the scattering anisotropy g (between -1 and 1) which by definition is the expected value of the cosine of the scattering angle. A g-value of 1 means purely forward directed scattering (i.e. no change of direction); a value of -1 means purely backward directed scattering. If $g = 0$, there is equal probability of scattering into all angles.

Here, *independent* scattering will be assumed. Strictly speaking, this means that the scattered field by one particle is not influenced by the field scattered by other particles of the studied medium. In other words, there is no strict relation between the phases of the different scattered fields. In terms of low-coherence illumination, this means that we assume that the scatterers are separated by more than one coherence length. However, in closely packed media, effects such as shielding may occur, effectively lowering the scattering cross section of the particle. In many practical cases, e.g. scattering from dense media such as blood, the assumption of independent scattering may not hold and as a direct consequence, the scaling of the microscopic cross sections to macroscopic coefficients with concentration can be questionable.

The interaction of light with a medium is due to fluctuations in the refractive index m (where the fluctuations can be discrete particles or more continuous variations). The refractive index is a complex quantity, and a function of frequency $m(\omega) = n(\omega) + i\kappa(\omega)$. By the principle of causality, $n(\omega)$ and $\kappa(\omega)$ are related through Kramers-Kronig relations,

$$n(\omega) - 1 = \frac{2}{\pi} P \int_0^\infty \frac{\omega' \kappa(\omega')}{\left(\omega^2 - \omega'^2\right)} d\omega' \tag{18.43}$$

and the inverse transform given by

$$\kappa(\omega) = \frac{2\omega}{\pi} P \int_0^\infty \frac{n(\omega') - 1}{\left(\omega^2 - \omega'^2\right)} d\omega' \tag{18.44}$$

Here, the symbol P denotes the principle value of the integral. A solution of the Maxwell equations for a homogeneous medium corresponds to a plane wave, traveling in the z-direction given by $E = \exp(jkmz - j\omega t)$, so:

$$E = \exp(-k\kappa z) \exp(jknz - j\omega t) \tag{18.45}$$

which means the wave is damped when the imaginary part of the refractive index is non-zero. The absorption coefficient μ_a of the medium is then given by $2k\kappa$ or, rearranging, the imaginary part of the refractive index is given by

$$\kappa(\omega) = \frac{c\mu_a(\omega)}{2\omega} \tag{18.46}$$

where c is the speed of light in a vacuum. Understandably, κ is often called absorption coefficient. Mie theory [27] provides a formal solution to Maxwell's equations for a homogeneous sphere of radius r. It depends on complex refractive index m, but the parameter $x = k \cdot m \cdot r$ (with k the wave number as before) is equally important, since the equations have to fulfill boundary conditions at the surface of the sphere. The calculations result in scattering anisotropy g, and the extinction and scattering efficiencies Q_{ext} and Q_{sca}, dimensionless quantities which are the ratios of the extinction (scattering) cross section to the geometrical cross section of the particle.

18.4.2 Single Backscattering

Consider a homogeneous sample in the sample arm. The locations of coherence and confocal gates are matched. We first calculate backscattering coefficient $h(z)$ in a one-dimensional approximation, and consider independent, single backscattering only. Then the probability of backscattering is independent of z, and the input field is attenuated according to Beer's law. Consequently, $h(z)$ is proportional to

$$h(z) \propto e^{-2\alpha z} e^{-jk2z} \tag{18.47}$$

where the factor 2 accounts for round-trip attenuation, and α is the field attenuation coefficient. The light intensity returning from the sample I_s will be proportional to $I_s \propto \left|E_s E_s^*\right| \propto h(z)h(z)$. The coefficient α is then related to the attenuation coefficient μ_t through:

$$I_s \propto e^{-4\alpha z} = e^{-2\mu_t z} \tag{18.48}$$

Using the $h(z)$ of Eqs. (18.47) and (18.48) yields for the OCT signal:

$$i(z) \propto \int_{-\infty}^{\infty} R\left(z - z'\right) e^{-\mu_t z'} dz' \approx e^{-\mu_t z} \tag{18.49}$$

The last approximation holds when the width of R is much less then the mean free path; i.e. when $l_c << 1/\mu_t$ the autocorrelation function can be approximated by a delta function $\delta(z–z')$. This is the widely used single backscattering model. Note that if the tissue is layered, with the given assumptions this equation holds inside each layer.

The attenuation coefficient is a function of wave number. From Eq. (18.19), the reflectivity in the spatial frequency domain, $H(k)$ can be written as the Fourier transform of $h(z)$ from Eq. (18.47):

$$H\left(k\right) \propto \int_{0}^{\infty} \left(e^{-\mu_t(k)z} e^{-jk2z}\right) e^{jk2z} dz = \int_{0}^{\infty} e^{-\mu_t(k)z} dz = \frac{1}{\mu_t\left(k\right)} \tag{18.50}$$

Here the Fourier transform pair is $k \Leftrightarrow \Delta l = 2z$ as before. The second integral on the right hand side of Eq. (18.50) shows that $H(k)$ is a the sum of attenuations $\exp(-\mu_t z)$ due to single backscattering at all depths z in the sample. The local spectral properties of a sample are of more interest. The local spectrum of the detector current $I_z(k) = S_{ii}(k)\, H_z(k)$ is calculated from:

$$H_z(k) = \int_{z}^{z+\Xi} e^{-\mu_t(k)z} dz = \frac{1}{\mu_t\left(k\right)} e^{-\mu_t(k)z} \cdot \left(1 - e^{-\mu_t(k)\Xi}\right) \approx \Xi e^{-\mu_t(k)z} \tag{18.51}$$

The approximation in the last term on the right hand side holds if the observation window Ξ is small compared to the mean free path length $1/\mu_t$.

Localized spectroscopic information at z is obtained from time-frequency analysis like a windowed short time Fourier transform (STFT), on a segment of $i_d(z)$. Note that for a non-homogeneous medium, still assuming single, independent backscattering, the amplitude of $H(k) \propto \exp(-\int \mu_t(k, z')\mathrm{d}z')$.

18.4.3 Backscattering and Multiple Forward Scattering

A model taking into account multiple scattering (neglecting absorption) was derived by Thrane et al. [28, 29]. The OCT signal is expressed as the mean square signal current and is computed from the integral of sample and reference fields over the detector area. To this end, first the mean backscattered irradiance distribution $< I_B(r) >$ at the 'reflection site', i.e. the location of the coherence gate is calculated:

$$\langle I_B(r)\rangle \propto \frac{\exp(-\mu_s z)\exp\left(-r^2/w_h^2\right)}{w_h^2} + \frac{\left[1-\exp(-\mu_s z)\right]\exp\left(-r^2/w_s^2\right)}{w_s^2} \quad (18.52)$$

Here r is the radial coordinate as before. The first term on the right-hand side is the remaining un-scattered distribution, i.e. attenuated according to Beer's law, thus proportional to $\exp(-\mu_s z)$; the second term represents a broader 'halo' due to multiple forward scattering, i.e. proportional to $[1-\exp(-\mu_s z)]$. The quantities W_h and W_s are the $1/e$ intensity radii of the probe beam without and with scattering, respectively. They are given by:

$$W_h^2 = W_0^2\left(\frac{z_{cf}-z}{nf}\right)^2 + \left(\frac{f}{kW_0} - \frac{z_{cf}-z}{nkW_0}\right)^2 \quad (18.53)$$

$$W_s^2 = W_h^2 + \left(\frac{2}{k\rho\left(z_{cf},z\right)}\left[f - \frac{z_{cf}-z}{n}\right]\right)^2 \quad (18.54)$$

Here, W_0 is $1/e$ intensity radius of the probe beam at the position of the focusing lens, z_{cf} and z are the locations of the confocal and coherence gate in the tissue; f is the focal length of the lens; n is the index of refraction; and $\rho_0(z_{cf}, z)$ is the lateral coherence length given by:

$$\rho_0\left(z_{cf},z\right) = \sqrt{\frac{3}{\mu_s z}}\frac{\lambda}{\pi\theta}\left(\frac{nf}{z} - \frac{z_{cf}-z}{z}\right) \quad (18.55)$$

where θ is the root mean square scattering angle which is non-linearly related to the scattering anisotropy of the sample. The OCT signal can then found to be proportional to:

$$i_d(z) \propto \frac{1}{W_h}\left[\exp(-2\mu_s z) + \frac{2\exp(-\mu_s z)\left[1-\exp(-\mu_s z)\right]}{1+W_s^2/W_h^2} + \left[1-\exp(-\mu_s z)\right]^2\frac{W_h^2}{W_s^2}\right]^{\frac{1}{2}} \quad (18.56)$$

The first term in the brackets is the single back-scattered contribution; the last two terms describe the multiple scattering contribution. The theory presented in this section is valid in the paraxial regime, i.e. small angle forward scattering ($g \geq 0.75$). Strictly speaking, this formalism only applies to the case where all of the scattered light is contained within the collection solid angle of the optical system which is similar to our assumption in Section 18.3, that the beam is not distorted due to scattering prior to, and after, the backscattering event.

The detection of multiple scattered light hampers both OCT imaging and the quantitative measurement of optical properties from the OCT data, and thus the identification of different (pathologic) tissues. Not surprisingly, the regimes of single and multiple scattering are a subject of many studies [30–34].

There are a few multiple scattering models for the OCT signal available in the literature [35, 36]. Unfortunately, most are not analytical. The most intuitive model

is introduced by Karamata et al. where it is shown that the OCT signal is described as the convolution of the autocorrelation function of the light source R_{ii} with the reflectance profile of the sample R. The latter needs to be obtained from Monte Carlo simulations.

18.4.4 Quantitative Attenuation Coefficient Measurements

The general approach for extracting attenuation coefficients from the measured OCT data is to establish a mathematical model for the OCT signal, and then fitting that to the measured data from some region of interest in the OCT image [13, 24]. For a moment, assuming that the right model can indeed be established, the fitting routine will give best estimates for the fit parameters (including the sought after attenuation coefficient) and the associated confidence intervals. The quality of accuracy of the results depends on the number of parameters in the fit, which means that, whenever possible, the system should be calibrated, and the calibrations should enter as 'constants' in the fitting procedure.

As we discussed in the preceding paragraphs, to establish the model we must choose between the single and multiple scattering model. When possible, the single scattering model is favored because of its simplicity and fewer parameters. If a spectral domain system is used, we need to account for the sensitivity roll-off in depth as described by Eq. (18.28). Luckily, for any given OCT system the roll-off can be calibrated by measuring the response from a mirror at different depths. The point spread function of Eq. (18.42) can be used to incorporate confocal gating (unless dynamic focusing is used in a time-domain system; then $T(z) = 1$). The PSF is completely characterized by the Rayleigh length which can be calibrated by measuring the response from a mirror at different positions with respect to the confocal gate and fitting to Eq. (18.40).

Use of this PSF in combination with the single backscatter model allowed determination of the attenuation coefficient of a dilute suspension of calibrated scattering particles. We showed that the range of validity of the single backscattering model and the PSF extends to weakly scattering media ($\mu_t < 6$ mm^{-1}) using a low NA (0.08) setup.

18.4.5 Clinical Implications

Conventionally, the differentiation between tissue structures in an OCT image is based on their reflectivity (gray scale value in the image) in combination with the (relative) position with respect to other layers and structures. The different structures can then be identified through comparison with histology. Sometimes specific anatomical landmarks appear, which eases this approach. Figure 18.12 shows a histology section (left) and corresponding OCT image [37] of normal esophageal wall tissue. The appearance of an elliptical structure, identified as a lymphoid follicle, improved accurate determination of the different tissue layers.

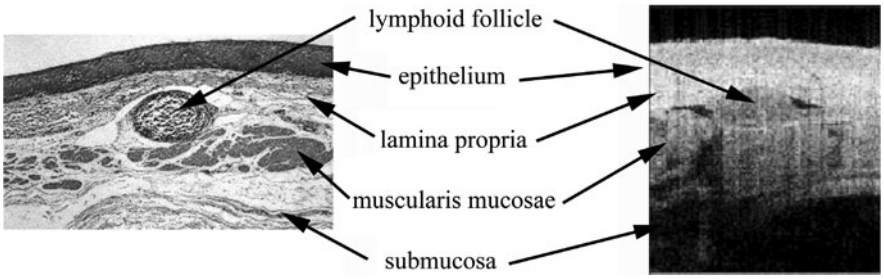


Fig. 18.12 Histologic section and OCT image of normal esophageal wall with an embedded ellipsoid structure. Histology (*left*): H&E staining, 40X magnification; OCT (*right*): 0.8 mm (depth) by 2.5 mm (width), fixed focus geometry, resolution 4 μm (*d*) by 5 μm (*w*), $\lambda_0 = 800$ nm

However, these structures may not always be present, the demarcation between different layers may not always be as clear, and the gray level of a certain tissue region may be influenced by its position relative to the focus of the imaging probe. Optical properties such as the attenuation coefficient μ_t are inherently different for different tissues and may therefore be used as markers in an OCT image.

OCT has shown to be valuable in the imaging of unstable atherosclerotic plaques because it is the only technique that allows for intravascular determination of plaque thickness. Unfortunately, differentiation between lipid rich and calcified lesions can currently only be made based on qualitative interpretation of the images. We determined the attenuation coefficient of human atherosclerotic arterial segments obtained at autopsy [38, 39] using a time-domain setup in fixed focus geometry. After imaging, the fitted regions of interest were classified using histology (20 lesions in 13 arterial segments). The results are shown in Fig. 18.13.

The attenuation coefficients of diffuse intimal thickened tissue (5.5 ± 1.2 mm^{-1}) and lipid-rich regions (3.2 ± 1.1 mm^{-1}) differed significantly from medial tissue (9.9 ± 1.8 mm^{-1}), calcifications (11.3 ± 4.9 mm^{-1}) and thrombi ($11.2 \pm$

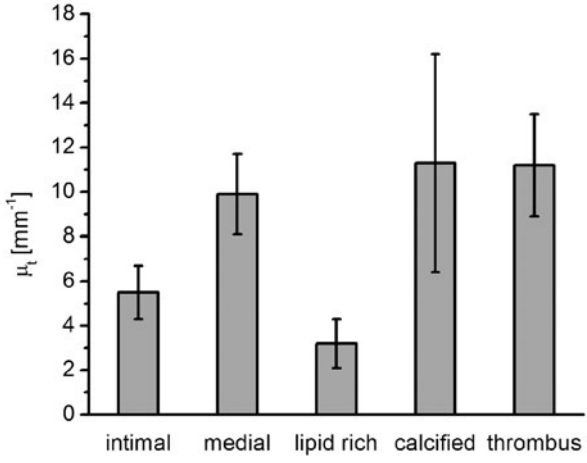


Fig. 18.13 Average values and standard deviations for the attenuation coefficient μ_t at 800 nm of intimal, medial, lipid-rich, and calcified tissue and thrombus

2.3 mm^{-1}), ($p < 0.01$). Even in this limited dataset, the clinically interesting lipid-rich regions are distinguishable from calcifications, which is not possible based on grey levels alone. Schmitt et al. were among the first to measure μ_t of rat aorta by OCT at 830 nm [40]. The larger reported $\mu_t = 14.9 \pm 2.3$ mm^{-1} can be attributed to the high content of elastin in their segments compared to our human arterial samples and differences in measurement procedure. Levitz et al. used the model of Eq. (18.39) to determine the scattering coefficient μ_s at 1300 nm of arterial tissue in a similar analysis [41].

Figure 18.14 presents a possible clinical implementation of these results. Panels A and B show OCT images of a calcified lesion ('*c*' in panel A) and a lipid rich lesion ('*l*' in panel B) respectively. Note that both the calcified and the lipid rich lesions appear as dark areas in the OCT image. The average A-scans of the regions of interest encompassed by the dotted rectangles in Fig. 18.14 A and B are depicted in 18.14 C and D, respectively, by the grey thin lines. The thick lines in these panels depict the individual fits using our model. Note the differences in μ_t (presented in $mm^{-1} \pm 95\%$ CI) for the different tissue types. A moving region of interest (~50 by 50 μm) was used to determine the local μ_t. In Fig. 18.14 E and F, color overlays are superimposed on the original OCT image presented in A and B, respectively. The scale indicates the local attenuation coefficient ranging from 0 to 15 mm^{-1}. In E, the calcified lesions of panel A are accentuated by thresholding the attenuation coefficients below 12 mm^{-1}; in F, the lipid rich lesion of B is accentuated by thresholding μ_t above 4 mm^{-1}.

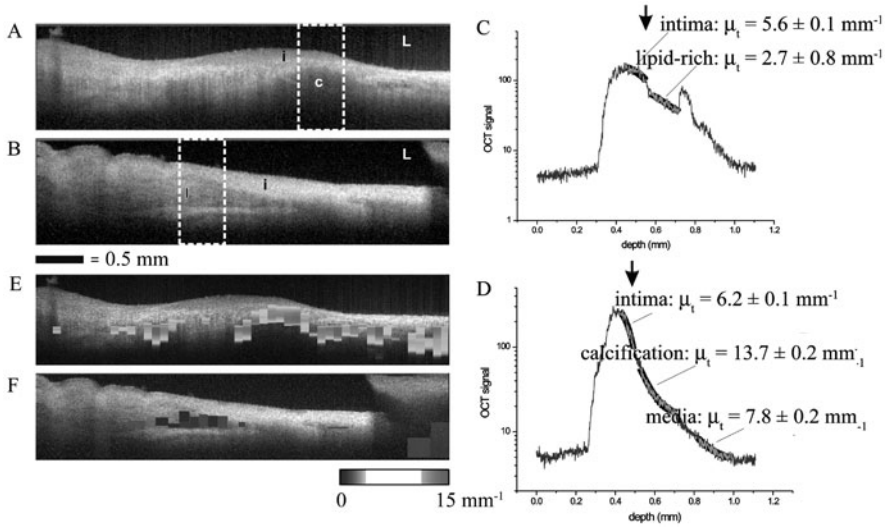


Fig. 18.14 OCT images of a calcified lesion (c in **A**) and a lipid rich lesion (l in **B**). (**C** and **D**) show average A-scans from the regions of interest in A and B respectively. Thick lines correspond to individual fits. (**E** and **F**) show false color overlays of μ_t on the image of panels 1 and 2. The color bar indicates the local attenuation coefficient ranging from 0 to 15 mm^{-1}. L: lumen, l: lipid rich lesion, c: calcified lesion, i: diffuse intimal thickening

References

1. Huang D, Swanson EA, Lin CP, Schuman JS, Stinson WG, Chang E, Hee MR, Flotte T, Gregory K, Puliafito CA, and Fujimoto JG. Optical coherence tomography. *Science*, 254:1178–1181 (1991).
2. Izatt JA, Hee MR, Owen GM, Swanson EA, and Fujimoto JG. Optical coherence microscopy in scattering media. *Opt. Lett.*, 19:590–592 (1994).
3. Schmitt JM, Knuettel A, and Yadlowsky M. Confocal microscopy in turbid media. *J. Opt. Soc. Am. A*, 11(8):2226–2235 (1994).
4. van Velthoven MEJ, Faber DJ, Verbraak FD, van Leeuwen TG, and de Smet MD. Recent developments in optical coherence tomography for imaging the retina. *Progr. Retinal Eye Res*., 26(1):57 (2007).
5. Nassif N, Cense B, Park BH, Yun SH, Chen TC, Bouma BE, Tearney GJ, and de Boer JF. In vivo human retinal imaging by ultrahigh-speed spectral domain optical coherence tomography. *Opt. Lett*., 29:480–482 (2004).
6. Izatt JA, Kulkarni MD, Yazdanfar S, Barton JK, and Welch AJ. In vivo bi-directional color Doppler flow imaging of picoliter blood volumes using optical coherence Tomography. *Opt. Lett*., 22:1439 (1997).
7. Aalders MC, Triesscheijn M, Ruevekamp M, de Bruin DM, Baas P, Faber DJ, and Stewart F. (Doppler) optical coherence tomography for monitoring the effect of photodynamic therapy on tissue morpology and perfusion. *J. Biomed. Opt*., 11(4):044011 (2006).
8. Knuettel A, Bonev S, and Knaak W. New method for evaluation of *in vivo* scattering and refractive index properties obtained with optical coherence tomography. *J. Biomed. Opt*., 9:265–273 (2004).
9. de Boer JF, Milner TE, van Gemert MJC, and Nelson JS. Two-dimensional birefringence imaging in biological tissue by polarization-sensitive optical coherence Tomography. *Opt. Lett*., 22:934 (1997).
10. Schmitt JM, Xiang SH, and Yung KM. Differential absorption imaging with optical coherence tomography. *J. Opt. Soc. Am. A*, 15:2288 (1998).
11. Faber DJ, Mik EG, Aalders MCG, and van Leeuwen TG. Light absorption of (oxy-hemoglobin assessed by spectroscopic optical coherence tomography. *Opt. Lett*., 28(16):1436 (2003).
12. Kholodnykh AI, Petrova IY, Motamedi M, and Esenaliev RO. Accurate measurement of total attenuation coefficient of thin tissue with optical coherence tomography. *IEEE J. Sel. Top. Quantum Electron*., 9:210–221 (2003).
13. Faber DJ, van der Meer FJ, Aalders MCG, and van Leeuwen TG. Quantitative measurements of attenuation coefficients of weakly scattering media using optical coherence tomography. *Opt. Express*, 12(19):4353 (2004).
14. Faber FJ, Aalders MCG, Mik EG, Hooper BA, van Gemert MJC, and van Leeuwen TG. Oxygen saturation-dependent absorption and scattering of blood. *Phys. Rev. Lett*., 93(2):028102-1 (2004).
15. Faber DJ, Mik EG, Aalders MCG, and van Leeuwen TG. Toward assessment of blood oxygen saturation by spectroscopic optical coherence tomography. *Opt. Lett*., 30(9):1015 (2005).
16. Sorin RV and Baney DM. A simple intensity noise reduction technique for optical low coherence reflectrometry. *IEEE Photonics Techn. Lett*., 4:1404 (1992).
17. Rollins AM and Izatt JA. Optimal interferometer designs for optical coherence tomography. *Opt. Lett*., 24:1484 (1999).
18. Podoleanou AG. Unbalanced versus balanced operation in an optical coherence tomography system. *Appl. Opt*., 39:173 (2000).
19. Leitgeb R, Hitzenberger CK, and Fercher AF. Performance of fourier domain vs. time domain optical coherence tomography. *Opt. Express*, 11(8):889–894 (2003).
20. de Boer JF, Cense B, Park BH, Pierce MC, Tearney GJ, and Bouma BE. Improved signal-to-noise ratio in spectral-domain compared with time-domain optical coherence tomography. *Opt. Lett*., 28(21):2067 (2003).

21. Choma MA, Sarunic MV, Yang CH, and Izatt JA. Sensitivity advantage of swept source and Fourier domain optical coherence tomography. *Opt. Express*, 11:2183 (2003).
22. Kodach V, Faber DJ, and van Leeuwen TG. Wavelength swept Ti:sapphire laser. *Opt. Commun*., 281(19):4975–4978 (2008).
23. Yun S, Tearney G, de Boer JF, Iftimia N, and Bouma B. High-speed optical frequency-domain imaging. *Opt. Express*, 11(22):2953 (2003).
24. van Leeuwen TG, Faber DJ, and Aalders MC. Measurement of the axial point spread function in scattering media using single-mode fiber-based optical coherence tomography. *IEEE J. Sel. Top. Quant. Electron*., 9(2):227 (2003).
25. Yuan S and Riza NA. General formula for coupling-loss characterization of single-mode fiber collimators by use of gradient-index rod lenses. *Appl. Opt*., 38(15):3214–3222 (1999).
26. de Grauw CJ, Sijtsema NM, Otto C, and Greve J. Axial resolution of confocal Raman microscopes: Gaussian beam theory and practice. *J. Microsc.*, 188:273–279 (1997).
27. van der Hulst HC. *Light scattering by small particles*. Dover, New York (1957).
28. Thrane L, Yura HT, and Andersen PE. Analysis of optical coherence tomography systems based on the extended HuygensFresnel principle. *J. Opt. Soc. Am. A*, 17:484–490 (2000).
29. Yura HT, Thrane L, and Andersen PE. Closed-from solution for the Wigner phase-space distribution function for diffuse reflection and small-angle scattering in a random medium. *J. Opt. Soc. Am. A*, 17(12):2464–2474 (2000).
30. Yadlowsky MJ, Schmitt JM, and Bonner RF. Multiple scattering in optical coherence microscopy. *Appl. Opt*., 34(25):5699–5707 (1995).
31. Pan Y, Birngruber R, and Engelhard R. Constrast limits of coherence-gated imaging in scattering media. *Appl. Opt*., 36(13):2979–2983 (1997).
32. Bizheva KK, Siegel AM, and Boas DA. Path-length-resolved dynamic light scattering in highly scattering random media: The transition to diffusing wave spectroscopy. *Phys. Rev. E*, 58(6):7664–7667 (1998).
33. Thrane L, Yura HT, and Andersen PE. Analysis of optical coherence tomography systems based on the extended HuygensFresnel principle. *J. Opt. Soc. Am. A*, 17:484–490 (2000).
34. Wax A, Yang C, Dasari RR, and Feld M. Path-length-resolved dynamic light scattering: Modeling the transition from single to diffusive scattering. *Appl. Opt*., 40(24):4222–4227 (2001).
35. Karamata B, Laubscher M, Leutenegger M, Bourquin S, Lasser T, and Lambelet P. Multiple scattering in optical coherence tomography. I. Investigation and modeling. *J. Opt. Soc. Am. A*, 22(7):1369 (2005).
36. Turchin IV, Sergeeva EA, Dolin LS, Kamensky VA, Shakova NM, and Richards-Kortum R. Novel algorithm of processing optical coherence tomography images for differentiation of biological tissue pathologies. *J. Biomed. Opt.*, 10(6):064024 (2005).
37. Çilesiz I, Fockens P, Kerindongo R, Faber DJ, Tytgat ten Kate F, and van Leeuwen TG. Comparative optical coherence tomography imaging of human esophagus: How accurate is localization of the *muscularis mucosae*. *Gastrointest. Endosc.*, 56(6):852–857 (2002).
38. van der Meer FJ, Faber DJ, Baraznji Sassoon DM, Aalders MC, Pasterkamp G, and van Leeuwen TG. Localized measurement of optical attenuation coefficients of atherosclerotic plaque constituents by quantitative optical coherence tomography. *IEEE Trans. Med. Imaging*, 24(10):1369–1376 (2005).
39. van der Meer FJ, Faber DJ, Peree J, Pasterkamp G, Baraznji Sassoon DMB, and van Leeuwen TG. Quantitative optical coherence tomography of arterial wall components. *Lasers Med. Sci.*, 20(1):45–51 (2005).
40. Schmitt JM, Knuttel A, Yadlowsky M, and Eckhaus MA. Optical-coherence tomography of a dense tissue: statistics of attenuation and backscattering. *Phys. Med. Biol*., 39:1705–1720 (1994).
41. Levitz D, Thrane L, Frosz MH, Andersen PE, Andersen CB, Andersson-Engels S, Valanciunaite J, Swartling J, and Hansen PR. Determination of optical scattering properties of highly-scattering media in optical coherence tomography images. *Opt. Express*, 12:249–259 (2004).

Chapter 19
Photoacoustic Tomography

Lihong V. Wang

19.1 Introduction

Photoacoustic tomography (PAT) refers to imaging that is based on the photoacoustic effect. Although the photoacoustic effect as a physical phenomenon was first reported on by Alexander Graham Bell in 1880 [1], PAT as an imaging technology was developed only after the advent of ultrasonic transducers, computers, and lasers [2–31]. A review on biomedical photoacoustics is available [32]. The motivation for PAT is to combine optical-absorption contrast with ultrasonic spatial resolution for deep imaging in the optical quasi-diffusive or diffusive regime. In PAT, the tissue is irradiated by usually a short-pulsed laser beam to achieve a thermal and acoustic impulse response (Fig. 19.1). Locally absorbed light is converted into heat, which is further converted to a pressure rise via thermo-elastic expansion. The

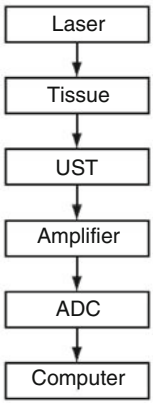


Fig. 19.1 Block diagram of the basic scheme for photoacoustic tomography. ADC: analog-to-digital converter; UST: ultrasonic transducer

L.V. Wang (✉)
Department of Biomedical Engineering, Washington University, St. Louis, MO 63130-4899, USA
e-mail: lhwang@seas.wustl.edu

A.J. Welch, M.J.C. van Gemert (eds.), *Optical-Thermal Response of Laser-Irradiated Tissue*, 2nd ed., DOI 10.1007/978-90-481-8831-4_19,

initial pressure rise – determined by the local optical absorption coefficient (μ_a), fluence (ψ) and other thermal and mechanical properties – propagates as an ultrasonic wave, which is referred to as a photoacoustic wave. The photoacoustic wave is detected by an ultrasonic transducer, producing an electric signal. The signal is then amplified, digitized and transferred to a computer. To form an image, we need to scan a single-element ultrasonic transducer; alternatively, an ultrasonic array can be used to acquire data in parallel.

The image contrast of PAT is based on optical absorption in the photoacoustic excitation phase. Selective optical absorption is associated with molecules such as oxygenated and deoxygenated hemoglobin and melanin. Concentrations of multiple chromophores whose spectra of absorption coefficient μ_a are different can be quantified simultaneously by varying the wavelength of the irradiating laser. Such quantification of oxygenated and deoxygenated hemoglobin, for example, can be subsequently used for functional imaging.

The spatial resolution of PAT is derived from ultrasonic detection in the photoacoustic emission phase. Due to strong optical scattering, pure optical imaging in biological tissue has either shallow imaging depth or low spatial resolution. Although light scattering broadens and attenuates light reaching an embedded object, the resulting ultrasonic signal provides better resolution than the backscattered optical signal in the optical quasi-diffusive or diffusive regime because ultrasonic scattering is 2–3 orders of magnitude weaker than optical scattering. The image resolution, as well as the maximum imaging depth, is scalable with the ultrasonic frequency within the reach of diffuse photons. Specifically, as the ultrasonic center frequency and bandwidth increase, the spatial resolution improves at the expense of imaging depth because ultrasonic attenuation increases with frequency. In addition, PAT provides images devoid of speckle artifacts, which are conspicuous in both OCT and ultrasonic images.

19.2 Initial Photoacoustic Pressure

Upon short laser pulse excitation, local fractional volume expansion dV/V of the heated tissue at position $\mathbf{r}$ can be expressed as

$$dV\big/V = -\kappa p\,(\mathbf{r}) + \beta T\,(\mathbf{r}) \tag{19.1}$$

Here, κ denotes the isothermal compressibility ($\sim 5 \times 10^{-10}$ Pa^{-1} for water or soft tissue); β denotes the thermal coefficient of volume expansion ($\sim 4 \times 10^{-4}$ K^{-1} for muscle); p and T denote the changes in pressure (Pa) and temperature (K), respectively. The isothermal compressibility κ can be expressed as

$$\kappa = \frac{C_p}{\rho v_s^2 C_v} \tag{19.2}$$

Here, ρ denotes the mass density ($\sim$1000 kg/m^3 for water and soft tissue); v_s denotes the speed of sound ($\sim$1480 m/s in water); C_p and C_v ($\sim$4000 J/(kg K) for muscle) denote the specific heat capacities at constant pressure and volume, respectively. It is important to distinguish between C_p and C_v for gasses but not for tissue.

If the laser pulse duration t_0 is less than the acoustic confinement time (which is less than the thermal confinement time), then an acoustic impulse response is produced. The necessary stress confinement requires

$$t_0 < \frac{\ell}{v_s} < \frac{\ell^2}{4\alpha} \tag{19.3}$$

where ℓ is the characteristic length of heat heterogeneity (the dimension of the optically absorbing target of interest or the decay constant of the optical energy deposition, whichever is smaller), and α is the thermal diffusivity.

For such a short laser pulse, the fractional volume expansion is negligible and the local pressure rise p_0 produced immediately after the laser excitation can be derived from Eq. (19.1) [33]:

$$p_0(\mathbf{r}) = \frac{\beta T(\mathbf{r})}{\kappa} \tag{19.4}$$

It can be shown that each mK temperature rise yields approximately an 8-mbar (or 800 Pa) pressure rise. If we assume that all absorbed optical energy is converted into heat, and nonthermal relaxation such as fluorescence is negligible, the temperature rise generated by the short laser pulse is

$$T(\mathbf{r}) = \frac{W_a(\mathbf{r})}{\rho C_v} \tag{19.5}$$

where W_a is the specific or volumetric optical absorption (J/cm^3, optical energy deposition). Since the local rate of heat generation S is the product of the local optical absorption coefficient at the laser wavelength μ_a and the fluence rate ϕ (W/cm^2, see Chapter 3), $S(\mathbf{r}, t)\, dt = \mu_a(\mathbf{r})\, \phi(\mathbf{r}, t)$, we have

$$W_a(\mathbf{r}) = \int_{-\infty}^{+\infty} S(\mathbf{r}, t)\, dt = \mu_a(\mathbf{r})\, \psi(\mathbf{r}) \tag{19.6}$$

where ψ is the optical fluence (J/cm^2) at the target. Although W_a and the radiant energy density W share the same unit, they have different physical meanings. In fact, we have $\phi = Wc$, where c is the speed of light in the medium.

From Eqs. (19.4) and (19.5), we have

$$p_0(\mathbf{r}) = \frac{\beta}{\kappa \rho C_v} W_a(\mathbf{r}) \tag{19.7}$$

We define the Grueneisen parameter (dimensionless) as

$$\Gamma = \frac{\beta}{\kappa \rho C_v} = \frac{\beta v_s^2}{C_p} \tag{19.8}$$

For water and diluted aqueous solutions, Γ can be estimated by the following empirical formula:

$$\Gamma_w(T_0) = 0.0043 + 0.0053 T_0 \tag{19.9}$$

where T_0 is the temperature in degrees Celsius. At body temperature, $\Gamma_w(37°\text{C}) = 0.20$. From Eqs. (19.8), (19.7) becomes

$$p_0(\mathbf{r}) = \Gamma W_a(\mathbf{r}) \tag{19.10}$$

19.3 Photoacoustic Equation

The photoacoustic wave generation and propagation in an inviscid medium is described by the following general photoacoustic equation:

$$\left(\nabla^2 - \frac{1}{v_s^2}\frac{\partial^2}{\partial t^2}\right) p(\mathbf{r}, t) = -\frac{\beta}{\kappa v_s^2}\frac{\partial^2 T(\mathbf{r}, t)}{\partial t^2} \tag{19.11}$$

where $p(\mathbf{r}, t)$ denotes the acoustic pressure at location $\mathbf{r}$ and time t, and T denotes the temperature rise. The left-hand side of this equation describes the wave propagation, whereas the right-hand side represents the source term.

For a short laser pulse, $\rho C_v \frac{\partial T(\mathbf{r},t)}{\partial t} = S(\mathbf{r}, t)$ holds. As a result, we obtain the following less general photoacoustic equation:

$$\left(\nabla^2 - \frac{1}{v_s^2}\frac{\partial^2}{\partial t^2}\right) p(\mathbf{r}, t) = -\frac{\beta}{C_p}\frac{\partial S(\mathbf{r}, t)}{\partial t} \tag{19.12}$$

Because the source term is related to the first time derivative of S, time-invariant heating does not produce a pressure wave; only time-variant heating does.

19.4 Forward Solution

The general photoacoustic equation shown in Eq. (19.11) can be solved by the Green's function approach [34]. The Green's function is defined here as the response to a spatial and temporal impulse source term:

$$\left(\nabla^2 - \frac{1}{v_s^2}\frac{\partial^2}{\partial t^2}\right) G\left(\mathbf{r}, t; \mathbf{r}', t'\right) = -\delta\left(\mathbf{r} - \mathbf{r}'\right)\delta\left(t - t'\right) \tag{19.13}$$

where $\mathbf{r}'$ and t' denote the source location and time, respectively. In infinite space, where no boundary exists, the Green's function is given by

$$G\left(\mathbf{r},t;\mathbf{r}',t'\right)=\frac{\delta\left(t-t'-\frac{\left|\mathbf{r}-\mathbf{r}'\right|}{v_s}\right)}{4\pi\left|\mathbf{r}-\mathbf{r}'\right|} \tag{19.14}$$

which represents an impulse diverging spherical wave. The following reciprocity relation holds:

$$G\left(\mathbf{r},t;\mathbf{r}',t'\right)=G\left(\mathbf{r}',-t';\mathbf{r},-t\right) \tag{19.15}$$

To see this relationship more clearly, one observes $G\left(\mathbf{r},t;\mathbf{r}',0\right)=G\left(\mathbf{r}',0;\mathbf{r},-t\right)$ by setting $t'=0$.

Applying the Green's function approach to Eq. (19.11) yields

$$p\left(\mathbf{r},t\right)=\int_{-\infty}^{t+}dt'\int d\mathbf{r}'G\left(\mathbf{r},t;\mathbf{r}',t'\right)\frac{\beta}{\kappa v_s^2}\frac{\partial^2 T\left(\mathbf{r}',t'\right)}{\partial t'^2} \tag{19.16}$$

which represents the pressure in response to an arbitrary source. Substituting Eq. (19.14) into Eq. (19.16) leads to

$$p\left(\mathbf{r},t\right)=\frac{\beta}{4\pi\kappa v_s^2}\int d\mathbf{r}'\frac{1}{\left|\mathbf{r}-\mathbf{r}'\right|}\left.\frac{\partial^2 T\left(\mathbf{r}',t'\right)}{\partial t'^2}\right|_{t'=t-\frac{\left|\mathbf{r}-\mathbf{r}'\right|}{v_s}} \tag{19.17}$$

In thermal confinement, substituting $\rho C_v\frac{\partial T(\mathbf{r},t)}{\partial t}=S\left(\mathbf{r},t\right)$ into Eq. (19.17) yields

$$p\left(\mathbf{r},t\right)=\frac{\beta}{4\pi C_p}\int d\mathbf{r}'\frac{1}{\left|\mathbf{r}-\mathbf{r}'\right|}\left.\frac{\partial S\left(\mathbf{r}',t'\right)}{\partial t'}\right|_{t'=t-\frac{\left|\mathbf{r}-\mathbf{r}'\right|}{v_s}} \tag{19.18}$$

or

$$p\left(\mathbf{r},t\right)=\frac{\beta}{4\pi C_p}\frac{\partial}{\partial t}\int d\mathbf{r}'\frac{1}{\left|\mathbf{r}-\mathbf{r}'\right|}S\left(\mathbf{r}',t-\frac{\left|\mathbf{r}-\mathbf{r}'\right|}{v_s}\right) \tag{19.19}$$

If the heating function can be decomposed as $S(\mathbf{r}',t')=W_a(\mathbf{r}')S_t\left(t'\right)$, where $\int S_t\left(t'\right)dt'=1$, Eq. (19.19) can be further simplified to

$$p\left(\mathbf{r},t\right)=\frac{\beta}{4\pi C_p}\frac{\partial}{\partial t}\int d\mathbf{r}'\frac{W_a\left(\mathbf{r}'\right)}{\left|\mathbf{r}-\mathbf{r}'\right|}S_t\left(t-\frac{\left|\mathbf{r}-\mathbf{r}'\right|}{v_s}\right) \tag{19.20}$$

If condition in Eq. (19.3) holds, we have approximately $S_t(t')=\delta(t')$. In this case, Eq. (19.20) becomes

$$p(\mathbf{r},t) = \frac{\beta}{4\pi C_p}\frac{\partial}{\partial t}\int d\mathbf{r}'\frac{W_a(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\delta\left(t-\frac{|\mathbf{r}-\mathbf{r}'|}{v_s}\right) \tag{19.21}$$

or

$$p(\mathbf{r},t) = \frac{\partial}{\partial t}\left[\frac{\beta}{4\pi C_p}\frac{1}{v_s t}\int d\mathbf{r}' W_a(\mathbf{r}')\,\delta\left(t-\frac{|\mathbf{r}-\mathbf{r}'|}{v_s}\right)\right] \tag{19.22}$$

where the quantity within the square brackets is the step-heating response of an arbitrary absorbing object and its time differentiation yields the delta-heating response. From Eqs. (19.8) and (19.10), we rewrite Eq. (19.22) as

$$p(\mathbf{r},t) = \frac{1}{4\pi v_s^2}\frac{\partial}{\partial t}\left[\frac{1}{v_s t}\int d\mathbf{r}' p_0(\mathbf{r}')\,\delta\left(t-\frac{|\mathbf{r}-\mathbf{r}'|}{v_s}\right)\right] \tag{19.23}$$

For a spherical object of radius R_s, Eq. (19.23) can be used to derive the delta-heating response analytically [31]. When an ultrasonic detector is placed $3R_s$ away from the center of the sphere, the detected pressure is shown in Fig. 19.2. Upon delta heating, an initial pressure p_0 – assumed to be constant across the entire heated sphere – is generated. This initial pressure is divided into two equal parts, each initiating a spherical wave. One travels outward as a diverging spherical compression wave, yielding the first-arriving positive pressure in the plot. The other travels inward as a converging spherical compression wave. When reaching the center of the heated spherical object, the converging spherical wave becomes a diverging spherical rarefaction wave, yielding the late-arriving negative pressure in the plot.

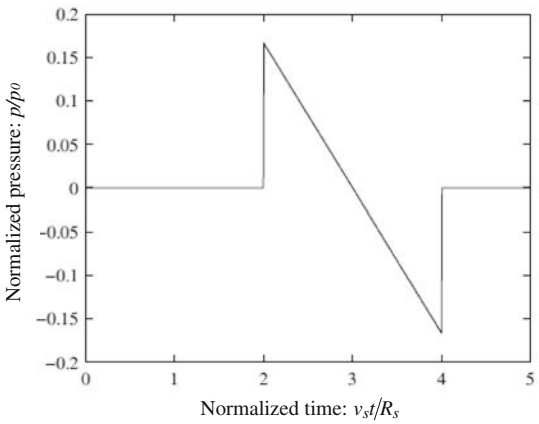


Fig. 19.2 Bipolar (positive followed by negative) pressure profile from a heated sphere versus time

19.5 Dark-Field Confocal Photoacoustic Microscopy

In this section, we introduce the dark-field confocal photoacoustic microscopy (PAM) (Fig. 19.3) [35, 36]. In conventional dark-field transmission optical microscopy, an opaque disc is placed between the light source and the condenser lens so that ballistic light is rejected; as a result, only non-ballistic light – which is scattered by the sample – is detected. In dark-field PAM, the excitation laser beam has a donut-shaped cross section; therefore, the photoacoustic signal from the tissue surface in the field of view is minimized. PAM can image optical-absorption contrast beyond the existing depth limit for high-resolution optical imaging; it can achieve a ratio of maximum imaging depth to depth resolution (un-interpolated pixel count) greater than 100, which is defined here as high relative spatial resolution.

Acoustic coupling requires contact measurement of the sample. The components within the dashed-box in the photograph (Fig. 19.3) are translated in a water bath. A window at the bottom of the water container is sealed with an optically and ultrasonically transparent disposable polyethylene membrane (thickness: 0.044 mm). After commercial ultrasound gel is applied to the region of interest on the sample for acoustic coupling, the sample is placed between the water container and the sample supporter for imaging. Note that sample here can also refer to live subjects.

Image formation starts from the laser. For the generation of photoacoustic waves, 6.5-ns laser pulses from a tunable dye laser that is pumped by an Nd:YAG laser are used. Such short laser pulses are able to generate photoacoustic waves of a >100-MHz bandwidth (approximately the reciprocal of laser pulse width) at the target. Laser light at a designated wavelength is delivered through an optical fiber to the PAM scanner. For hemoglobin imaging using PAM, the laser wavelength is typically set in the Q-band of hemoglobin (∼560 nm) to achieve good signal-to-noise ratio. The energy of each laser pulse is recorded by a photodiode for shot-to-shot

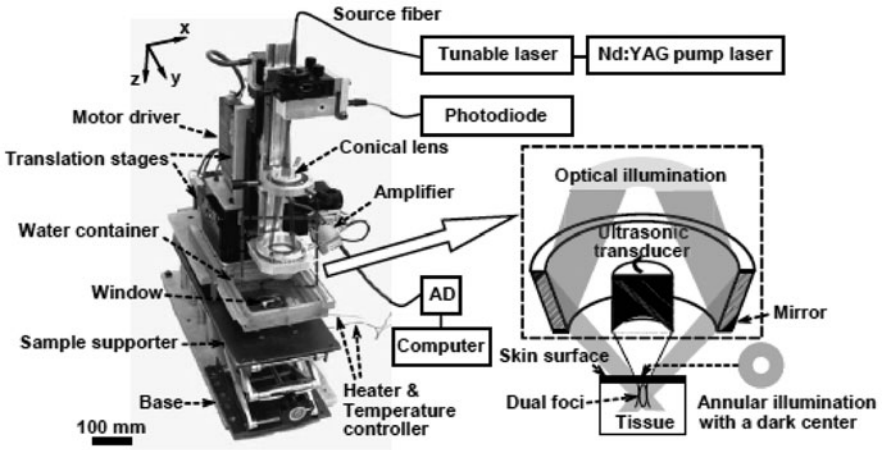


Fig. 19.3 Photoacoustic microscopy (PAM) system. (Reprinted with permission from [36])

calibration. The laser beam from the fiber passes through a conical lens to form a ring-shaped illumination pattern. It is then weakly focused into the tissue with the focal region coaxially overlapping the ultrasonic focus inside the tissue. The optical illumination on the skin surface is donut-shaped with a dark center. In an optically clear medium, the optical focus is measured to be 2 mm in diameter, which is much wider than the ultrasonic focus. The time-resolved photoacoustic wave is recorded at each location of the ultrasonic transducer for 2 μs and, subsequently, converted into a one-dimensional (1D) depth-resolved image (A-line) based on the sound velocity in soft tissue (1.54 mm/μs). Then, raster scanning of the dual optical-ultrasonic foci in the horizontal (*x*–*y*) plane with a step size of 50 μm produces a three-dimensional (3D) image. No signal averaging is performed in data acquisition.

Spatial resolution is an important index for imaging. The lateral resolution of PAM is determined by the focal diameter of the ultrasonic transducer at the center frequency [37], and the axial resolution is inversely related to the bandwidth of the ultrasonic transducer.

The lateral resolution is usually computed by $0.61(\lambda_0/NA)$, where λ_0 is the center acoustic wavelength and *NA* is the numerical aperture of the ultrasonic transducer. Therefore, achieving high spatial resolution requires the use of an ultrasonic transducer with a large *NA*, a high center frequency, and a wide bandwidth. At center frequencies greater than ~10 MHz, it is the penetration depth of ultrasound, rather than the penetration depth of diffuse light, that limits the maximum imaging depth owing to the strong frequency-dependent acoustic attenuation. As a result, both the spatial resolution and the maximum imaging depth of PAM are scalable with ultrasonic parameters within the reach of excitation photons. In the PAM reported here, the ultrasonic detector (V214-BB-RM, Panametrics, WA; element diameter: 6 mm) has a 50-MHz center frequency and a 35-MHz nominal bandwidth. A spherically focusing lens provides an *NA* of 0.44, a focal length of 6.7 mm, and a focal zone of 0.3 mm.

The actual resolution of the PAM system is quantified using a well-controlled sample. A cross section of a carbon fiber (diameter: 6 μm) immersed in an optically scattering medium is imaged (Fig. 19.4). The 1D line spread functions (LSF) along the *z* (axial) and the *x* (lateral) axes shown in Figs. 19.4(a) and (c), respectively, are extracted from the 2D cross-sectional image. To quantify the spatial resolution in each direction, we shift the 1D LSF by a distance and added it to the original LSF. The smallest shifted distance that still allows resolving the two peaks with a ~5% dip is considered as the resolution along that direction. This PAM system is found to have an axial resolution of ~15 μm and a lateral resolution of ~45 μm. An imaging depth of more than 3 mm in biological tissue has been reached in animals (Fig. 19.5). Therefore, this PAM system has a depth-to-resolution ratio of 200 and is considered to possess high relative spatial resolution.

The major optical absorbers in biological tissue in the visible spectral region include deoxyhemoglobin (HbR), oxyhemoglobin (HbO_2), and melanin. Consequently, PAM, functioning as a sensitive blood detecting technology with high contrast and specificity, is suitable for imaging the volumetric morphology of the subcutaneous microvasculature. At the 584-nm isosbestic optical wavelength,

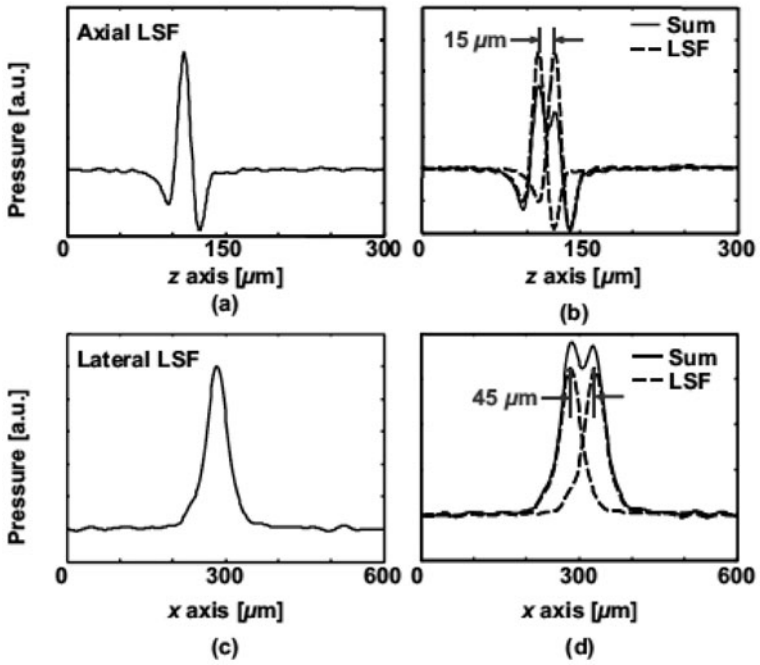


Fig. 19.4 Spatial resolution of the 50-MHz PAM system. (**a**) Axial LSF of position along the acoustic (*z*) axis. (**b**) Shift-and-sum of the axial LSF. (**c**) Lateral LSF of position perpendicular to the acoustic (*z*) axis. (**d**) Shift-and-sum of the lateral LSF

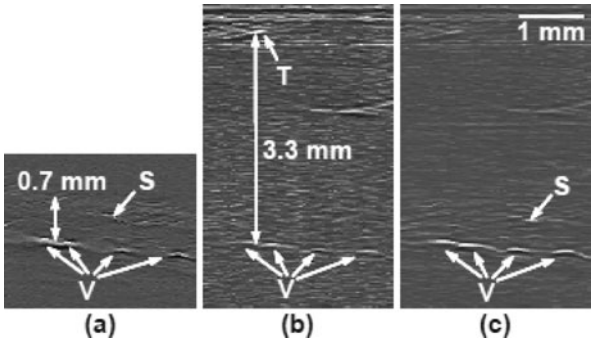


Fig. 19.5 Penetration of PAM in tissue. (**a**) Image of subcutaneous vessels in a Sprague-Dawley rat acquired in vivo without signal averaging. Vessels V are ~0.7 mm below skin surface S. (**b**) Image of the same vessels shown in *a* acquired in vivo without signal averaging when the skin is covered by a chicken breast tissue slab. The vessels are now 3.3 mm below tissue surface T. (**c**) Same as *b* except that 40-time signal averaging is taken. The SNR is improved as anticipated. S: skin surface; V: blood vessels; T: top surface of chicken tissue

where the two forms of hemoglobin have an identical molar extinction coefficient, the contrast of PAM is sensitive to the total hemoglobin concentration but insensitive to the hemoglobin oxygenation.

By extending single-wavelength imaging to dual-wavelength imaging, PAM can separate the contributions from two chromophores based on their different molar absorption spectra. A good example is melanoma imaging. At the 584-nm optical wavelength, PAM can image the morphological relationship between a melanoma tumor and the surrounding blood vessels in the *x*–*y* plane because both melanin and hemoglobin have comparably strong absorption at this wavelength. However, visible light at this wavelength is unsuitable for measuring the tumor thickness because it cannot easily penetrate through the melanin-rich tumor. To overcome this limitation, we form another image using NIR light at the 764-nm wavelength, where light can infiltrate the tumor because of both the decreased optical absorption of melanin and the minimal absorption of blood. The combination of the two images from these two spectral regions reveals the 3D morphology of both the melanoma and the surrounding vasculature, where some parallel arterioles and venules are evident (Fig. 19.6). In Fig. 19.6b, microvessels with diameters of less than a single pixel (50 μm), presumably resulting from angiogenesis, are observed clearly surrounding the tumor. At the 584-nm wavelength, the average ratio of the vessel to the background in photoacoustic signal amplitude is 13±0.89, and the average ratio of the melanoma tumor to the surrounding blood vessels is 0.92±0.02. However, at the 764-nm wavelength, the average ratio of the melanoma tumor to the blood vessels is 29±3, and the average ratio of the melanoma tumor to the background is as high as 68±5. Such high optical contrast and specificity are due to the minimal background absorption and the strong absorption of both hemoglobin and melanin. The average contrast-to-noise ratio for the melanoma is 67 and 64 dB at the 584-nm and 764-nm wavelengths, respectively.

In addition to the functional imaging of total hemoglobin concentration, PAM is capable of providing functional imaging of oxygen saturation of hemoglobin (SO_2) using multi-wavelength measurements [38]. Although the optical measurement of SO_2 has been investigated for decades, no non-invasive in vivo SO_2 imaging of single blood vessels has succeeded owing to either the lack of adequate spatial resolution or low sensitivity. PAM quantifies SO_2 in the same manner as do the traditional optical techniques. HbR and HbO_2 are treated as the dominant absorbing compounds, and their relative concentrations are calculated based on their molar extinction spectra from spectral measurements of the total absorption coefficient [39]. However, PAM can pinpoint individual blood vessels with high resolution and evaluate their individual SO_2 levels.

PAM is used to image the static SO_2 of subcutaneous vessels in normoxia, where the measured photoacoustic amplitudes from blood vessels are compensated for by the average wavelength-dependent optical attenuation in the dermis. The imaged SO_2 measured around 0.97±0.02 in arterial blood and 0.77±0.04 in venous blood (Fig. 19.7). While other noninvasive techniques measure volume-averaged SO_2, PAM provides a vessel-by-vessel mapping of SO_2 (Fig. 19.7b).

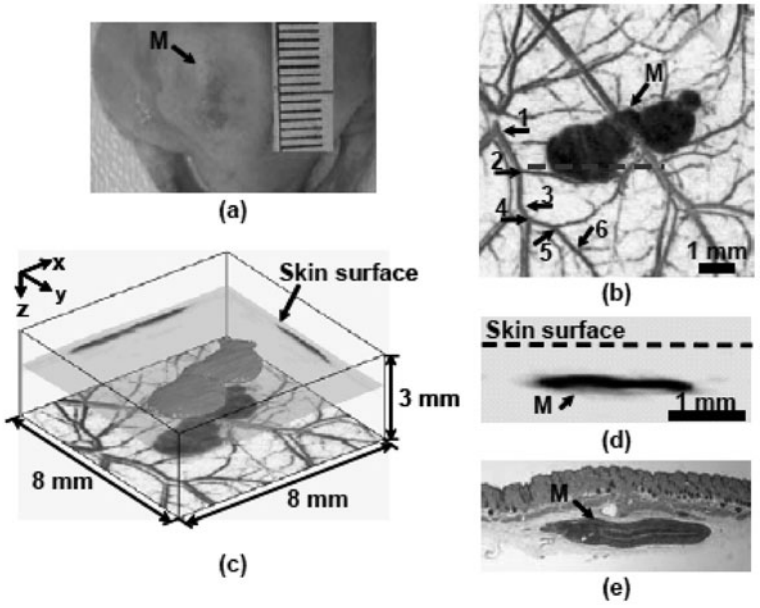


Fig. 19.6 Images of a subcutaneously inoculated B16-melanoma in a nude mouse acquired in vivo by PAM at 584 and 764 nm, respectively. (**a**) Photograph of a melanoma. (**b**) Composite of the two maximum amplitude projections (MAP) images projected along the *z* axis. Six orders of vessel branching (No. 1–6) can be observed. (**c**) 3D rendering of the melanoma acquired at the 764-nm wavelength. Two MAP images at this wavelength projected along the *x* and *y* axes are shown on the side walls. The composite image in panel *b* is redrawn at the *bottom*. The *top* of the tumor is 0.32 mm below the skin surface, and the thickness of the tumor is 0.3 mm. (**d**) A close-up B-scan image of the melanoma parallel with the *z*-*x* plane at the location marked with a dashed line in panel *b*. (**e**) HE-stained section at the same marked location. M: melanoma. (Reprinted with permission from [36])

PAM is safe for human subjects since it delivers a transient fluence of only ~6 mJ/cm^2 at the optical focus even if the sample is transparent, which is well within the ANSI safety standard (20 mJ/cm^2) in the visible spectrum (400–700 nm) [41]. If a point on the skin is exposed to laser light for more than 10 s, ANSI stipulates that the mean irradiance should not exceed 200 mW/cm^2, which results in a maximum permissible repetition rate of ~30 Hz here. Since PAM works in raster scanning mode, the exposure duration of any given point on the skin surface is typically less than 10 s. In this case, ANSI requires that the maximum permissible exposure is limited by $1100t_e^{1/4}$ in mJ/cm^2, where t_e denotes the exposure duration in seconds. This limit translates into $E \times \sqrt[4]{\mathrm{PRF}} \leq 2.75 \times 10^2 \pi d^{5/4} \sigma^{3/4}$, where E denotes the pulse energy in mJ, PRF denotes the pulse repetition frequency in Hz, d denotes the diameter of the illumination area at the ultrasonic focus in cm, and σ denotes the scanning step size in cm [42]. From the parameters of the current PAM, PRF can be as high as 13 kHz.

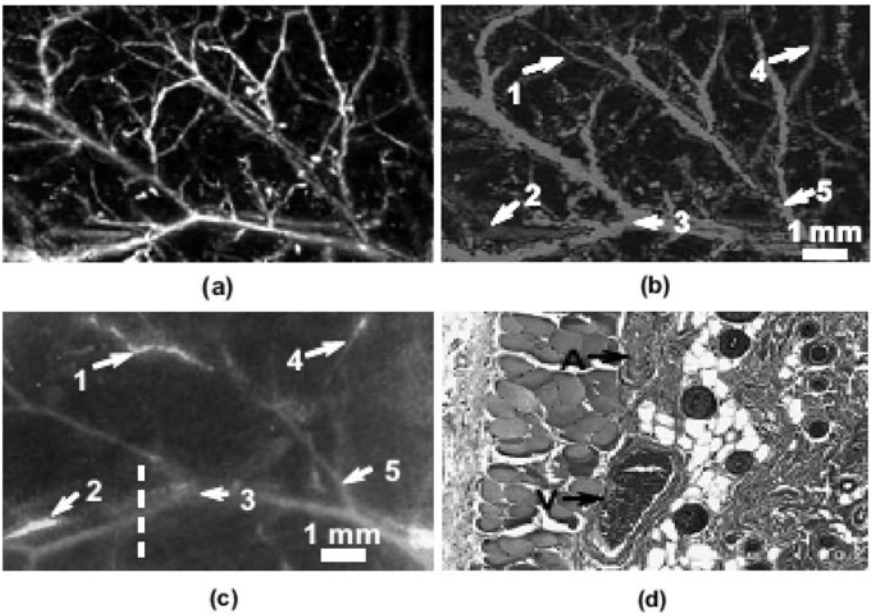


Fig. 19.7 Functional imaging of SO_2 by PAM in vivo in a 150-g Sprague-Dawley rat. Four wavelengths (578, 584, 590, and 596 nm) were employed to acquire four images. (**a**) Structural image acquired at 584 nm. (**b**) Vessel-by-vessel SO_2 mapping based on a least-squares fitting of the four images. Grayscale indicates the SO_2. (**c**) Microsphere-perfusion image of arterioles and venules. The perfusion study was modified from the ink-and-gelatin procedure [40] by using 1-ml suspension of 20-mg red-fluorescent microspheres. Vessel features numbered 1–5 match those in *b*, indicating that the vessels with high and low SO_2 are arterioles and venules, respectively. (**d**) HE-stained section showing the vessel structure along the dashed line in *c*. The arteriole and venule can be distinguished by their wall thicknesses. A: arteriole; V: venule. (Reprinted with permission from [36])

A volumetric image of the subcutaneous micro-vasculature is acquired with PAM at the 584-nm optical wavelength from a human palm in vivo (Fig. 19.8). The largest vessel in the image is approximately 350 μm in diameter, whereas the smallest is within one pixel (50 μm). The average ratio of the vessel to the background in photoacoustic signal amplitude is 35±2, and the contrast-to-noise ratio is 51 dB on average.

PAM is fundamentally a high-speed technology. The acquisition time of PAM is currently limited by the 10-Hz laser pulse repetition frequency. While scanning for a single B-scan image takes only 10 s, a two-dimensional scan over an 8 mm by 8 mm area takes more than 18 min for single-wavelength imaging and 160 min for multi-wavelength SO_2 imaging. The data acquisition can be shortened considerably, however. Each A-line is acquired within only 2 μs, which equals the maximum imaging depth (3 mm) divided by the speed of sound (~1.5 mm/μs). The time lapse between two consecutive laser pulses should be long enough for the photoacoustic

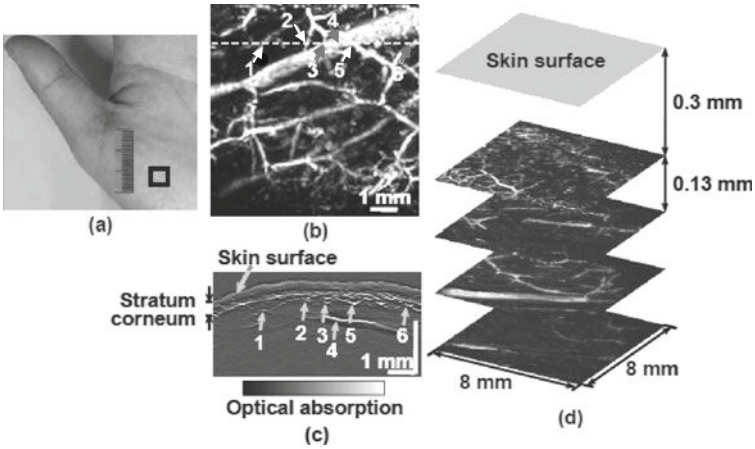


Fig. 19.8 Imaging of the total hemoglobin concentration in subcutaneous vasculature of a human palm acquired in vivo by PAM at 584 nm. (**a**) Photograph of the imaged area. (**b**) MAP image of the photoacoustic amplitude below the stratum-corneum layer projected along the z axis. (**c**). B-scan image in the z–x plane at the dashed line in panel b. The skin surface and the stratum corneum as well as blood vessels are visualized. Numbers 1–6 indicate the corresponding vessels in the MAP and B-scan images. (**d**) A series of MAP images from different layers parallel with the skin surface. The first layer starts from 0.17 mm beneath the skin surface, whereas the layer spacing is 0.13 mm. (Reprinted with permission from [36])

wave induced by the first one to die down. If the time lapse is set to 20 μs, for example, the laser pulse repetition rate can be as high as 50 kHz, which is greater than the aforementioned 13 kHz limited by the ANSI safety standard. The author's laboratory is working on accelerating data acquisition using a much higher laser pulse repetition frequency (2 kHz).

19.6 Reconstruction-Based Photoacoustic Tomography

In the previous section, photoacoustic images are formed directly by scanning a focused ultrasonic transducer. In this section, we consider reconstruction-based photoacoustic tomography or photoacoustic computed tomography, which is formed by scanning an unfocused ultrasonic transducer – ideally a point transducer – followed by image reconstruction. The initial photoacoustic pressure excited by laser pulse $\delta(t)$ equals $p_0(\mathbf{r})$ (Eq. (19.10)). The acoustic pressure $p(\mathbf{r}_0, t)$ at position $\mathbf{r}_0$ and time t, initiated by source $p_0(\mathbf{r})$, is measured around the tissue by an unfocused ultrasonic transducer.

Spherical, cylindrical, and planar detection configurations are considered (Fig. 19.9), where the detection surface S_0 encloses the source $p_0(\mathbf{r})$. Since the derivation of the reconstruction algorithm is beyond the scope of this book, we provide the final back-projection formula [43] only:

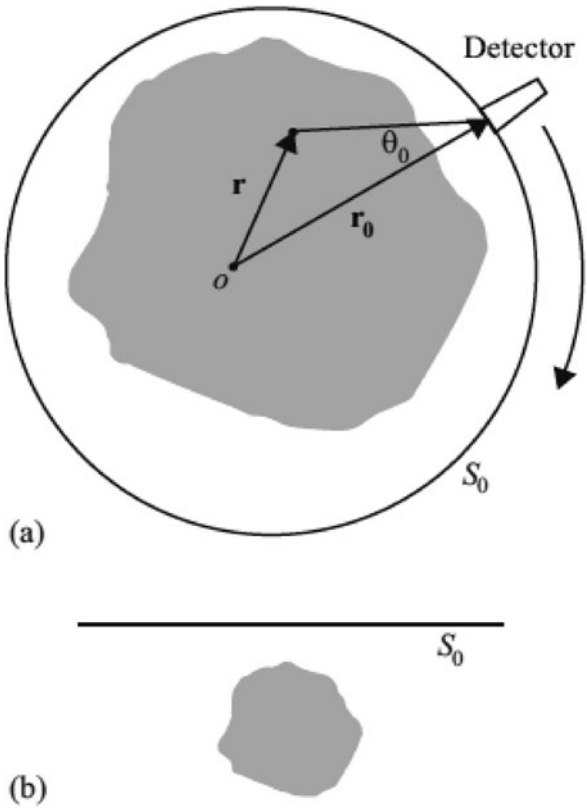


Fig. 19.9
Reconstruction-based photoacoustic tomography. (**a**) Spherical or cylindrical detection configuration. During measurement, an ultrasonic point detector at position $\mathbf{r}_0$ on surface S_0 receives photoacoustic signals emitted from source $p_0(\mathbf{r})$. During image reconstruction, a quantity related to the measurement at position $\mathbf{r}_0$ projects backward via a spherical surface (not shown) centered at $\mathbf{r}_0$. (**b**) Planar detection configuration

$$p_0(\mathbf{r}) = \frac{1}{\Omega_0} \int_{S_0} \left[2p(\mathbf{r}_0, t) - 2t \partial p(\mathbf{r}_0, t) / \partial t \right] \frac{\cos \theta_0}{|\mathbf{r}, \mathbf{r}_0|^2} dS_0 \tag{19.24}$$

Here, $\Omega_0 = 2\pi$ for the planar geometry and $\Omega_0 = 4\pi$ for the spherical or cylindrical geometry, and θ_0 denotes the angle between the detection surface normal and the vector pointing to the reconstruction point $\mathbf{r}$ as illustrated in Fig. 19.9a.

A circular scanning configuration of PAT is implemented to image small animal brains (Fig. 19.10a) [44]. A Q-switched Nd:YAG laser provides light pulses (532-nm wavelength, 6.5-ns pulse duration, and 10-Hz pulse repetition frequency). The laser beam is expanded and homogenized to provide relatively uniform incident fluence, which is less than 10 mJ/cm^2 on the skin surface. Photoacoustic waves are coupled through water to an ultrasonic transducer. The ultrasonic transducer has a center frequency of 3.5 MHz and is cylindrically focused in the direction perpendicular to the imaging plane to achieve a ~1 mm out-of-plane resolution.

Since a circle – rather than a full spherical surface – is scanned, the above reconstruction algorithm is only approximately applicable. Nevertheless, good images are

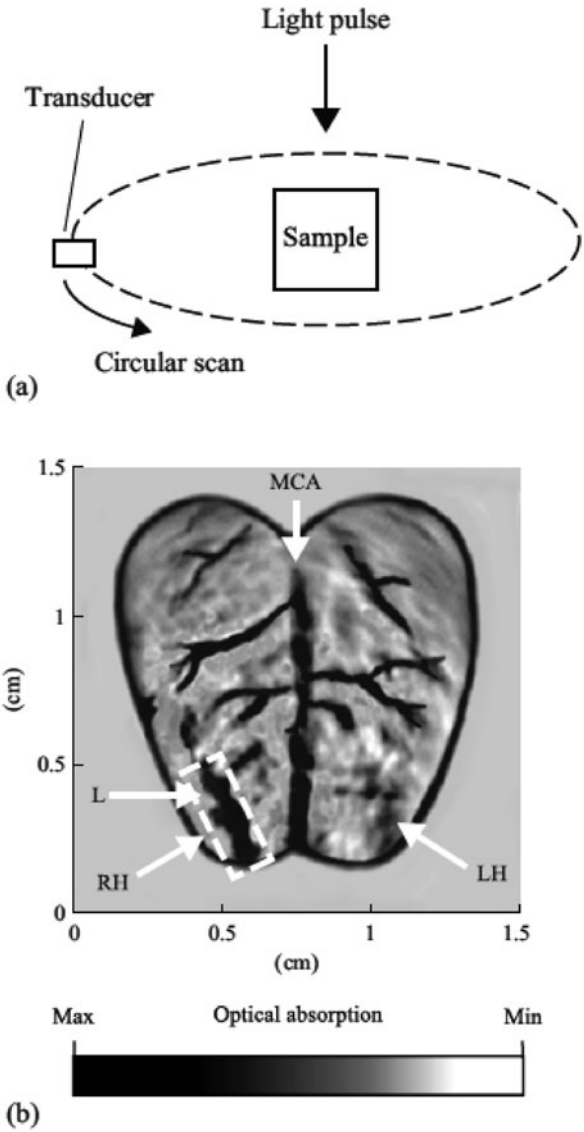


Fig. 19.10 (**a**) Diagram of a circular-scanning PAT system for small-animal imaging. (**b**) A cross-sectional PAT image of a rat brain. RH represents right cerebral hemisphere; LH, left cerebral hemisphere; L, lesion; MCA, middle cerebral artery. (Adapted with permission from [44])

still attainable. Blood vessels in the cortical surface of small animals can be imaged transcranially with the scalp and skull intact although the hair is removed for light delivery (Fig. 19.10b). At this optical wavelength, the contrast of hemoglobin is high; the imaging depth is limited to about 1 cm, which is greater than the dimension of the entire brain of a small animal such as a rat.

Several differences between the above described two representative implementations of photoacoustic imaging are summarized here. First, the confocal PAM

operates in reflection (backward) mode, whereas the reconstruction-based PAT works in orthogonal mode. Second, the confocal PAM provides lateral resolution using spherical focusing of the ultrasonic transducer and the axial resolution using the time resolution of the transducer. In contrast, the circular-scanning PAT derives in-plane resolution from reconstruction and out-of-plane resolution from cylindrical focusing of the ultrasonic transducer. To some extent, the reconstruction can be thought of as a numerical acoustic lens. Third, the confocal PAM is based on a scanning single-element ultrasonic transducer, whereas reconstruction-based PAT can be implemented with an array of ultrasonic transducers for faster data acquisition. It should be mentioned that photoacoustic imaging can be implemented in other forms depending on the applications in mind.

Acknowledgment The author thanks H. Zhang and K. Maslov for providing unpublished Fig. 19.4. This work was sponsored by National Institutes of Health grants R01 EB000712 and R01 NS46214.

References

1. Bell AG. On the production and reproduction of sound by light. *Am. J. Sci*., 20:305–324 (1880).
2. Karabutov AA, Podymova NB, and Letokhov VS. Time-resolved laser optoacoustic tomography of inhomogeneous media. *Appl. Phys. B-Lasers Opt*., 63(6):545–563 (1996).
3. Esenaliev RO, Karabutov AA, and Oraevsky AA. Sensitivity of laser optoacoustic imaging in detection of small deeply embedded tumors. *IEEE J. Sel. Top. Quant. Electron*., 5(4):981–988 (1999).
4. Kruger RA, Reinecke DR, and Kruger GA. Thermoacoustic computed tomography-technical considerations. *Med. Phys*., 26(9):1832–1837 (1999).
5. Wang LHV, Zhao XM, Sun HT, and Ku G. Microwave-induced acoustic imaging of biological tissues. *Rev. Sci. Instrum*., 70(9):3744–3748 (1999).
6. Karabutov AA, Savateeva EV, Podymova NB, and Oraevsky AA. Backward mode detection of laser-induced wide-band ultrasonic transients with optoacoustic transducer. *J. Appl. Phys*., 87(4):2003–2014 (2000).
7. Ku G and Wang LHV. Scanning thermoacoustic tomography in biological tissue. *Med. Phys*., 27(5):1195–1202 (2000).
8. Kostli KP, Frauchiger D, Niederhauser JJ, Paltauf G, Weber HP, and Frenz M. Optoacoustic imaging using a three-dimensional reconstruction algorithm. *IEEE J. Sel. Top. Quant. Electron*., 7(6):918–923 (2001).
9. Kostli KP, Frenz M, Weber HP, Paltauf G, and Schmidt-Kloiber H. Optoacoustic tomography: Time-gated measurement of pressure distributions and image reconstruction. *Appl. Opt*., 40(22):3800–3809 (2001).
10. Ku G and Wang LHV. Scanning microwave-induced thermoacoustic tomography: Signal, resolution, and contrast. *Med. Phys*., 28(1):4–10 (2001).
11. Xu MH, Ku G, and Wang LHV. Microwave-induced thermoacoustic tomography using multisector scanning. *Med. Phys*., 28(9):1958–1963 (2001).
12. Xu Y and Wang LHV. Signal processing in scanning thermoacoustic tomography in biological tissues. *Med. Phys*., 28(7):1519–1524 (2001).
13. Paltauf G, Viator JA, Prahl SA, and Jacques SL. Iterative reconstruction algorithm for optoacoustic imaging. *J. Acoustical Soc. Am*., 112(4):1536–1544 (2002).
14. Xu MH and Wang LHV. Time-domain reconstruction for thermoacoustic tomography in a spherical geometry. *IEEE Trans. Med. Imaging*, 21(7):814–822 (2002).

15. Xu Y, Feng DZ, and Wang LHV. Exact frequency-domain reconstruction for thermoacoustic tomography – I: Planar geometry. *IEEE Trans. Med. Imaging*, 21(7): 823–828 (2002).
16. Xu Y, Xu MH, and Wang LHV. Exact frequency-domain reconstruction for thermoacoustic tomography – II: Cylindrical geometry. *IEEE Trans. Med. Imaging*, 21(7):829–833 (2002).
17. Andreev VG, Karabutov AA, and Oraevsky AA. Detection of ultrawide-band ultrasound pulses in optoacoustic tomography. *IEEE Trans. Ultrasonics Ferroelectrics Frequency Control*, 50(10):1383–1390 (2003).
18. Finch D, Patch SK, and Rakesh. Determining a function from its mean values over a family of spheres. *SIAM J. Math. Anal*., 35(5):1213–1240 (2003).
19. Karabutov AA, Savateeva EV, and Oraevsky AA. Optoacoustic tomography: New modality of laser diagnostic systems. *Laser Phys*., 13(5):711–723 (2003).
20. Kostli KP and Beard PC. Two-dimensional photoacoustic imaging by use of Fourier-transform image reconstruction and a detector with an anisotropic response. *Appl. Opt*., 42(10):1899–1908 (2003).
21. Wang XD, Pang YJ, Ku G, Stoica G, and Wang LHV. Three-dimensional laser-induced photoacoustic tomography of mouse brain with the skin and skull intact. *Opt. Lett*., 28(19):1739–1741 (2003).
22. Xu MH and Wang LHV. Analytic explanation of spatial resolution related to bandwidth and detector aperture size in thermoacoustic or photoacoustic reconstruction. *Phys. Rev. E*, 67(5):056605 (2003).
23. Eghtedari M, Copland JA, Kotov NA, Oraevsky AA, and Motamedi M. Optoacoustic imaging of nanoparticle labeled breast cancer cells: A molecular based approach for imaging of deep tumors. *Lasers Surg. Med*., 52–52 (2004).
24. Haltmeier M, Scherzer O, Burgholzer P, and Paltauf G. Thermoacoustic computed tomography with large planar receivers. *Inverse Problems*, 20(5):1663–1673 (2004).
25. Ku G, Wang XD, Stoica G, and Wang LHV. Multiple-bandwidth photoacoustic tomography. *Phys. Med. Biol*., 49(7):1329–1338 (2004).
26. Xu Y and Wang LHV. Time reversal and its application to tomography with diffracting sources. *Phys. Rev. Lett*., 92(3):033902 (2004).
27. Cox BT and Beard PC. Fast calculation of pulsed photoacoustic fields in fluids using k-space methods. *J. Acoustical Soc. Am*., 117(6):3616–3627 (2005).
28. Ku G and Wang LHV. Deeply penetrating photoacoustic tomography in biological tissues enhanced with an optical contrast agent. *Opt. Lett*., 30(5):507–509 (2005).
29. Ku G, Wang XD, Xie XY, Stoica G, and Wang LHV. Imaging of tumor angiogenesis in rat brains in vivo by photoacoustic tomography. *Appl. Opt*., 44(5):770–775 (2005).
30. Zhang J, Anastasio MA, Pan XC, and Wang LHV. Weighted expectation maximization reconstruction algorithms for thermoacoustic tomography. *IEEE Trans. Med. Imaging*, 24(6):817–820 (2005).
31. Wang LHV and Wu HI. *Biomedical optics: Principles and imaging*. Wiley, Hoboken, NJ (2007).
32. Xu M and Wang LHV. Biomedical photoacoustics. *Rev. Sci. Instruments*, 96(16):163902 (2006).
33. Gusev VE and Karabutov AA. *Laser optoacoustics*. American Institute of Physics, New York (1993).
34. Morse PM and Feshbach H. *Methods of theoretical physics*. McGraw-Hill, Boston, MA (1999).
35. Maslov K, Stoica G, and Wang LHV. In vivo dark-field reflection-mode photoacoustic microscopy. *Opt. Lett*., 30(6):625–627 (2005).
36. Zhang HF, Maslov K, Stoica G, and Wang LHV. Functional photoacoustic microscopy for high-resolution and noninvasive in vivo imaging. *Nat. Biotechnol*., 24(7):848–851 (2006).
37. Briggs GAD. *Acoustic microscopy*. Clarendon, Oxford (1992).
38. Jobsis FF. Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters. *Science* 198:1264–1267 (1977).

39. Chance B, Leigh JS, Miyaka H, Smith DS, Niola DS, Greenfeld R, Finander M, Kaufmann K, Levy W, Young M, Chen P, Yoshioka P, and Boretsky R. Comparison of time-resolved and -unresolved measurements of deoxyhemoglobin in brain. *Proc. Natl. Acad. Sci. USA*, 85:4971–4975 (1988).
40. Ambach G and Palkovits M. Blood supply of the rat hypothalamus I. nucleus supraopticus. *ACTA Morphologica Academiae Scientiarum Hungaricae*, 22:291–310 (1974).
41. Laser Institute of America, American National Standard for Safe Use of Lasers ANSI Z136. 1-2000, American National Standards Institute, New York, NY (2000).
42. Zhang HF, Maslov K, and Wang LHV. In vivo imaging of subcutaneous structures using functional photoacoustic microscopy. *Nat. Protocols*, 2:797– 804 (2007).
43. Xu MH and Wang LHV. Universal back-projection algorithm for photoacoustic computed tomography. *Phys. Rev. E*, 71(1):016706 (2005).
44. Wang XD, Pang YJ, Ku G, Xie XY, Stoica G, and Wang LHV. Noninvasive laser-induced photoacoustic tomography for structural and functional in vivo imaging of the brain. *Nat. Biotechnol*., 21(7):803–806 (2003).

Chapter 20
Steady State Fluorescence Spectroscopy for Medical Diagnosis

Anita Mahadevan-Jansen and Steven C. Gebhart

20.1 Introduction

Light can react with tissue in different ways and provide information for identifying the physiological state of tissue or detecting the presence of disease. The light used to probe tissue does so in a *non-intrusive* manner and typically uses very low levels of light far below the requirements for therapeutic applications. The use of fiber optics simplifies the delivery and collection of this light in a *minimally invasive* manner. Since tissue response is virtually instantaneous, the results are obtained in *real-time* and the use of data processing techniques and multi-variate statistical analysis allows for *automated detection* and therefore provides an *objective* estimation of the tissue state. These then form the fundamental basis for the application of optical techniques for the detection of tissue physiology as well as pathology. These distinct advantages have encouraged many researchers to pursue the development of the different optical interactions for biological and medical detection.

20.2 Optical Spectroscopy

Study of the interaction between electromagnetic radiation (EM) and matter is formally defined as *spectroscopy.*

$$\text{Spectroscopy} = \text{interaction of EM radiation with matter}$$

An optical spectrum is a plot of the intensity of light (from the ultraviolet to the infrared) as a function of wavelength.

A. Mahadevan-Jansen (✉)
Departments of Biomedical Engineering and Neurosurgery, Vanderbilt University, Nashville, TN 37235, USA
e-mail: anita.mahadevan-jansen@vanderbilt.edu

A.J. Welch, M.J.C. van Gemert (eds.), *Optical-Thermal Response of Laser-Irradiated Tissue*, 2nd ed., DOI 10.1007/978-90-481-8831-4_20,

Tissue response to light is typically measured in either imaging or spectroscopy. Optical spectroscopy is the measurement of light intensity as a function of wavelength with high spectral resolution but without much spatial information. Although multi-fiber spectroscopic probes provide some spatial information, optical imaging is the technique used when spatial information is needed. Imaging, however, provides limited spectral information. Spectral imaging is a method by which light spectrum at every picture element, i.e. pixel, of a two-dimensional image is measured. The images acquired using such a method consist of a cube of information $I_{x,y}(\lambda)$, which contains the full spectrum at each pixel position (x, y). In this chapter I represents intensity and, depending upon its detection and calibration of the system used, is assigned arbitrary units (au). Thus spectral imaging combines the features of spectroscopy and imaging to provide spatial as well as spectral information. The various responses of tissue to light can be displayed and utilized in any of these forms depending on the application of the technique.

When a photon is incident on a molecule, it may be transmitted, reflected, absorbed or scattered. Different techniques arise from these different light-tissue interactions. In the context of tissue detection, these techniques include:

- *Absorption spectroscopy*, where the absorption of light as a function of wavelength is measured. For biologic materials, near-infrared (near IR) to infrared (IR) wavelengths provide the most valuable information that is relevant to detection. Infrared wavelengths represent transitions between vibrational energy levels and therefore provide biochemical information [1]. However, the most significant chromophore in tissue in the IR is water, which limits the application of IR spectroscopy in vivo. For example, visible and near-IR wavelengths provide functional information associated with blood [2, 3].
- *Elastic Scattering and Diffuse Reflectance Spectroscopy* form their basis on the occurrence of elastic scattering as light impacts tissue [4]. Their spectrum measures the intensity of reflected light as a function of wavelength. Diffuse reflectance yields a measure of the optical properties of the tissue [5]. This technique is particularly useful in accounting for blood absorption in tissue [6].
- *Fluorescence Spectroscopy* is the measurement of fluorescence intensity as a function of wavelength. Fluorescence results following transitions between electronic energy levels.
- *Raman Spectroscopy* relies on the occurrence of inelastic scattering between photons and the vibrational levels of molecular bonds [7]. A Raman spectrum is a plot of scattered intensity as a function of the energy difference between the incident and scattered photons. It provides a molecular specific fingerprint that is a snapshot of tissue biochemistry [8].

All other methods of optical detection rely on these optical phenomena or a combination thereof. Each of these methods has pros and cons and their utility for tissue detection depends on the application. This chapter presents the concept, instrumentation and application of fluorescence spectroscopy as a steady state technique for tissue detection.

20.3 What is Fluorescence?

To explain the phenomenon of fluorescence, we must consider the molecular structure of matter. Molecules possess rotational, vibrational and electronic behavior and corresponding energy levels. The absorption and emission of light by a molecule is illustrated by the Jablonski diagram [9] where the ground, first and second electronically excited states are depicted by S_0, S_1 and S_2, respectively (Fig. 20.1). Within each of these states, vibrational energy levels are represented by a number in parenthesis – e.g. $S_0(0)$.

At room temperature, molecules primarily reside in the ground state. When stimulators such as photons are absorbed (each photon has energy, $h\nu_a$ where h is the Planck's constant (6.626×10^{-34}J·s) and ν_a is the frequency at which absorption occurs), molecules occupying the lowest vibrational level of the ground state $S_0(0)$ are excited to higher electronic levels (S_1,S_2).

Following absorption, several processes can occur. In most organic compounds, the molecules will typically relax to $S_1(0)$ by heat generation – referred to as *internal conversion*. This occurs virtually instantaneously; the time scale of this event is 10^{-14}–10^{-12} s. (One known exception is azulene, which emits fluorescence from S_2 as well) [9]. From $S_1(0)$, the molecule can relax to the ground state by three processes: (1) they can internally convert (by heat generation) to any of the S_0 states; (2) they can emit fluorescence; and (3) they can emit phosphorescence. In *fluorescence*, the emitted photons have energy, $h\nu_f$ equal to the energy difference between the excited and ground states, where ν_f is the frequency of the fluorescent photon.

$$h\,\nu_f = E_{S1(0)} - E_{S0(0)}[\mathrm{J}] \tag{20.1}$$

The average time a molecule remains in the excited state before emitting fluorescence (*fluorescence lifetime*) is 10^{-10}–10^{-8} s.

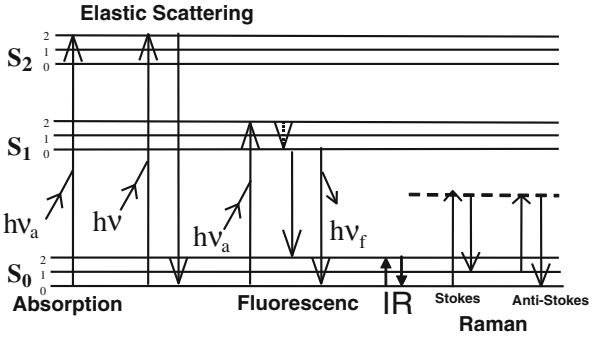


Fig. 20.1 Modified Jablonski diagram showing the vibrational (0,1,2) and electronic (S) energy levels and processes that can take place with the absorption and scattering of light

Molecules in the S_1 state can also undergo conversion to the first triplet state T_1. Emission from T_1 to return to S_0 is termed as *phosphorescence*. *Phosphorescence lifetimes* are in the order 10^{-3}–10^{0} s.

20.4 Fluorescence Spectroscopy

The fluorescence emitted can be mathematically represented in Eq. (20.2).

For a dilute solution with only 1 fluorophore (defined as a molecule/compound that exhibits fluorescence), *fluorescence intensity* is a function of both excitation and emission wavelengths and can be written as;

$$I(\lambda_{\text{exc}}, \lambda_{\text{em}}) = I_0(\lambda_{\text{exc}})\, c\, \varepsilon(\lambda_{\text{exc}})\, \ell\ Q(\lambda_{\text{em}}) \ \ [\text{au}] \tag{20.2}$$

Which is the same as;

$$I(\lambda_{\text{exc}}, \lambda_{\text{em}}) = I_0(\lambda_{\text{exc}})\ \mu_a(\lambda_{\text{exc}})\, \ell\ Q(\lambda_{\text{em}}) \tag{20.3}$$

since

$$c\ \varepsilon(\lambda_{\text{exc}}) = \mu_a(\lambda_{\text{exc}}) \ \ [\text{cm}^{-1}]$$

where

$I_0(\lambda_{\text{exc}})$ is the incident intensity of light at the excitaion wavelength, λ_{exc}
c is the concentration of the fluorophore
$\varepsilon(\lambda_{\text{exc}})$ is the molar extinction coefficient at the excitation wavelength
$\mu_a(\lambda_{\text{exc}})$ is the absorption coefficient at the excitation wavelength
ℓ is the length or thickness of the fluorophore
$I_0(\lambda_{\text{exc}})\, \mu_a(\lambda_{\text{exc}})$ represents the amount of light absorbed
$Q(\lambda_{\text{em}})$ is the quantum yield of the fluorophore at the emission wavelength and represents the amount of light absorbed that is converted to fluorescence

The *fluorescence quantum yield Q* is a measure of the proportion of molecules emitting fluorescence as compared to the total number of molecules excited and is a function of the emission wavelength. Fluorescence intensity may be measured

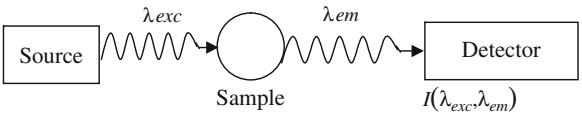


Fig. 20.2 A schematic of a basic fluorescence setup is shown for the purposes of understanding the mathematical representation of fluorescence

with respect to the excitation or emission wavelength, the two variables in the above equations.

- An *excitation spectrum* depicts the fluorescence intensity as a function of excitation wavelength for a fixed emission wavelength.

$$I(\lambda_{em}) = I_0(\lambda_{exc})\,\mu_a(\lambda_{exc})\,\ell\,\underbrace{Q(\lambda_{em})}_{\text{constant}} \tag{20.4}$$

  Since $Q(\lambda_{em})$ is a constant, the *excitation spectrum* looks like the absorption spectrum for that fluorophore.
- An *emission spectrum* depicts the fluorescence intensity as a function of emission wavelength for a fixed excitation wavelength.

$$I(\lambda_{em}) = I_0(\lambda_{exc})\;\underbrace{\mu_a}_{\text{constant}}\;(\lambda_{exc})\;\ell\;\;Q(\lambda_{em}) \tag{20.5}$$

  Since the absorption term is a constant, the emission spectrum represents the quantum yield for that fluorophore.
- *Excitation emission matrices* (EEMs) are assembled from a series of fluorescence emission spectra collected at sequential excitation wavelengths. For example,

  For λ_{exc1}, measure intensity as a function of λ_{em}
  For λ_{exc2}, measure intensity as a function of λ_{em} and so on

This data results in a matrix where the first row consists of the excitation wavelengths, the first column consists of the emission wavelength and the rest of the matrix contains the fluorescence intensities at the corresponding excitation and emission wavelength as shown below

$$\begin{bmatrix} & \lambda_{exc} & \ldots & \ldots \\ \lambda_{em} & I_1 & I_2 & I_n \\ . & .. & .. & .. \\ . & I_m & .... & I_{m,n} \end{bmatrix} \tag{20.6}$$

and

$$I\,(\lambda_{exc},\,\lambda_{em})\;\;= I_0\,(\lambda_{exc})\,c\,\varepsilon(\lambda_{exc})Q(\lambda_{em})$$

EEMs form surface plots that are usually presented as contour plots where each contour line connects points of equal fluorescence intensity (similar to isobar and isothermal charts) as shown in Fig. 20.3.

Emission spectra may vary depending on the chemistry of the fluorophore and the environment. The position of the fluorescence maxima is dependent on the environment and the dynamics of the fluorophore. A turbid sample such as tissue contains not only multiple fluorophores, which may or may not interact with each other, but

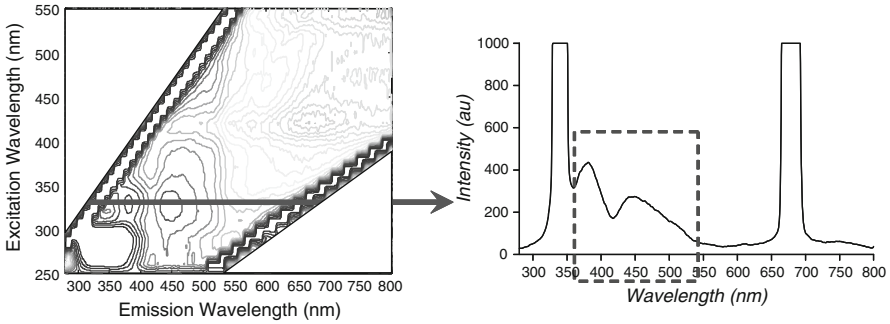


Fig. 20.3 An EEM of normal brain tissue is shown along with an emission spectrum at 340 nm. The two spikes represent the first and second order of elastic scattering

also contain scatterers and absorbers. Thus, an EEM of such a material shows the fluorescence excitation emission maximum of each fluorophore present altered from its pure state by its environment (see Fig. 20.3).

Any scatterers present typically manifest themselves in a broad band attenuation of the collected spectrum. The presence of an absorber is indicated by valleys at both the excitation and emission wavelengths at which it absorbs. Thus, fluorescence gives a measure of the composition as well as the environment of the sample of interest.

20.5 Properties of Fluorescence

20.5.1 Stokes' Shift

In general, the emission wavelength is observed to be longer than the excitation wavelength. This can be explained by the conservation of energy. Molecules excited into higher electronic and vibrational levels quickly relax into the lowest vibrational level of S_1. Thus

$$E_{\text{abs}} = E_{\text{IC}} + E_{\text{f}} \tag{20.7}$$

where E_{abs} is the absorbed or incident energy, E_{IC} is the energy relaxed by internal conversion and E_{f} is the energy released by fluorescence.

$$\begin{aligned} &h\,\nu_a = \text{heat} + h\,\nu_f \text{ (from Fig. 20.1)} \\ &h\,\nu_a > h\,\nu_f \\ &\nu_a > \nu_f \\ &\lambda_a < \lambda_f \end{aligned} \tag{20.8}$$

This shift in frequency (and wavelength) between absorption (excitation) and fluorescence (emission) is called Stokes' shift. In practice, this means that the emission wavelength is always longer than the excitation wavelength.

20.5.2 Emission Wavelength is Independent of the Excitation Wavelength

In an emission spectrum, by definition, the fluorescence intensity is only dependent on the quantum yield. Thus, the shape of the emission spectrum is independent of the incident frequency (excitation wavelength), because emission typically occurs from the lowest vibrational state of the lowest electronically excited state i.e. $S_1(0)$ [9].

20.5.3 Mirror Image Rule

The excitation spectra, $I(\lambda_{exc})$ are indicative of the vibrational levels of the excited states whereas the emission spectra indicate those of ground state. Since generally excitation does not alter these energy levels (with some exceptions), the emission spectrum appears as a mirror image of the excitation spectrum (Fig. 20.4). This symmetry is due to the same vibrational levels being involved in absorption and emission as indicated by the Frank-Condon's principle [10].

20.5.4 Fluorescence Lifetimes and Quantum Yield

Quantum yield, Q is defined as the proportion of photons emitted by fluorescence as compared to the photons absorbed.

$$Q = \frac{\text{number of photons emitted by fluorescence}}{\text{number of photons absorbed}}$$

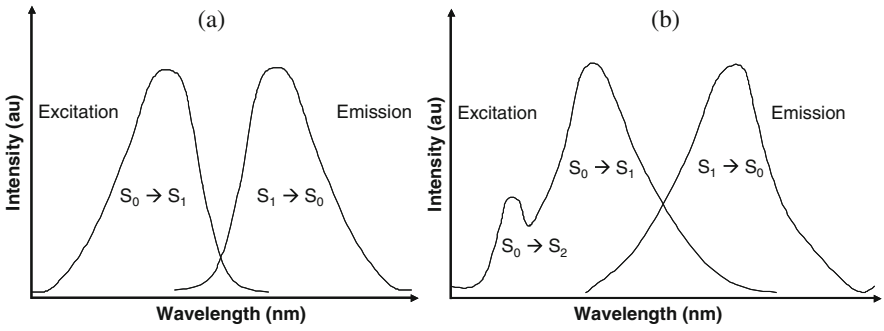


Fig. 20.4 (**a**) Demonstrates the energy transitions that lead to the mirror image rule. (**b**) Demonstrates exceptions to this rule when absorption occurs from $S_2 \rightarrow S_0$

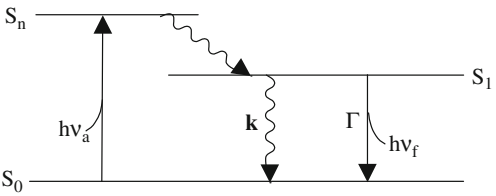


Fig. 20.5 Simplified Jablonski diagram showing the rate processes involved in fluorescence. Γ is the rate of emission of photons by fluorescence, k is the rate of internal conversion (IC)

This can also be written in terms of the rate processes (as shown in Fig. 20.5) as $Q = \Gamma/(\Gamma + k)$. When $\Gamma >>> k$, i.e. fluorescence is the dominant phenomenon, then $Q = 1$. However, due to Stokes' shift, some of the energy absorbed is always lost by heat when decaying to S_1, which in turn implies that Q cannot be 1. Therefore, conservation of energy determines that Q is always less than 1. Fluorescence dyes such as Rhodamine and Fluorescein have high quantum yields (~ 0.8) and thus exhibit bright fluorescence.

Fluorescence lifetime τ_f, is defined as the average time that a molecule spends in S1 before decaying to the ground state. It can be written as,

$$\tau_f = \frac{1}{\Gamma + k} \tag{20.9}$$

Fluorescence lifetimes vary in different molecules, providing researchers with a diagnostic tool. This measure also forms the basis for such imaging tools as fluorescence lifetime imaging microscopy (FLIM) [9].

20.6 What Fluoresces?

Molecules that exhibit fluorescence are called fluorophores. Fluorescence is exhibited primarily by organic molecules with abundant delocalized electrons i.e. molecules with numerous conjugated bonds. For example, aromatic compounds with several benzene rings are known to exhibit fluorescence.

20.6.1 Natural Biological Fluorophores

There are numerous biological fluorophores that exist in tissue that exhibit intrinsic fluorescence.

(1) Proteins: The most common amino acid that exhibits fluorescence is tryptophan. Its excitation emission maximum occurs at (290, 330 nm). Its fluorescence lifetime is 2.6 ns. Most proteins exhibit this peak associated with tryptophan in their EEM. Another amino acid that exhibits fluorescence is tyrosine. Its excitation emission maximum occurs at (230, 300 nm) with a lifetime of 3.6 ns.

However, although tyrosine fluoresces in solution, its fluorescence is very weak in proteins and tissues.

(2) Nucleic Acids: Nucleic acids do not in general, exhibit fluorescence. One known exception is Yeast-tRNA.

(3) Co-factors: Some metabolic products exhibit fluorescence. NAD(P)H has an excitation emission maximum at (340, 450 nm). Its fluorescence lifetime is 0.5 ns. However, interestingly, NAD^+ does not fluoresce.

(4) Flavins: Flavins such as flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD) exhibit fluorescence at (450, 515 nm). Their lifetimes are 4.8 ns and 2.3 ns, respectively.

20.6.2 Extrinsic Fluorophores

The natural fluorescence properties of biological molecules are often weak and need to be enhanced. To improve the signal-to-noise of the fluorescence signal, extrinsic fluorophores are often added to display improved spectra. Some commonly used dyes include fluorescein, rhodamine isocyanates and isothiocyanates, often used to label proteins. Fluorescent agents including green and yellow fluorescent protein (GFP and YFP), Cy3 and Cy5, are commonly used to tag and track molecular targets [9]. A table of the excitation and emission maxima of some these contrast agents can be found at http://dbc.bio.uci.edu/OBCresources/fluorophores.pdf.

The isolation of porphyrins and the subsequent discovery of their tumor-localizing properties (and phototoxic effects on tumor tissue) led to the implementation of these agents for fluorescence enhancement in tissue applications. Two derivatives of porphyrins that are approved for clinical use are 5-aminolevulenic acid (5-ALA) [11] and Photofrin [12]. These agents been applied for the detection of several cancers including bladder cancers [13], skin and oral cancers [14].

Other contrast agents include quantum dots and nanoparticles. Quantum dots are single fluorescent semiconductor nanocrystals that represent one of the most effective methods of imaging in vivo cells in real time. Made from a cadmium core and composed of 200–10,000 atoms, quantum dots are manufactured in a vast array of sizes and compositions, thus altering the absorption and emission properties [15]. This makes the quantum dots capable of being resolved at multiple wavelengths simultaneously, creating multicolor images with wavelengths from 400 to 4000 nm [16]. When compared to organic dyes, they offer many advantages including their large extinction coefficient, large Stokes shift and photostability, allowing for the long-term acquisition of brighter and more crisp in vivo images. Quantum dots allow long-term vasculature imaging which is impossible with fluorescent dyes and proteins due to the fact that the dyes are quickly removed from the bloodstream [17]. Quantum dots are also able to track the movement and differentiation of single cells and molecules over time due to their longer lifetimes. One drawback to quantum dots is their potential toxicity due to heavy-metal composition and nanoscale properties [18]. They are found to be highly toxic when under UV illumination or in oxidative environments for extended periods of time since this dissolves the

semi-conductor core, releasing cadmium ions. They can be made biocompatible when coated with the polymer layer and in the absence of UV irradiation; however, this increases the size of the quantum dot, thus modifying its properties.

Presently, quantum dots are used for in vivo imaging of tumors and diagnosis in cells and animals, but could be developed to simultaneously image and treat cancer cells by utilizing the correct antibody coating [18]. Future quantum dot applications could also include using them as a real time surgical guide in conjunction with MRI or CT scans, allowing surgeons to correctly remove the tumor without damaging the surrounding healthy tissue.

20.7 Quantitative Representation of Fluorescence

The generation and propagation of fluorescence by a fluorophore can be quantitatively described as a function of concentration, excitation and emission wavelengths, and position within a solvent or turbid medium such as tissue.

20.7.1 In a Dilute Solution

The intensity of fluorescence generated by a fluorophore in a dilute solution that is trans-illuminated by excitation light can be described as a function of the distance from the surface of the solution. From Beer's Law, the one-dimensional longitudinal distribution of excitation light $I(\lambda_{exc}, z)$ from the surface to a given depth z is:

$$I\left(\lambda_{exc},\, z\right) \;=\; I_o\left(\lambda_{exc}\right) e^{-\mu_a(\lambda_{exc})z} \tag{20.10}$$

where

- $I(\lambda_{exc})$ is the intensity of light at the excitation wavelength at the solution surface ($z = 0$)
- $\mu_a(\lambda_{exc})$ is the absorption coefficient of the fluorophore solution at the excitation wavelength.

Since optical density is given by $-\mathrm{Log}_{10}[I(z)/I_0]$, which is equal to $c\varepsilon(z)$ and $2.303\mathrm{Log}_{10}x = \mathrm{Ln}x$, to convert Eq. (20.10) in terms of the concentration of the fluorophore, we use the relationship that $\mu_a(\lambda_{exc}) = 2.303c\varepsilon(\lambda_{exc})$, where c is the concentration of the fluorophore and $\varepsilon(\lambda_{exc})$ is the decadic molar extinction coefficient at the excitation wavelength.

Therefore:

$$I\left(\lambda_{exc}\right) \;=\; I_o\left(\lambda_{exc}\right) e^{-2.30c\varepsilon(\lambda_{exc})z} \tag{20.11}$$

For a dilute solution, the optical density is small ($\mathrm{OD} = 2.303\varepsilon(\lambda_{exc})cz << 1$). Using the fact that e^z can be replaced by $(z + 1)$ for a small exponent z, the distribution of excitation light can be expressed as:

$$I\left(\lambda_{\text{exc}}, z\right) = I_o\left(\lambda_{\text{exc}}\right)\left[1 - 2.303\varepsilon\left(\lambda_{\text{exc}}\right) cz\right] \tag{20.12}$$

Fluorescence generation at the emission wavelength $F(\lambda_{\text{exc}}, \lambda_{\text{em}})$ is defined as the product of

(1) the excitation light absorbed $I_{\text{abs}}(\lambda_{\text{exc}})$
(2) the quantum yield $Q(\lambda_{\text{exc}}, \lambda_{\text{em}})$ of the fluorophore at the excitation and emission wavelengths:

$$F\left(\lambda_{\text{exc}}, \lambda_{\text{exc}}\right) = I_{\text{abs}}\left(\lambda_{\text{exc}}\right) Q\left(\lambda_{\text{exc}}, \lambda_{\text{exc}}\right) \tag{20.13}$$

Given that the total excitation light absorbed by the solution between the surface and depth z is the incident light minus the light transmitted to depth z:

$$I_{\text{abs}}\left(\lambda_{\text{exc}}\right) = I_o\left(\lambda_{\text{exc}}\right) - I\left(\lambda_{\text{exc}}, z\right) \tag{20.14}$$

the total fluorescence generation between the surface and that depth can be expressed as:

$$\begin{aligned} F\left(\lambda_{\text{exc}}, \lambda_{\text{em}}, z\right) &= \left[I_o\left(\lambda_{\text{exc}}\right) - I\left(\lambda_{\text{exc}}, z\right)\right] Q\left(\lambda_{\text{exc}}, \lambda_{\text{em}}\right) \\ &= \left[I_o\left(\lambda_{\text{exc}}\right) - I_o\left(\lambda_{\text{exc}}\right)\left[1 - 2.303\varepsilon\left(\lambda_{\text{exc}}\right) cz\right]\right] Q\left(\lambda_{\text{exc}}, \lambda_{\text{em}}\right) \\ &= 2.303 I_o\left(\lambda_{\text{exc}}\right) \varepsilon\left(\lambda_{\text{exc}}\right) czQ\left(\lambda_{\text{exc}}, \lambda_{\text{em}}\right) \end{aligned} \tag{20.15}$$

In sum, the total fluorescence generation in a dilute fluorophore solution of thickness L can be written as:

$$F\left(\lambda_{\text{exc}}, \lambda_{\text{em}}\right) = 2.303 I_o\left(\lambda_{\text{exc}}\right) \varepsilon\left(\lambda_{\text{exc}}\right) cQ\left(\lambda_{\text{exc}}, \lambda_{\text{em}}\right) L \tag{20.16}$$

20.7.2 In a Concentrated Solution

To derive the fluorescence which escapes the surface of a concentrated fluorophore solution, we divide the thickness of the solution into infinitesimally shallow layers (each with thickness dz), treat each of them as a dilute solution, and integrate those layers across the entire thickness of the solution.

$$F\left(\lambda_{\text{exc}}, \lambda_{\text{em}}, z\right) = 2.303 I_o\left(\lambda_{\text{exc}}, z\right) \varepsilon\left(\lambda_{\text{exc}}\right) cQ\left(\lambda_{\text{exc}}, \lambda_{\text{em}}\right) dz \tag{20.17}$$

Since $I_o(\lambda_{\text{exc}}, z)$ is equal to $I(\lambda_{\text{exc}}, z)$ in Eq. (20.17) above,

$$F\left(\lambda_{\text{exc}}, \lambda_{\text{em}}, z\right) = 2.303 I_o\left(\lambda_{\text{exc}}\right) e^{-2.303\varepsilon\left(\lambda_{\text{exc}}\right) cz} \varepsilon\left(\lambda_{\text{exc}}\right) cQ\left(\lambda_{\text{exc}}, \lambda_{\text{em}}\right) dz \tag{20.18}$$

Given the isotropic nature of fluorescence generation, only half of fluorescent photons are directed toward the surface of the solution. Further, the fluorescent photons are absorbed by the solution based upon its molar extinction coefficient at the

emission wavelength according to Beer's Law. Based on these two behaviors, the fluorescence emitted at the solution surface that was generated within layer dz can be expressed as:

$$F_{\text{top}}\left(\lambda_{\text{exc}}, \lambda_{\text{em},\, z}\right) = \frac{F\left(\lambda_{\text{exc}}, \lambda_{\text{em},\, z}\right)}{2} e^{-2.303\varepsilon(\lambda_{\text{exc}})cz} = \frac{2.303}{2} I_o\left(\lambda_{\text{exc}}\right) e^{-2.303[\varepsilon(\lambda_{\text{exc}})+\varepsilon(\lambda_{\text{em}})]cz} \varepsilon\left(\lambda_{\text{exc}}\right) cQ\left(\lambda_{\text{exc}}, \lambda_{\text{em}}\right) dz \tag{20.19}$$

Integrating this expression across the thickness L of the solution, the total fluorescence emitted from the top surface of the solution is

$$F_{\text{top}}\left(\lambda_{\text{exc}}, \lambda_{\text{em}}\right) = \frac{2.303}{2} I_o\left(\lambda_{\text{exc}}\right) \varepsilon\left(\lambda_{\text{exc}}\right) cQ\left(\lambda_{\text{exc}}, \lambda_{\text{em}}\right) \int_0^L e^{-2.303[\varepsilon(\lambda_{\text{exc}})+\varepsilon(\lambda_{\text{em}})]cz} dz = \frac{I_o\left(\lambda_{\text{exc}}\right) Q\left(\lambda_{\text{exc}}, \lambda_{\text{em}}\right)}{2} \frac{\varepsilon\left(\lambda_{\text{exc}}\right)}{\varepsilon\left(\lambda_{\text{exc}}\right) + \varepsilon\left(\lambda_{\text{em}}\right)} \left[1 - e^{-2.303[\varepsilon(\lambda_{\text{exc}})+\varepsilon(\lambda_{\text{em}})]cL}\right] \tag{20.20}$$

20.7.3 Extracting Intrinsic Fluorescence

The previous section discussed the calculation of intrinsic fluorescence properties of turbid media when measured using an ideal measurement geometry. Extracting intrinsic fluorescence becomes significantly more complicated once boundary conditions of non-infinite excitation and collection geometries come into play, since these geometries significantly affect measured spectral lineshapes in ways which cannot be predicted with analytical models. An array of researchers has tackled the problem by recognizing that reflectance spectra are often affected in similar ways and can be used to extract intrinsic fluorescence from measured spectra [19–21].

Most models developed to date extract intrinsic fluorescence by normalizing fluorescence spectra with reflectance spectra measured with an identical geometry. These models are couched in the idea that fluorescence and reflectance photon propagation are similarly governed by tissue absorption, scattering, and boundary conditions such that the photon paths traveled are almost identical. The only disparities between the photon paths result from: (1) fluorescence generation is isotropic while reflectance is forward scattering in soft tissues, and (2) tissue optical properties at the excitation and emission wavelengths differ. The differences in the models are borne from the various analyses used to develop corrections made to the reflectance spectra used for normalization to compensate for these exceptions.

Researchers from Michael Feld's laboratory developed a series of models [21] which compensate for the exceptions by using "effective" tissue parameters for anisotropy and albedo:

$$g_{\text{eff}} = \frac{N - 1}{N} g \tag{20.21}$$

where N is the number of scattering events before fluorescence emission.

$$a_{\text{eff}}(\lambda_i) = r_d a(\lambda_i) + (1 - r_d)\, a(\lambda_i) \left[\frac{\mu_t^z(\lambda_{\text{em}})}{\mu_t^z(\lambda_{\text{exc}})} \right] \tag{20.22}$$

which use a priori optical property information at both the excitation and emission wavelengths and where r_d is the diffuse reflectance at the tissue surface [22]. The measured fluorescence spectrum is corrected using the equation:

$$\varphi(\lambda_{\text{em}}) = k(\lambda_{\text{exc}}) \frac{F(\lambda_{\text{em}})}{R_{\text{eff}}(\lambda_{\text{em}})} \tag{20.23}$$

where

$\varphi(\lambda_{\text{em}})$ is the intrinsic fluorescence quantum yield spectrum
$k(\lambda_{\text{exc}})$ is a single proportionality constant that is calculated using reflectance at the excitation wavelength
$F(\lambda_{\text{em}})$ is the measured fluorescence spectrum across the spectral range of emission wavelengths
$R_{\text{eff}}(\lambda_{\text{em}})$ is a corrected reflectance spectrum across the emission wavelengths of the fluorescence spectrum

The proportionality constant and corrected reflectance spectrum are calculated using the effective anisotropy and albedo as inputs into Monte Carlo models of reflectance propagation. While the full derivation of the model is beyond the scope of this chapter, Eq. (20.23) demonstrates that the measured fluorescence emission spectrum is thus normalized by an "effective" reflectance developed from Monte Carlo propagation models to remove the effects of absorption, scattering, and measurement boundary conditions [22].

Gardner et al. developed a similar method based on a one-dimensional analysis of Monte Carlo simulations designed to analytically describe, as functions of diffuse reflectance, the fluence rate of excitation light and the escape function of fluorescence photons within the tissue [23]. From their model, the intrinsic fluorescence coefficient β (defined as the product of the absorption coefficient at the excitation wavelength and the quantum yield of the fluorophore) can be calculated from the collection efficiency D of the measurement device as well as the penetration depth and diffuse reflectance emission at the excitation and emission wavelengths:

$$\beta(\lambda_{\text{exc}}, \lambda_{\text{em}}) = \frac{F(\lambda_{\text{exc}}, \lambda_{\text{em}})}{D(\lambda_{\text{em}})\, X_{1D}(Rd_{\text{exc}}, Rd_{\text{em}}, \delta_{\text{exc}}, \delta_{\text{em}})} \tag{20.24}$$

where

$F(\lambda_{\text{exc}}, \lambda_{\text{em}})$ is the measured fluorescence spectrum
$X_{1D}(Rd_{\text{exc}}, Rd_{\text{em}}, \delta_{\text{exc}}, \delta_{\text{em}})$ is a complex calculated reflectance spectrum which uses reflectance (Rd) and a priori penetration depth (δ) information at the excitation and emission wavelengths.

A similar model using a complex analysis of diffusion theory has also been developed by Zhadin and Alfano to correct the measured reflectance spectra [24]:

$$T_{EM}\left(\lambda_{\text{exc}}, \lambda_{\text{em}}\right) = \frac{F\left(\lambda_{\text{exc}}, \lambda_{\text{em}}\right)}{N\left(\lambda_{\text{exc}}\right) Y\left(\lambda_{\text{exc}}, \lambda_{\text{em}}\right)} \tag{20.25}$$

where

$F(\lambda_{\text{exc}}, \lambda_{\text{em}})$ is the measured fluorescence spectrum
$N(\lambda_{\text{exc}})$ is a normalization constant based on the mean free path of photons at the excitation wavelength
$Y(\lambda_{\text{exc}}, \lambda_{\text{em}})$ is a complex calculated reflectance spectrum which uses measured reflectance intensities at the excitation and emission wavelengths

The advantage of their method is that it requires no a priori optical property information at the excitation or emission wavelengths, only measured reflectance at those wavelengths.

In contrast to these complex analytical correction models, Lin et al. developed a rudimentary but elegant method to correct fluorescence spectra solely for the effects of hemoglobin absorption, the major chromophore in soft tissue [6]. Using Beer's Law, the effect of a superficial layer of blood of thickness L on measured fluorescence can be described as:

$$F\left(\lambda_{\text{em}}\right) = F_o\left(\lambda_{\text{em}}\right) e^{-[\mu_a(\lambda_{\text{exc}}) + \mu_a(\lambda_{\text{em}})]L} = F_o\left(\lambda_{\text{em}}\right) e^{-(k+1)\mu_a(\lambda_{\text{em}})L} \tag{20.26}$$

where

F_o is the true fluorescence emission
$\mu_a(\lambda_{\text{exc}})$, $\mu_a(\lambda_{\text{em}})$ are the absorption coefficients of blood at the excitation and emission wavelengths, respectively
k is the ratio of the two coefficients $-k = \mu_a(\lambda_{exc})/\mu_a(\lambda_{em})$

Similarly, the effect on measured reflectance can be written as:

$$R\left(\lambda_{em}\right) = R_0\left(\lambda_{em}\right) e^{-2\mu_{a,em}L} \tag{20.27}$$

Lin et al. recognized that simply computing the ratio of measured fluorescence to measured reflectance raised to the $(k+1)/2$ power effectively removes the exponentials in the equations (i.e. removes the effect of the blood absorption) [6]:

$$\frac{F\left(\lambda_{em}\right)}{\left[R\left(\lambda_{em}\right)\right]^{\left(\frac{k+1}{2}\right)}} = \frac{F_0\left(\lambda_{em}\right) e^{-(k+1)\mu_{a,em}L}}{\left[R_0\left(\lambda_{em}\right)\right]^{\left(\frac{k+1}{2}\right)} e^{-\left(\frac{k+1}{2}\right)(2\mu_{a,em}L)}} = \frac{F_0\left(\lambda_{em}\right)}{\left[R_0\left(\lambda_{em}\right)\right]^{\left(\frac{k+1}{2}\right)}} \tag{20.28}$$

leaving only the true fluorescence and diffuse reflectance emission spectra. While this represents a very simple way of removing the effects of blood absorption of

measured fluorescence spectra, it has its limitation in that the resulting ratio is neither a true fluorescence nor a reflectance spectrum, but a hybrid of the two. As such, it can be difficult to interpret spectral features and discern tissue properties from the spectra.

As can be seen from the described methods, extracting intrinsic fluorescence information from measured spectra can be a daunting task. Doing so in a medium as complex as soft tissue often involves complex models, simplifying assumptions, and/or intricate calibrations. Using reflectance spectra to normalize fluorescence spectra requires a multi-modal spectroscopy system whose spectral range encompasses both the excitation and emission wavelengths, calling for more complex instrumentation. While methods to acquire optical property information in vivo have progressed greatly in recent years, requiring optical property information for the model also often forces the user to make guesses about the composition of the tissue he/she is trying to discriminate in order to perform diagnoses in real time. For these reasons, teasing apart intrinsic fluorescence from the effects of measurement geometry and other optical parameters such as absorption and scattering that obfuscate differences between tissues remains a significant challenge to the utility of fluorescence spectroscopy for robust tissue diagnosis.

20.8 Instrumentation

The typical system used to measure fluorescence spectra is often referred to as a spectrofluorometer. The basic system consists of a light source, a wavelength selector, a sample delivery and collection system, a dispersive element and a detector (see Fig. 20.6).

20.8.1 Light Source

This can be a laser or a white light source depending on the study design. In a system where the fluorescence properties of the sample are unknown, it may be necessary to measure the EEM of the sample. In this case, a white light source that is typically a mercury or xenon arc lamp is used. Measuring the EEM, however, adds significant time to the measurement and may not be practical in an in vivo setting. In this case,

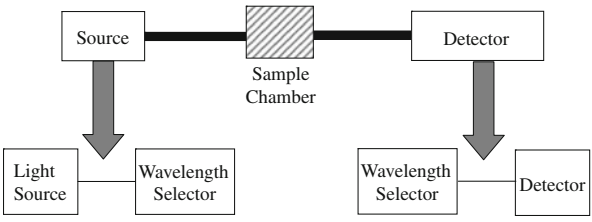


Fig. 20.6 Schematic of a basic fluorescence spectroscopy measurement system

1–3 optimal wavelengths are selected that provide the information necessary for the study. In this case, a laser emitting the specific wavelengths designed may be used. This laser is often a nitrogen laser or a nitrogen-pumped dye laser.

20.8.2 Wavelength Selector (Excitation)

When multiple wavelengths are required, such as when building an EEM, a device that disperses the white light as a function of wavelength is used. The output can then be tuned to select each specific wavelength. The wavelength selector may be a monochromator or, in a simpler case, may consist of a series of bandpass filters in a filter wheel. Most current monochromators use diffraction gratings rather than prisms for wavelength selection. At the excitation leg, a monochromator with high efficiency in the ultraviolet-visible region is typically chosen.

20.8.3 Wavelength Dispersion (Emission)

The wavelength selector at the emission end may be a spectrograph or monochromator. While a monochromator allows only one wavelength of light to exit at a given time, an imaging spectrograph disperses wavelengths along one axis (typically the *x*-axis) to exit. The selection of the grating is critical as the dispersion of the grating will determine the spectral resolution of the system. Further, the gratings in combination with the size of the detector used determine the range of wavelengths that are measured. Since biological fluorescence is broad (∼10–20 nm), gratings with 150–300 g/mm rulings are typically selected.

20.8.4 Detector

If the emission leg uses a monochromator for wavelength dispersion, intensity at each emission wavelength is sequentially measured with a photomultiplier tube (PMT) as a detector. PMTs are selective and have the highest efficiency for detection, i.e. are most sensitive. A PMT consists of a photocathode which ejects electrons when struck with photons. The generation efficiency of photoelectrons is proportional to the incident wavelength. The electrons then go through a series of dynodes that amplify the signal. Since a PMT needs a single detector, a monochromator serially selects wavelengths to build a spectrum.

However, current technology has seen tremendous development of charge coupled device cameras (CCDs) that consist of a two-dimensional array of detectors. When CCDs are coupled with an imaging spectrograph, it is possible to simultaneously detect all wavelengths along the *x*-axis while providing limited spatial information along the *y*-axis. When multiple emission fibers are used in a fiber optic probe, the spectra from individual fibers can be spatially separated and analyzed using the vertical axis of the CCD. Since biological fluorescence typically occurs in the UV-VIS, a front-illuminated camera with a UV window is used in a fluorescence system.

20.8.5 Delivery

Excitation light must be delivered to the sample of interest, while emitted light must be collected and delivered to the dispersion optics and detector. In a device such as a spectrofluorometer, light is simply relayed via optical trains which focus the excitation light directly onto the sample and image the excitation spot onto the detector. Despite advances in miniaturization of CCD-based spectrometers, spectrofluorometers often employ monochromators for wavelength selection during both excitation and emission and a PMT for detection, giving the instrument the versatility to acquire excitation spectra, emission spectra, or full EEMs from a given sample. However, the significant bulk of direct delivery systems often limits their utility to in vitro or ex vivo applications.

For clinical in vivo applications, access to hard-to-reach areas inside the body is frequently required. This is accomplished through the working channel of an endoscope. For these applications, researchers have developed fiber-optic probes to deliver and collect light from tissue. Fiber-optic probes are often terminated with standard SMA connectors to permit direct and consistent coupling to laser excitation sources and miniaturized CCD-based spectrometers. Fiber optics provide excellent light transmission across a wide range of wavelengths from the UV to the NIR and allow the instrumentation for excitation and emission to be physically separated from the sample by several meters. Flexible fiber optics also allow the probe to be directly housed in, or compatible with, devices such as catheters and endoscopes that are tailor-made for access to hard-to-reach body areas, such that optical diagnosis can be incorporated into minimally invasive procedures.

Fiber-optic probes for fluorescence spectroscopy have been designed in a variety of configurations [25], with the major distinction based on the number of fibers contained in the probe:

a. *Single-fiber probes* – A single fiber is employed for both delivery of excitation light and collection of fluorescence emission. This configuration guarantees overlap of the excitation and collection spots on the tissue surface, to collect photons from the region of peak fluorescence emission. These probes often employ an evanescent fiber coupler to split the excitation and collection light between the fiber legs which lead to the excitation source and detector, causing a 50% reduction in light intensity along both the excitation and emission light paths (see Chapter 17).
b. *Multi-fiber probes* – Multiple fibers are employed with one fiber usually dedicated to delivery of the excitation light and the remaining for collection of emitted fluorescence. Exclusivity between excitation and collection fibers results in no collection at the excitation-fiber face where peak emission occurs. However, the reduction in collected fluorescence by each individual fiber is overcome by eliminating losses due to evanescent coupling and by enlarging the collection area with multiple collection fibers. Multi-fiber probes also hold the capability for multi-modal spectroscopy (e.g. combined fluorescence and diffuse reflectance), since more than one fiber can be coupled to multiple excitation and illumination sources.

c. *Probe designs tailored to target specific tissue layers* – The spatial geometry of the fiber configuration where the probe interfaces the tissue sample has also been repeatedly shown to affect the measured spectral intensity and lineshape, even in homogeneous tissues [26–28]. Fiber diameters and the distances between them dictate the degree to which photons which are capable of being collected have interacted with chromophores and scatterers inside the sample between excitation and emission. Greater separation distances leads to collection of more diffuse photons, which have the potential to give more information on tissue composition, especially with respect to tissue chromophores. However, emission intensity decreases with distance from the excitation source fiber, creating a trade-off between collected intensity and tissue information that can be gleaned from the spectra. This behavior is discussed in more detail in Section 20.8.7 where probe-based spectroscopy is compared to non-contact spectral imaging.

20.8.6 Spectral Imaging

Fiber-optic probe-based spectroscopy is useful for correlating spectral features to histopathological diagnosis at a specific interrogation site. In principle, optical spectroscopy has the capability to rapidly, objectively, and non-invasively differentiate between tissue types; however, probe-based spectroscopy has limited clinical applicability. It produces single-point diagnostic measurements without providing spatial information regarding the extent of disease. This information is necessary to efficiently guide biopsy removal, tissue resection, or therapy monitoring. For optical biopsy to be clinically useful, single-point spectroscopy systems must be extended to spectral imaging in which spectral information is acquired at every pixel within a two-dimensional field of view, yielding spatial and spectral tissue information for a comprehensive snapshot of tissue pathology in real time.

Successful translation from probe-based spectroscopy to macroscopic spectral imaging faces three main challenges: (1) reasonable image acquisition time despite significant decreases in excitation irradiance due to larger target areas, (2) a convenient clinical interface for seamless integration into the clinical setting, and (3) maintenance of accurate spectral discrimination between tissue types. While spectral discrimination capability will be discussed in Section 20.8.8, speed of spectral image acquisition and creating a versatile, compact imaging system for use in the clinic depend to a large degree on optical design and spectral imaging implementation. The following section outlines the various modalities for spectral imaging.

20.8.6.1 Spectral Imaging Modalities

Acquiring spatially resolved spectral information via spectral imaging yields three-dimensional data cubes with two dimensions of spatial information and one dimension which varies with wavelength (Fig. 20.7). Spectral imaging methods are

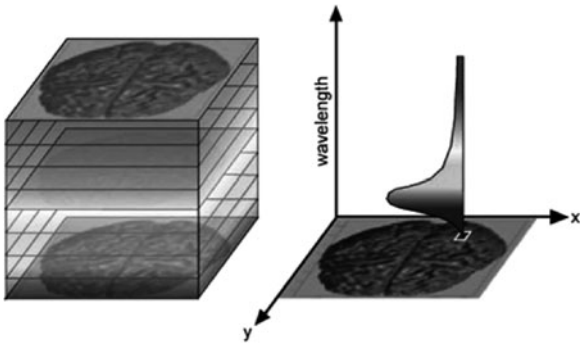


Fig. 20.7 Spectral image data cube with two dimensions (*x,y*) of spatial information and one dimension (wavelength) of spectral information

common in the fields of astronomy [29–31], remote sensing [32–34], and chemical compound analysis [35–37] to elucidate the composition and characteristics of terrestrial, atmospheric, and extra-terrestrial objects. Current spectral imaging systems for life science applications are coupled to benchtop microscopes for fluorescence in situ hybridization [38–41], general cell visualization [42–44], and fluorescent-protein cell trafficking (FRET) [45, 46].

The surge of optical methods for tissue discrimination have led to many optical imaging systems for tissue contrast; however, few of the systems designed to date perform true spectral imaging, a term generally reserved for the acquisition of detailed spectral information at every pixel, and only recently have any been quantitatively tested for helping guide disease management. Most tissue-based optical imaging systems acquire limited spectral information (a few emission bands), trading decreased spectral information for increased frame rates by employing filter-based assemblies containing bandpass-filter wheels, dichroic filters, and/or beamsplitters and oftentimes optically implementing tissue classification algorithms in hardware [47, 48]. These systems have been most commonly applied to exploit the preferential deposition of exogenous fluorescent dyes to tumor tissue [49–51] and to target the hemoglobin absorption bands of soft tissue for perfusion imaging [52, 53]. While the speed of these optical imaging systems is impressive, hardware acquisition of partial spectral information limits the complexity of viable classification algorithms, which are often unique to a given organ system, requiring a separate imaging system and training algorithm for each tissue type.

Spectral imaging systems capable of good resolution across all three dimensions can be grouped according to the method by which they acquire three dimensions of information using a two-dimensional optical detector (CCD camera): line-scanning, tunable-filters, and multiplexing methods such as Fourier-transform interferometry. Line-scanning systems typically utilize motors to scan a 1-D excitation-collection line of light across the tissue surface [54–56]. At each scan position, the line of light collected from the tissue surface is aligned along one dimension of a CCD array while a grating disperses the spectral components along the orthogonal dimension. Thus, each image acquired by the CCD camera contains the complete spectra from one spatial row of the image, and successive CCD frames are acquired as the

excitation-collection line is scanned across the tissue surface. Line-scanning methods are simple to implement; however, the motors often cause vibration artifacts within the images and the images require as many CCD frames as the spatial dimensionality of the image, resulting in time-consuming data readout from the camera. The beam is usually scanned using scanning mirrors.

Tunable-filter spectral imaging systems [57–59] continuously excite and collect emission from the entire tissue surface and employ an electronically tunable bandpass filter, such as a liquid-crystal tunable filter (LCTF) or an acousto-optic tunable filter (AOTF), to transmit or diffract, respectively, one band of the emission spectrum to the CCD camera at a time. In this way, tunable filter spectral imaging systems use the filter to directly parse the wavelength dimension of the spectral image. A frame of CCD data is acquired for each band of the emission spectrum, using the entire CCD array for spatial resolution. The spectral dimension of the image is constructed by acquiring successive CCD frames as the filter tuning wavelength is swept across the spectral range of interest. The most recent generations of these filters possess a wide spectral range across the visible or infrared regions of the spectrum with excellent spectral resolution (±5 nm) and can be tuned on the order of microseconds (AOTF) to milliseconds (LCTF).

Liquid-crystal tunable filters are constructed from a stack of linear polarizers with interspersed layers of liquid crystals and birefringent crystals (Fig. 20.8a). The retardance (R) of each stage, defined as the product of its thickness (d) and the difference between n_o and n_e (the refractive indices for polarizations perpendicular (ordinary) and parallel (extraordinary) to the axis of anisotropy respectively – see Chapter 4), $R = d(n_o - n_e)$, can be modulated by applying a voltage to the liquid crystal and changing its extraordinary index of refraction. Altering the stage's retardance affects the phase delay between the ordinary and extraordinary light rays emerging from the stage, thereby changing its transmission (T) as a function of wavelength (λ), $T = 1/2\cos^2(x)$ where $x = R/\lambda$. The thicknesses (d) of the birefringent crystals in the filter stack and the voltages applied to the liquid crystals are

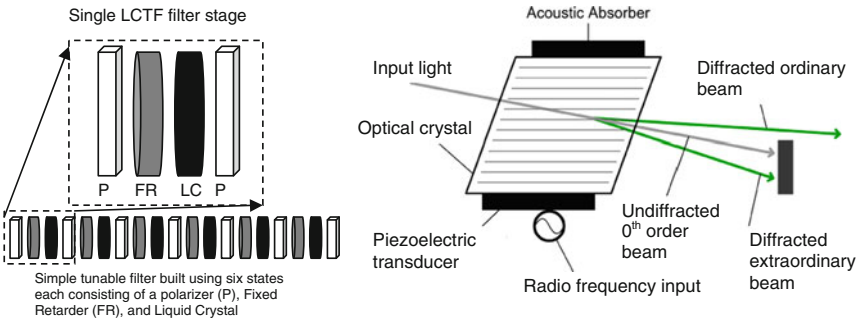


Fig. 20.8 Schematics showing (**a**) the construction of a liquid-crystal tunable filter and (**b**) the principles of operation for an acousto-optic tunable filter

often selected to produce a binary series of retardances – R, $2R$, $4R$, etc. – such that all of the stages transmit the selected passband while destructive interference at the other wavelengths "trims" the stopband light. The number of stages employed in the filter dictates the width of the passband and the degree of stopband attenuation. The advantage to LCTF implementation in a spectral imaging system is the simplicity in the optical design. Without disturbance of the passband beam or moving parts to create vibration artifacts in the images, the system requires only a camera lens to focus the transmitted light onto the CCD focal plane. However, the peak transmission of liquid-crystal devices for co-polarized light ranges from 5% (at 400 nm) to 60% (at 700 nm) across the visible spectrum, and a major drawback to any filter-based imaging system is the inherent trade-off between the system's optical throughput (3–10%) – defined as the fraction of light collected by the imaging system which is ultimately detected by the camera – and its spectral resolution.

Acousto-optic tunable filters apply radio-frequency (RF) acoustic waves to modulate the index of refraction of the filter crystal. If the incident wavelength satisfies a momentum-matching condition based upon the frequency of the RF wave, the linear polarization states of the incoming light are diffracted to the opposite state and the light is deflected upon its exit from the crystal, with two symmetrically diffracted passband beams and a transmitted zero-order stopband beam exiting the filter (Fig. 20.8b). Since the diffracted and zero-order beams must be separated for spectral imaging, optical design of AOTF systems requires either:

(1) a long object distance to allow physical separation of the diffracted and zero-order beams; or,
(2) orthogonally oriented polarizers on either side of the filter to polarize the input light and allow only orthogonally oriented output light (i.e. the passband) to reach the camera.

Although AOTF diffraction efficiency is on the order of 90–95% for polarized input light (and 70% for unpolarized light), both beam separation solutions result in a loss of at least 50% of the light collected from the sample.

Multiplexing methods for optical measurements were first developed for spectral analysis in the infrared region of the electromagnetic spectrum due to a dearth of low-noise infrared detectors. Rather than isolating and measuring individual spectral bands of a signal, multiplexing methods weighed the spectral components in groups using orthogonal basis functions (cosines in the case of Fourier spectroscopy) and measure the total intensity of each weighted group. Measuring multiple spectral components simultaneously and back-calculating their intensities from the weighted measurements provided greater optical throughput than spectrometer and filter-based systems with subsequent increase in signal-to-noise for equivalent integration times.

Fourier-based imaging systems use Michelson or Sagnac interferometers [60] where the optical path difference (OPD) between two beams created and subsequently recombined by a 50/50 beamsplitter is varied between successive CCD frames (Fig. 20.9a). Varying the OPD modulates the interference pattern seen by

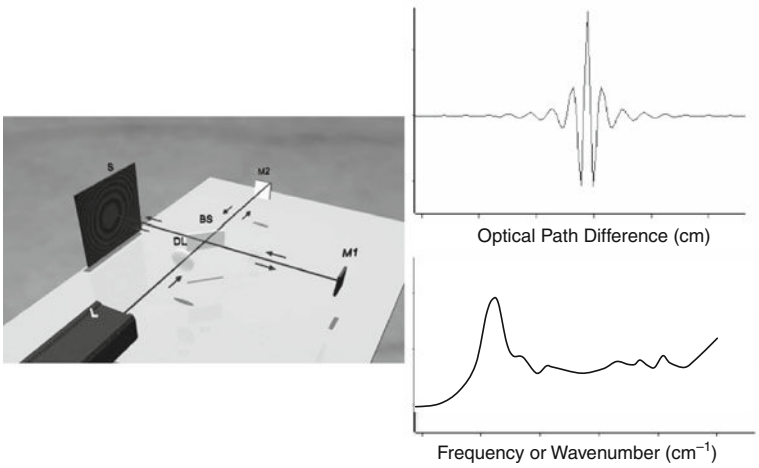


Fig. 20.9 (**a**) Michelson interferometer with laser input showing interference fringes as a result of changes in optical path difference with radial position. [http://www.sewanee.edu/Chem/Chem&Art/Detail_Pages/ColorProjects_2004/Rodgers/Rodgers.htm]. In spectral imaging, one of the mirrors is replaced with the sample of interest, and the other mirror is scanned along the optical axis to create multiple OPD values. (**b**) Example inverse Fourier transformation from interferogram as a function of optical path difference to spectrum as a function of wavenumber (inverse of wavelength)

the CCD camera, building interferograms at each pixel as the OPD is shifted and the CCD frames are acquired. After complete image acquisition, each pixel interferogram represents the Fourier transform of the pixel spectrum and can be individually converted to such via the inverse Fourier transform (Fig. 20.9b). Employing a beamsplitter-based interferometer for multiplexing provides a theoretical system optical throughput of 50% (half the light is lost back out the entrance port) over the entire spectrum and a multiplex advantage of $\sqrt{[N/8]}$ (N = number of frames) over measuring the wavelengths individually with a filter system. However, considerable overhead is expended in image acquisition time since spectral extraction must be performed for every spatial pixel within the image. This process requires N^2 additions or multiplications at each pixel, a process that can become computationally intensive for a highly resolved image. The relationship between sampling the interferogram and the spectrum is not intuitive, to the point that certain combinations of image acquisition parameters can lead to incomplete spectra or images with spatial harmonics. Furthermore, the optical designs of multiplexing systems are somewhat complex and system performance is often extremely sensitive to optical alignment.

20.8.7 Spectroscopy Versus Imaging

Ideally, non-contact spectral imaging would produce spectral lineshapes equivalent to probe-based spectroscopy; otherwise, time and resources spent developing

discrimination algorithms based upon spectral differences observed in probe-based data during large, multi-patient proof-of-principle studies would be squandered, since new imaging-based algorithms must be determined. Investigators in the field of probe-based spectroscopy and spectral imaging have always been diligent in accounting for spectral variations in source illumination and detector sensitivity to remove the effects on measured lineshape of factors extrinsic to the sample, a topic which will be discussed in Section 20.8.8.

At first glance, producing equivalent lineshapes between probe-based spectroscopy and spectral imaging is as simple as applying the standard spectral calibration procedure to spectral imaging data (as described in Section 20.8.8). However, investigations over the past 20 years have repeatedly demonstrated that measured spectral lineshapes are as affected by the excitation-collection geometry of the measurement system as the spectral response of the system hardware. The results of these investigations indicate that spectral lineshape is modulated by two main factors. First, the sample optical properties dictate the impulse response of light distribution within the tissue in terms of the relative propagation and attenuation of different wavelengths and the subsequent intensity, spatial distribution, and angular profile of light emitted from the tissue surface. Second, the excitation-collection geometry of the measurement system dictates the convolution of that impulse response with the spatial dimensions of the excitation and collection regions and the excitation-collection angles to determine what fraction of the light emitted at each wavelength is ultimately measured. Therefore, an excitation-collection geometry which involves spectral variations in the fraction of remitted light collected by the system impacts the measured spectral lineshape despite correction for excitation spectrum and detector sensitivity.

The transition from probe-based spectroscopy to spectral imaging often involves a drastic shift in excitation-collection geometry, both spatially and in terms of angular relationship to the tissue surface (Fig. 20.10). Spectroscopy probes often contain

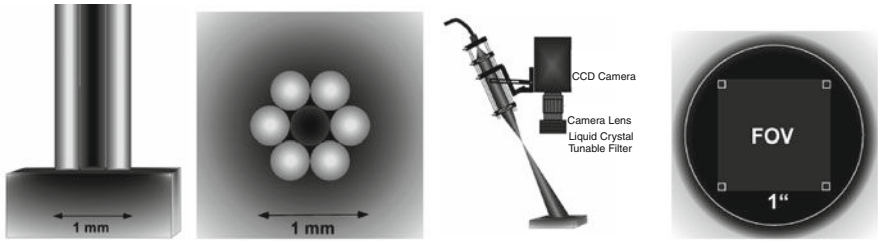


Fig. 20.10 Excitation-collection geometries for example probe-based spectroscopy and non-contact spectral imaging systems. (**a**) Fiber-optic probe with a single excitation fiber, fibers polished normally, and direct contact between the fibers and the tissue surface. (**b**) Cross-section of the probe-tissue interface for a six-around-one probe configuration with a central excitation fiber and six surrounding collection fibers relative to the light distribution in the tissue. (**c**) Spectral imaging system with non-contact excitation and collection. (**d**) Spectral imaging field of view and individual pixels relative to the excitation spotsize and subsequent light distribution

multiple fibers, some of which are used for excitation delivery and the rest for emission collection. Direct contact of the probe tip with the tissue surface and the specific configuration of the probe – the relative positions of the excitation and collection fibers, the gap between the fiber tips and the tissue surface, and the fiber polish angles – produce consistent excitation and collection angles for every sample but create a unique set of source-detector separation distances (Fig. 20.10a, b). Since the spatial distribution of light remitted from tissue varies with optical properties, its relative relationship to the fiber configuration will vary with wavelength, such that lineshapes measured with probe-based spectroscopy will be influenced by spectral variations in collection efficiency.

While multiple modalities exist for spectral imaging implementation, most spectral imaging systems acquire spectral data with non-contact excitation and collection (Fig. 20.10c). When the excitation spot is much larger than the system field of view, as is often the case, each pixel in the image can be treated as a small collection region surrounded by a semi-infinite excitation spot (Fig. 20.10d), and its spatial excitation-collection geometry as a convolution of all possible source-detector separation distances. However, non-contact imaging produces a geometry whose excitation and collection angles vary not only from image to image but across pixels within a single image when the sample surface topography is non-uniform. Therefore, while the collection efficiency of spectral imaging is not sensitive to optical property variations in the spatial distribution of remitted light, it is sensitive to the relative relationship between the excitation-collection angles and the angular emission profile of remitted light. If angular profiles of surface emission for a given excitation angle depend upon sample optical properties, angular emission will vary between wavelengths and their relative measured intensities will be modulated by detector collection that is limited to a narrow range of emission angles. If angular emission profiles depend upon incident excitation angle, an undulating sample surface can similarly induce spectral distortions within imaging spectra and extrinsic differences between them.

While previous studies in the literature, outlined below, have investigated fluorescence and diffuse reflectance propagation with results which hint at some of the aforementioned effects of excitation-collection measurement geometry on measured spectral lineshape, only recently has the effect of the shift in spatial and angular excitation-collection geometry from probe-based spectroscopy to non-contact spectral imaging been thoroughly examined.

Using Monte Carlo simulation, Qu et al. reported “nearly Lambertian” angular emission profiles for fluorescence and diffuse reflectance when excitation is normal to the tissue surface [61]. Angular emission profiles for diffuse reflectance deviated from Lambertian behavior when the excitation was delivered at oblique angles (30° and 75° with respect to the surface normal) such that the emission was biased away from the source, suggesting that angular emission of diffuse reflectance was dependent on sample optical properties when the excitation was delivered at an oblique excitation angle. Churmakov et al. reported similar roughly Lambertian emission behavior for diffuse reflectance when excited normal to the surface [62]. These

results were the first to hint at angular emission profiles of diffuse reflectance which are dependent upon incident excitation angle.

Research into diffuse reflectance propagation in tissue began with one-dimensional depth-resolved investigations in which diffuse reflectance was modeled as a function of optical properties using photon migration theory [63]. Several groups simulated or measured radially resolved profiles of diffuse reflectance for wide ranges of optical properties; the results were used to develop inverse models for in vivo optical property determination [64, 65]. Bargo et al. measured the collection efficiency of a single excitation-collection fiber and of an excitation-collection fiber pair as a function of optical properties and separation distance, but defined collection efficiency as the fraction of light impinging on the face of the collection fiber that falls within its acceptance cone, without investigating how the absolute emission intensity varied with optical properties and thereby affected diffuse reflectance lineshape [66, 67].

Several studies demonstrated changes in diffuse reflectance lineshape subsequent to a shift in excitation-collection geometry. Liu et al. simulated and experimentally measured diffuse reflectance for a subset of discrete optical property values and three different probe geometries [68]. A subsequent paper from the same research group used one of the investigated probes to demonstrate that fluorescence and diffuse reflectance lineshapes from breast tissue change between three source-detector separation distances, ascribing the disparities to probing different depths of the tissue [26].

The only studies in the literature which have examined the combined effects of optical properties and excitation-collection geometry on diffuse reflectance lineshape were conducted by Schmitt and Kumar [69, 70]. Diffuse reflectance spectra were measured in the near infrared for a single excitation-collection fiber pair in an effort to determine how collected emission varied with separation distance ($d = 1-6$ mm) for various optical properties ($\mu_a = 0.0-1.0$ mm^{-1}, $\mu_s' = 0.5-2.0$ mm^{-1}). They demonstrated that diffuse reflectance lineshape varies with source-detector separation distance and explained the shifts using diffusion theory. However, diffusion-based predictions broke down at small source-detector separation distances and, thus far, no group has validated their results in the visible spectrum where increased scattering and absorption and the shorter source-detector separation distances of most fiber-optic probes ($d = 0-1$ mm) render diffusion theory inaccurate.

Recent investigations into fluorescence propagation can be divided into two major categories: those that investigate depth-resolved fluorescence [71–75] and those that seek to extract intrinsic fluorescence information from fluorescence measurements [21–23, 76, 77]. Depth-resolved fluorescence investigations predominantly aim to identify probe geometries that will collect fluorescence from specific layers in epithelial tissues, but rarely worry about how optical properties and probe geometry affect measured lineshape. While extracting intrinsic fluorescence information from fluorescence measurements recognizes the effects of optical properties on lineshape, and actually seeks to remove those effects, most of the

methods nullify the effects of probe geometry by using diffuse reflectance information for fluorescence correction. Since the fluorescence and diffuse reflectance spectra are generally collected with the same excitation-collection geometry, diffuse reflectance is similarly modulated, eliminating any wavelength-specific geometry effects. While depth-resolved and intrinsic fluorescence studies provide knowledge into fluorescence behavior, neither of them give insight into how tissue optical properties and excitation-collection geometry combine to modulate fluorescence lineshape. Surprisingly, while the recent investigations outlined above have strayed from the issue, two of the first investigations into fluorescence propagation initially demonstrated a dependence of fluorescence lineshape on excitation-collection geometry. Richards-Kortum et al. compared measured fluorescence spectra from regions inside and outside of the excitation spot [78]. Keijzer et al. simulated and measured fluorescence spectra as a function of radial distance from the source [79]. Both studies found that fluorescence lineshape drastically changes as collection is moved out of and further from the excitation spot. Shifts in lineshape are attributed to increased probing depth as the source-detector separation distance increases, subjecting the fluorescence to greater modulations from absorption. Since then, other groups have observed similar results [80, 81]; but until recently, none had investigated spectral lineshapes measured with traditional fiber-optic probe configurations or non-contact imaging systems.

Gebhart et al. characterized fluorescence propagation and angular emission profiles as functions of incident excitation angle and sample optical properties to predict the potential effects of the transition from probe-based spectroscopy to non-contact spectral imaging on measured spectral lineshapes [82]. The study found that fluorescence emission is consistently Lambertian regardless of optical properties or incident excitation angle, most likely due to the isotropic nature of fluorescence generation within tissue. While forward scattering of diffuse reflectance biases its emission away from an angled illumination source, isotropic generation de-couples the directions of fluorescence excitation and emission photons, resulting in consistently shaped fluorescence emission from tissue. However, the study also found evidence that the amplitude of Lambertian fluorescence emission varied with optical properties differently for different incident excitation angles, suggesting that fluorescence lineshapes may be affected by surface topography during non-contact imaging.

20.8.8 Calibrating Your Spectra (Or Data Processing)

Fluorescence spectra $F(\lambda)$ are modulated by the sensitivity and response of the detection system used to measure them:

$$F_{\text{measured}}(\lambda) = S(\lambda) F_{\text{emission}}(\lambda) \tag{20.29}$$

To recover the true fluorescence emission spectrum, the detection sensitivity of the detection system must be measured. The most common method of spectral

sensitivity calibration is to measure the spectrum of a light source with a known spectral output, such as a calibration tungsten halogen lamp:

$$F_{\text{calibrated}}(\lambda) = S(\lambda) I_{\text{lamp}}(\lambda) \tag{20.30}$$

Once the spectrum of the calibrated source is measured with the detection system, the effects of the system's sensitivity,

$$S(\lambda) = \frac{F_{\text{calibrated}}(\lambda)}{I_{\text{lamp}}(\lambda)} \tag{20.31}$$

can be removed from past and future spectra measured from unknown samples:

$$F_{\text{emission}}(\lambda) = \frac{F_{\text{measured}}(\lambda)}{S(\lambda)} = \frac{F_{\text{measured}}(\lambda)}{F_{\text{calibrated}}(\lambda)/I_{\text{lamp}}(\lambda)} = \frac{F_{\text{measured}}(\lambda) I_{\text{lamp}}(\lambda)}{F_{\text{calibrated}}(\lambda)} \tag{20.32}$$

The response of the individual components within the detection system is prone to change over time; the transmission of fiber optics and the alignment of spectrometers can shift. As such, it is recommended that any fluorescence system be calibrated on a routine basis. Calibrating the system also allows the fluorescence spectra measured across different systems with similar collection geometry to be compared.

20.8.9 Data Analysis

Fluorescence spectra by themselves typically cannot be used to predict a tissue condition, a disease state or metabolic content. Fluorescence spectroscopy ultimately relies on the statistical occurrence of consistent differences to characterize the tissue. The algorithms to be developed depend on the clinical problem under study. In most applications of fluorescence spectroscopy, the goal is to detect the presence of a particular tissue state relative to all others that can occur. For example, one might measure spectra from normal and abnormal areas of a particular tissue in a certain number of patients with suspicion of a particular disease. Based on consistent differences observed in the fluorescence spectra measured from the various tissue areas under study, diagnostic algorithms are developed. Algorithms may be developed that can (a) separate the presence of disease from all other tissue types and/or (b) separate a particular disease state from all other possible non-normal conditions. These algorithms may be simple empirical comparisons based on changes in intensity or ratios of intensities [83]. These empirical algorithms indicate the specific changes that occur in the spectra acquired and provide an indication about the biochemical processes that result in these spectral differences.

Empirical methods, however, have two important limitations. First, clinically useful diagnostic information is typically contained in more than just the few wavelengths surrounding peaks or valleys observed in tissue; a method of analysis and classification that includes all the available spectral information can potentially improve the accuracy of detection. Second, empirical algorithms are optimized for the spectra within the study. Hence, the estimates of algorithm performances will be biased towards that patient population. An unbiased estimate of the performance of the algorithms is required for an accurate evaluation of the performance of fluorescence spectroscopy. To address these limitations, multivariate statistical techniques have become routinely used to develop and evaluate algorithms that differentiate between normal, non-normal, and disease states.

In recent years, the potential for using multivariate techniques for spectroscopic data analysis in disease detection has been exploited with great success [84–88]. Discrimination techniques such as linear regression as well as classification techniques such as neural networks have been used. Following data processing, most analyses typically follow a four-step process.

(1) *Dimension reduction* – Measured spectra typically consist of intensities at as many emission wavelengths as the number of pixels in the detector. Thus, the number of variables available for algorithm development usually far exceeds the number of samples/patients in the study. And thus methods such as Principal Component Analysis (PCA) [89] are used to dimensionally reduce the data to fewer orthogonal (uncorrelated) variables while maximizing the variance.
(2) *Feature Extraction/Selection* – Many different approaches have been followed for extracting the features that maximize separation of the groups to be classified. A simple approach is to select diagnostically important principal components (output of the previous step) by using an unpaired Student's t-test to identify those components that have scores that are statistically different between the different groups [90]. Other methods include non-linear approaches such as Kernel PC and machine learning techniques [91, 92].
(3) *Classification* – This step realizes the predictive capability of the optical method as compared to the gold standard, which in clinical applications typically tends to be histopathology. Development of a classification algorithm can rely on complex methods such as neural networks [87, 93, 94] and simpler models such as nearest neighbor classifiers [85] and decision trees (such as hierarchical cluster analysis (HCA)) [95, 96]. Routine methods in spectroscopy include linear regression, logistic regression and Fisher Discriminant Analysis (FDA). FDA is a classification technique that uses linear combinations of independent variables (such as principal components) to discriminate between two or more groups such that the degree of misclassification is minimized [97]. Although this technique can be used for all data distributions, FDA yields optimal results for normally distributed data when the prior probability, which accounts for the differences in sample size of the two groups, and the cost of misclassifying a sample in a particular group are included. Another, more powerful, technique is based on logistic discrimination analysis (LDA) in which the posterior

probability that of an unknown sample belonging to each of the possible tissue categories is calculated. Thus a normal sample would have a posterior probability close to 1 for belonging in the normal group and 0 for belonging in the precancerous group. This is then plotted to give a probability distribution of each sample belonging to a particular tissue type [97]. However, more complex methods such as sparse multinomial linear regression, support vector machine systems and even genetic algorithms have also been attempted [98].

(4) *Validation* – Ultimately, any algorithm tested will yield an evaluation of the diagnostic technique under study that is only as robust as the method of validation used. Unbiased evaluation of algorithm performance may be performed using cross validation techniques such as leave-one-out, leave-five-out etc. [99, 100]. A more robust approach is obtained by splitting the data set into training and validation sets [101]. Thus, an algorithm is developed using the training set and applied to the validation set to obtain performance estimates that is unbiased and translatable to a wider population. The true test of a new diagnostic method is obtained by independent validation. The algorithm will be applied to an unbiased group of samples to independently validate the diagnostic performance of the technique.

20.9 Applications

A variety of biological molecules contain naturally occurring or endogenous fluorophores. Some of the best known fluorophores include aromatic amino acids such as tryptophan, tyrosine and phenylalanine, the cofactor – reduced nicotinamide adenine dinucleotide (NADH), vitamin A, flavins (flavin adenine dinucleotide (FAD)), riboflavins (flavin mononucleotide (FMN)) and porphyrins.

Although a complete understanding of the quantitative information contained within a tissue fluorescence spectrum has not been achieved, many groups have attempted to use fluorescence spectroscopy for automated, fast and non-invasive characterization and extent of disease [102]. It has been established that the fluorescence spectra and relaxation times of normal and diseased tissues differ. This difference is attributed to changes in the environment surrounding the fluorophore and their concentration [9]. Auto and dye induced fluorescence have been applied for the detection of atherosclerosis and various types of cancers. Exogenous fluorescent dyes, most notably hematoporphyrin derivatives (Photofrin, 5-aminoluvolinic acid), have also been shown to selectively collect in tumor tissue and have been used for enhancement of fluorescent contrast between normal and neoplastic tissue in the human lung [103], brain [104, 105] and colon [106]. More recently, molecular targeting agents such as quantum dots and nanoparticles have been developed to enhance contrast [107–110].

However, analysis using autofluorescence eliminates any likelihood of toxic reactions due to the injection of an external agent. Auto-fluorescence spectra of normal and diseased tissues have been measured from several organ sites, both in vitro and in vivo [102]. Studies which exploit spectral differences for tissue discrimination

have been carried out for the brain [111, 112], bronchus [113, 114], lung [115, 116], colon [117–119], cervix [120, 121], bladder [122, 123], esophagus [124, 125], skin [126, 127], breast [128, 129], and arterial wall [130]. These applications range from diagnosis of suspicious lesions in a clinic to guidance of therapy in the operating room.

There are many pros and cons to the application of autofluorescence for biomedicine. Many instrinsic biologic fluorophores such as NaD(P)H, flavins, porphyrins and structural proteins are known to differ with disease state. Thus autofluorescence can provide both biochemical as well as morphologic information about tissues. On a practical level, a clinically usable fluorescence system is straight-forward to assemble; such a system is compact and easily portable. The signals obtained are bright (i.e. such a system does not require highly sensitive detectors) with millisecond level integration times and can be implemented in vivo with negligible effect to the patient. Many researchers have used this ability to demonstrate that fluorescence can differentiate between normal and malignant tissues in vivo. While extensive studies continue to apply fluorescence for disease diagnosis, there are some fundamental limitations to this method. Most of these studies show that the specificity in discriminating malignant and *pre-malignant* from *non-malignant* tissues is reduced [131]. This result validated in different organs indicate that fluorescence spectra of malignant and pre-malignant tissues of, for example, the cervix and colon are similar to the fluorescence spectra of benign abnormalities such as inflammation, hyperplasia and metaplasia [131, 132], thus reducing the specificity of the technique. This limitation likely arises from the fact that only a limited number of fluorophores contribute to tissue fluorescence. Thus, discrimination based on autofluorescence relies on subtle changes in these fluororphores which, in turn, hinders differential diagnosis. Further, since most biological fluorophores fluoresce in the UV–VIS, the probing depth of these signals is limited to the superficial layers of the tissues (few hundred microns). Thus, successful application of fluorescence depends on the biomedical problem to be addressed and tissue organ/site under consideration.

The limited performance of fluorescence spectroscopy in disease diagnosis has led many researchers to add multiple wavelengths of excitation, combine different optical and non-optical techniques with fluorescence and add contrast agents with limited success. Multi-modality is recognized by many as the wave of the future. Researchers have combined fluorescence with one, two or even three other techniques in an effort to improve diagnosis. Some of these combinations include fluorescence + OCT [133], fluorescence at one λ_{exc} + diffuse reflectance [111, 126], fluorescence at multiple λ_{exc} + diffuse reflectance [28, 134], fluorescence + diffuse reflectance + light scattering [135] and others [136]. However, it is important that the selection of the different modalities be governed by the problem to be solved, not by the availability of the technologies. Researchers should consider tissue biology, physiology, and geometry, as well as the clinical question to be addressed in determining the optimal optical solution to be applied, if indeed it is even feasible.

Fluorescence has been used to its maximum benefit in microscopy with the aid of targeted contrast agents of molecules. The development of many dyes and other

agents such as green fluorescent protein, Cy3, Cy5, etc. has led to the widespread use of fluorescence microscopy in fundamental to advanced studies in molecular biology. Intrinsic or autofluorescence spectroscopy and imaging has been pursued as a clinical diagnostic tool by many researchers. The most successful application of this technique has been when differentiation is focused on normal versus malignancy, and the depth of study is limited to the superficial few hundred microns. This is not the technique to use when one needs to distinguish between subtle tissue categories. However, the potential advantages of the technique continue to make it a viable tool for specific clinical problems such as guidance of tumor resection and cancer imaging.

References

1. Krafft C and Sergo V. Biomedical applications of Raman and infrared spectroscopy to diagnose tissues. *Spectrosc. Int. J*., 20(5–6):195–218 (2006).
2. Wolf M, Ferrari M, and Quaresima V. Progress of near-infrared spectroscopy and topography for brain and muscle clinical applications. *J. Biomed. Opt*., 12(6):062104 (2007).
3. Hamaoka T, McCully KK, Quaresima V, Yamamoto K, and Chance B. Near-infrared spectroscopy/imaging for monitoring muscle oxygenation and oxidative metabolism in healthy and diseased humans. *J. Biomed. Opt*., 12(6):062105 (2007).
4. Mourant JR and Bigio IJ. Elastic scattering spectroscopy and diffuse reflectance. In: T Vo Dinh (ed) *Biomedical photonics handbook*. CRC, Boca Raton, FL, pp. 29.21–29.19 (2003).
5. Zonios G and Dimou A. Modeling diffuse reflectance from semi-infinite turbid media: Application to the study of skin optical properties. *Opt. Express*, 14(19):8661–8674 (2006).
6. Lin WC, Toms SA, Jansen ED, and Mahadevan-Jansen A. Intraoperative application of optical spectroscopy in the presence of blood. *J. Sel. Top. Quant.*, 7(6):996–1003 (2001).
7. Ferraro JR, Nakamoto K, and Brown CW. *Introductory Raman spectroscopy*. 2nd ed., Academic, Amsterdam, Boston (2003).
8. Mahadevan-Jansen A. Raman spectroscopy: From benchtop to bedside. In: T Vo Dinh (ed) *Biomedical photonics handbook*, CRC, Boca Raton, FL, pp. 30.31–30.27 (2003).
9. Lakowicz JR. *Principles of fluorescence spectroscopy*. 3rd ed., Springer, New York (2006).
10. Mishra AK. Photoluminescence spectroscopy. In: T Pradeep (ed) *Advances in physical chemistry*. Allied, Mumbai, pp. 195–214 (1999).
11. Fotinos N, Campo MA, Popowycz F, Gurny R, and Lange N. 5-aminolevulinic acid derivatives in photomedicine: Characteristics, application and perspectives. *Photochem. Photobiol.*, 82(4):994–1015 (2006).
12. Allison RR and Sibata CH. Photodiagnosis for cutaneous malignancy: A brief clinical and technical review. *Photodiagn. Photodyn.*, 5(4):247–250 (2008).
13. Jichlinski P, Guillou L, Karlsen SJ, Malmstrom PU, Jocham D, Brennhovd B, Johansson E, Gartner T, Lange N, van den Bergh H, and Leisinger HJ. Hexyl aminolevulinate fluorescence cystoscopy: A new diagnostic tool for the photodiagnosis of superficial bladder cancer – A multicenter study. *J. Urol.*, 170(1):226–229 (2003).
14. Kelty CJ, Brown NJ, Reed MWR, and Ackroyd R. The use of 5-aminolaevulinic acid as a photosensitiser in photodynamic therapy and photodiagnosis. *Photochem. Photobiol. Sci*., 1(3):158–168 (2002).
15. Ferrari M. Cancer nanotechnology: Opportunities and challenges. *Nat. Rev. Cancer*, 5(3):161–171 (2005).
16. Pope-Harman A, Cheng MMC, Robertson F, Sakamoto J, and Ferrari M. Biomedical nanotechnology for cancer. *Med. Clin. North Am.*, 91(5):899–927 (2007).

17. Jayagopal A, Russ PK, and Haselton FR. Surface engineering of quantum dots for in vivo vascular Imaging. *Bioconjug. Chem.*, 18(5):1424–1433 (2007).
18. Drbohlavova J, Adam V, Kizek R, and Hubalek J. Quantum dots – characterization, preparation and usage in biological systems. *Int. J. Mol. Sci*., 10(2):656–673 (2009).
19. Volynskaya Z, Haka AS, Bechtel KL, Fitzmaurice M, Shenk R, Wang N, Nazemi J, Dasari RR, and Feld MS. Diagnosing breast cancer using diffuse reflectance spectroscopy and intrinsic fluorescence spectroscopy. *J. Biomed. Opt*., 13(2):024012 (2008).
20. Zhu CF, Palmer GM, Breslin TM, Harter J, and Ramanujam N. Diagnosis of breast cancer using fluorescence and diffuse reflectance spectroscopy: A Monte-Carlo-model-based approach. *J. Biomed. Opt*., 13(3):034015 (2008).
21. Muller MG, Georgakoudi I, Zhang QG, Wu J, and Feld MS. Intrinsic fluorescence spectroscopy in turbid media: disentangling effects of scattering and absorption. *Appl. Opt.*, 40(25):4633–4646 (2001).
22. Wu J, Feld MS, and Rava RP. Analytical model for extracting intrinsic fluorescence in turbid media. *Appl. Opt.*, 32(19):3585–3595 (1993).
23. Gardner CM, Jacques SL, and Welch AJ. Fluorescence spectroscopy of tissue: Recovery of intrinsic fluorescence from measured fluorescence. *Appl. Opt.*, 35(10):1780–1792 (1996).
24. Zhadin NN and Alfano RR. Correction of the internal absorption effect in fluorescence emission spectra from absorbing and highly scattering media: theory and experiment. *J. Biomed. Opt*., 3(2):171–186 (1998).
25. Utzinger U and Richards-Kortum RR. Fiber optic probes for biomedical optical spectroscopy. *J. Biomed. Opt*., 8(1):121–147 (2003).
26. Zhu CF, Palmer GM, Breslin TM, Xu FS, and Ramanujam N. Use of a multiseparation fiber optic probe for the optical diagnosis of breast cancer. *J. Biomed. Opt*., 10(2):024032 (2005).
27. Nieman LT, Jakovljevic M, and Sokolov K. Compact beveled fiber optic probe design for enhanced depth discrimination in epithelial tissues. *Opt. Express*, 17(4):2780–2796 (2009).
28. Schwarz RA, Gao W, Daye D, Williams MD, Richards-Kortum R, and Gillenwater AM. Autofluorescence and diffuse reflectance spectroscopy of oral epithelial tissue using a depth-sensitive fiber-optic probe. *Appl. Opt.*, 47(6):825–834 (2008).
29. Glenar DA, Samuelson RE, Pearl JC, Bjoraker GL, and Blaney D. Spectral imaging of martian water ice clouds and their diurnal behavior during the 1999 aphelion season (L-s=130 degrees). *Icarus*, 161(2):297–318 (2003).
30. Kim EE, Choi YW, Yang HS, Kang MS, Jeong SK, Yang SU, Kim ED, Lee JH, Rasheed AAA, Nasir HM, and Rosdi MRM. A high-resolution multi-spectral imaging system for small satellites. *Acta Astronaut.*, 52(9–12):813–818 (2003).
31. Miskey CL and Bruhweiler FC. STIS spectral imagery of the ob stars in NGC 604. I. Description of the extraction technique for a crowded stellar field. *Astron. J.*, 125(6):3071–3081 (2003).
32. Galand M, Lummerzheim D, Stephan AW, Bush BC, and Chakrabarti S. Electron and proton aurora observed spectroscopically in the far ultraviolet. *J. Geophys. Res.-Space Phys*., 107(A7): SIA4.1–SIA4.4 (2002).
33. Kerekes JP and Baum JE. Spectral imaging system analytical model for subpixel object detection. *IEEE Trans. Geosci. Remote Sensing*, 40(5):1088–1101 (2002).
34. White SN, Chave AD, Reynolds GT, and Van Dover CL. Ambient light emission from hydrothermal vents on the Mid-Atlantic Ridge. *Geophys. Res. Lett.*, 29(15):34.1–34.4 (2002).
35. Kulcke A, Gurschler C, Spock G, Leitner R, and Kraft M. On-line classification of synthetic polymers using near infrared spectral imaging. *J. Near Infrared Spectrosc*., 11(1):71–81 (2003).
36. Howard J, Carraro L, Puiatti ME, Sattin F, Scarin P, Valisa M, Zaniol B, Konig R, and Chung J. Coherence-imaging approach to time-resolved charge-exchange recombination spectroscopy in high-temperature plasma. *Rev. Sci. Instrum.*, 74(3):2060–2063 (2003).
37. Demers DR, Schroch PM, Radke RJ, Anderson JK, Craig D, and Den Hartog DJ. Spectroscopic ion beam imaging for investigations into magnetic field mapping of a plasma. *Rev. Sci. Instrum.*, 74(3):2103–2106 (2003).

38. Slovak ML, Zhang FY, Tcheurekdjian L, Bobadilla D, Bedell V, Arber DA, Persons DL, Sosman JA, and Murata-Collins JL. Targeting multiple genetic aberrations in isolated tumor cells by spectral fluorescence in situ hybridization. *Cancer Detect. Prev.*, 26(3):171–179 (2002).
39. Levenson RM. Spectral imaging and pathology: Seeing more. *Lab. Med.*, 35(4):244–252 (2004).
40. Padilla-Nash HM, Barenboim-Stapleton L, Difilippantonio MJ, and Reid T. Spectral karyotyping analysis of human and mouse chromosomes *Nat. protoc*., 1(6):3129–3142 (2006).
41. Weier HUG, Munne S, Lersch RA, Hsieh HB, Smida J, Chen XN, Korenberg JR, Pedersen RA, and Fung J. Towards a full karyotype screening of interphase cells: 'FISH and chip' technology. *Mol. Cell. Endocrinol.*, 183:S41–S45 (2001).
42. Maiti D, Sennoune S, and Martinez-Zaguilan R. Proton gradients in human breast cancer cells determined by confocal, multiphoton, and spectral imaging microscopy. *FASEB J.*, 17(4):A467–A467 (2003).
43. Papadakis AN, Stathopoulos E, Delides G, Berberides K, Nikiforidis G, and Balas C. A novel spectral microscope system: Application in quantitative pathology. *IEEE Trans. Biomed. Eng.*, 50(2):207–217 (2003).
44. Synytsya A, Kral V, Synytsya A, Volka K, and Sessler JL. In vitro interaction of macrocyclic photosensitizers with intact mitochondria: A spectroscopic study. *Biochimica Et Biophysica Acta-General Subjects*, 1620(1–3):85–96 (2003).
45. Greenbaum L, Schwartz D, and Malik Z. Spectrally resolved microscopy of GFP trafficking. *J. Histochem. Cytochem.*, 50(9):1205–1212 (2002).
46. Haraguchi T, Shimi T, Koujin T, Hashiguchi N, and Hiraoka Y. Spectral imaging fluorescence microscopy. *Genes Cells* 7(9):881–887 (2002).
47. Qu JNY, Chang HP, and Xiong SM. Fluorescence spectral imaging for characterization of tissue based on multivariate statistical analysis. *J. Opt. Soc. Am. a-Opt. Image Sci. Vision*, 19(9):1823–1831 (2002).
48. Ramanujam N, Chen JX, Gossage K, Richards-Kortum R, and Chance B. Fast and noninvasive fluorescence imaging of biological tissues in vivo using a flying-spot scanner. *IEEE Trans. Biomed. Eng.*, 48(9):1034–1041 (2001).
49. Lazarev VV, Roth RA, Kazakevich Y, and Hang J. Detection of premalignant oral lesions in hamsters with an endoscopic fluorescence imaging system. *Cancer*, 85(7):1421–1429 (1999).
50. Andersson-Engels S, Canti G, Cubeddu R, Eker C, af Klinteberg C, Pifferi A, Svanberg K, Svanberg S, Taroni P, Valentini G, and Wang I. Preliminary evaluation of two fluorescence imaging methods for the detection and the delineation of nasal cell carcinomas of the skin. *Lasers Surg. Med.*, 26(1):76–82 (2000).
51. Hewett J, Nadeau V, Ferguson J, Moseley H, Ibbotson S, Allen JW, Sibbett W, and Padgett M. The application of a compact multispectral imaging system with integrated excitation source to in vivo monitoring of fluorescence during topical photodynamic therapy of superficial skin cancers. *Photochem. Photobiol.*, 73(3):278–282 (2001).
52. Nadeau RG and Groner W. The role of a new noninvasive imaging technology in the diagnosis of anemia. *J. Nutr.*, 131(5):1610s–1614s (2001).
53. Thomale UW, Schaser KD, Unterberg AW, and Stover JF. Visualization of rat pial microcirculation using the novel orthogonal polarized spectral (OPS) imaging after brain injury. *J. Neurosci. Methods* 108(1):85–90 (2001).
54. Schultz RA, Nielsen T, Zavaleta JR, Ruch R, Wyatt R, and Garner HR. Hyperspectral imaging: A novel approach for microscopic analysis. *Cytometry*, 43(4):239–247 (2001).
55. Schlucker S, Schaeberle MD, Huffman SW, and Levin IW. Raman microspectroscopy: A comparison of point, line, and wide-field imaging methodologies. *Anal. Chem.*, 75(16):4312–4318 (2003).
56. Otto C and Greve J. Progress in instrumentation for Raman micro-spectroscopy and Raman imaging for cellular biophysics. *Internet J. Vibrational Spectrosc*., 2(3) (1998).

57. Khait O, Smirnov S, and Tran CD. Multispectral imaging microscope with millisecond time resolution. *Anal. Chem.*, 73(4):732–739 (2001).
58. Tran CD. Development and analytical applications of multispectral imaging techniques: An overview. *Fresenius J. Anal. Chem.*, 369(3–4):313–319 (2001).
59. Balas C. A novel optical imaging method for the early detection, quantitative grading, and mapping of cancerous and precancerous lesions of cervix. *IEEE Trans. Biomed. Eng.*, 48(1):96–104 (2001).
60. Garini Y, Katzir N, Cabib D, Buckwald RA, Soenksen DG, and Malik Z. Spectral bio-imaging. In: XF Wang and B Herman (eds) *Fluorescence imaging spectroscopy and microscopy*. Wiley, New York (1996).
61. Qu JNY, Huang ZJ, and Hua JW. Excitation-and-collection geometry insensitive fluorescence imaging of tissue-simulating turbid media. *Appl. Opt.*, 39(19):3344–3356 (2000).
62. Churmakov DY, Meglinski IV, and Greenhalgh DA. Influence of refractive index matching on the photon diffuse reflectance. *Phys. Med. Biol.*, 47(23):4271–4285 (2002).
63. Wu J, Partovi F, Field MS, and Rava RP. Diffuse reflectance from turbid media – an analytical model of photon migration. *Appl. Opt.*, 32(7):1115–1121 (1993).
64. Pfefer TJ, Matchette LS, Bennett CL, Gall JA, Wilke JN, Durkin AJ, and Ediger MN. Reflectance-based determination of optical properties in highly attenuating tissue. *J. Biomed. Opt*., 8(2):206–215 (2003).
65. Bevilacqua F, Piguet D, Marquet P, Gross JD, Tromberg BJ, and Depeursinge C. In vivo local determination of tissue optical properties: Applications to human brain. *Appl. Opt.*, 38(22):4939–4950 (1999).
66. Bargo PR, Prahl SA, and Jacques SL. Optical properties effects upon the collection efficiency of optical fibers in different probe configurations. *J. Sel. Top. Quant.*, 9(2):314–321 (2003).
67. Bargo PR, Prahl SA, and Jacques SL. Collection efficiency of a single optical fiber in turbid media. *Appl. Opt.*, 42(16):3187–3197 (2003).
68. Liu Q, Zhu CF, and Ramanujam N. Experimental validation of Monte Carlo modeling of fluorescence in tissues in the UV-visible spectrum. *J. Biomed. Opt*., 8(2):223–236 (2003).
69. Schmitt JM and Kumar G. Spectral distortions in near-infrared spectroscopy of turbid materials. *Appl. Spec.*, 50(8):1066–1073 (1996).
70. Kumar G and Schmitt JM. Optimal probe geometry for near-infrared spectroscopy of biological tissue. *Appl. Opt*., 36(10):2286–2293 (1997).
71. Welch AJ, Gardner C, RichardsKortum R, Chan E, Criswell G, Pfefer J, and Warren S. Propagation of fluorescent light. *Lasers Surg. Med.*, 21(2):166–178 (1997).
72. Pfefer TJ, Matchette LS, and Drezek R. Influence of illumination-collection geometry on fluorescence spectroscopy in multilayer tissue. *Med. Biol. Eng. Comput*., 42(5):669–673 (2004).
73. Pfefer TJ, Schomacker KT, Ediger MN, and Nishioka NS. Multiple-fiber probe design for fluorescence spectroscopy in tissue. *Appl. Opt.*, 41(22):4712–4721 (2002).
74. Zhu CF, Liu Q, and Ramanujam N. Effect of fiber optic probe geometry on depth-resolved fluorescence measurements from epithelial tissues: A Monte Carlo simulation. *J. Biomed. Opt*., 8(2):237–247 (2003).
75. Chang SK, Arifler D, Drezek R, Follen M, and Richards-Kortum R. Analytical model to describe fluorescence spectra of normal and preneoplastic epithelial tissue: Comparison with Monte Carlo simulations and clinical measurements. *J. Biomed. Opt*., 9(3):511–522 (2004).
76. Durkin AJ, Jaikumar S, Ramanujam N, and Richardskortum R. Relation between fluorescence-spectra of dilute and turbid samples. *Appl. Opt.*, 33(3):414–423 (1994).
77. Zhang QG, Muller MG, Wu J, and Feld MS. Turbidity-free fluorescence spectroscopy of biological tissue. *Opt. Lett*., 25(19):1451–1453 (2000).
78. Richards-Kortum R, Mehta A, Hayes G, Cothren R, Kolubayev T, Kittrell C, Ratliff NB, Kramer JR, and Feld MS. Spectral diagnosis of atherosclerosis using an optical fiber laser catheter. *Am. Heart J.*, 118(2):381–391 (1989).

79. Keijzer M, Richardskortum RR, Jacques SL, and Feld MS. Fluorescence spectroscopy of turbid media – autofluorescence of the human aorta. *Appl. Opt.*, 28(20):4286–4292 (1989).
80. Avrillier S, Tinet E, Ettori D, Tualle JM, and Gelebart B. Influence of the emission-reception geometry in laser-induced fluorescence spectra from turbid media. *Appl. Opt.*, 37(13):2781–2787 (1998).
81. Swartling J, Svensson J, Bengtsson D, Terike K, and Andersson-Engels S. Fluorescence spectra provide information on the depth of fluorescent lesions in tissue. *Appl. Opt.*, 44(10):1934–1941 (2005).
82. Gebhart SC, Mahadevan-Jansen A, and Lin WC. Experimental and simulated angular profiles of fluorescence and diffuse reflectance emission from turbid media. *Appl. Opt.*, 44(23):4884–4901 (2005).
83. Ramanujam N, Mitchell MF, Mahadevan A, Thomsen S, Silva E, and Richardskortum R. Fluorescence spectroscopy – a diagnostic-tool for cervical intraepithelial neoplasia (cin). *Gynecol. Oncol.*, 52(1):31–38 (1994).
84. Majumder SK, Gupta A, Gupta S, Ghosh N, and Gupta PK. Multi-class classification algorithm for optical diagnosis of oral cancer. *J. Photochem. Photobiol. B-Biol.*, 85(2):109–117 (2006).
85. Kamath SD, Bhat RA, Ray S, and Mahato KK. Autofluorescence of normal, benign, and malignant ovarian tissues: A pilot study. *Photomed. Laser Surg.*, 27(2):325–335 (2009).
86. Wang GF, Platz CP, and Geng ML. Probability-based differential normalized fluorescence bivariate analysis for the classification of tissue autofluorescence spectra. *Appl. Spec.*, 60(5):545–550 (2006).
87. Wang LQ and Mizaikoff B. Application of multivariate data-analysis techniques to biomedical diagnostics based on mid-infrared spectroscopy. *Anal. Bioanal. Chem.*, 391(5): 1641–1654 (2008).
88. Escandar GM, Damiani PC, Goicoechea HC, and Olivieri AC. A review of multivariate calibration methods applied to biomedical analysis. *Microchem. J.* 82(1):29–42 (2006).
89. Dillon RW and Goldstein M. *Multivariate analysis: Methods and applications.* Wiley, New York (1984).
90. Ramanujam N, Mitchell MF, Mahadevan A, Thomsen S, Malpica A, Wright T, Atkinson N, and Richards-Kortum R. Development of a multivariate statistical algorithm to analyze human cervical tissue fluorescence spectra acquired in vivo. *Lasers Surg. Med.*, 19(1):46–62 (1996).
91. Sajda P. Machine learning for detection and diagnosis of disease. *Annu. Rev. Biomed. Eng.*, 8:537–565 (2006).
92. McCarthy JF, Marx KA, Hoffman PE, Gee AG, O'Neil P, Ujwal ML, and Hotchkiss J. Applications of machine learning and high-dimensional visualization in cancer detection, diagnosis, and management. *Appl. Bioinform. Canc. Det.*, 1020:239–262 (2004).
93. Sigurdsson S, Philipsen PA, Hansen LK, Larsen J, Gniadecka M, and Wulf HC. Detection of skin cancer by classification of Raman spectra. *IEEE Trans. Biomed. Eng.*, 51(10): 1784–1793 (2004).
94. Carrara M, Bono A, Bartoli C, Colombo A, Lualdi M, Moglia D, Santoro N, Tolomio E, Tomatis S, Tragni G, Santinami M, and Marchesini R. Multispectral imaging and artificial neural network: Mimicking the management decision of the clinician facing pigmented skin lesions. *Phys. Med. Biol.*, 52(9):2599–2613 (2007).
95. Baek SJ, Park A, Kim JY, Na SY, Won Y, and Choo J. Detection of basal cell carcinoma by automatic classification of Confocal Raman spectra. *Lect. Notes Comput. Sci.*, 4115: 402–411 (2006).
96. de Veld DCG, Schut TCB, Skurichina M, Witjes MJH, Van der Wal JE, Roodenburg JLN, and Sterenborg HJCM. Autofluorescence and Raman microspectroscopy of tissue sections of oral lesions. *Lasers Med. Sci.*, 19(4):203–209 (2005).

97. Albert A and Harris EK. *Multivariate interpretation of clinical laboratory data.* Marcel Dekker, New York (1987).
98. Halter RJ, Hartov A, Paulsen KD, Schned A, and Heaney J. Genetic and least squares algorithms for estimating spectral EIS parameters of prostatic tissues. *Physiol. Meas.*, 29(6):S111–S123 (2008).
99. Majumder SK, Keller MD, Boulos FI, Kelley MC, and Mahadevan-Jansen A. Comparison of autofluorescence, diffuse reflectance, and Raman spectroscopy for breast tissue discrimination. *J. Biomed. Opt*., 13(5):054009 (2008).
100. Smit S, Hoefsloot HCJ, and Smilde AK. Statistical data processing in clinical proteomics. *J. Chromatogr. B-Anal. Technol. Biomed. Life Sci.*, 866(1–2):77–88 (2008).
101. Diana G and Tommasi C. Cross-validation methods in principal componant analysis: A comparison. *Stat. Methods Appl*., 11:71–82 (2002).
102. Ramanujam N. Fluorescence spectroscopy of neoplastic and non-neoplastic tissues. *Neoplasia*, 2(1–2):89–117 (2000).
103. Baas P, Triesscheijn M, Burgers S, van Pel R, Stewart F, and Aalders M. Fluorescence detection of pleural malignancies using 5-aminolaevulinic acid. *Chest*, 129(3):718–724 (2006).
104. Pichlmeier U, Bink A, Schackert G, Stummer W, and Grp AGS. Resection and survival in glioblastoma multiforme: An RTOG recursive partitioning analysis of ALA study patients. *Neuro-Oncol*., 10(6):1025–1034 (2008).
105. Eljamel MS, Goodman C, and Moseley H. ALA and Photofrin (R) fluorescence-guided resection and repetitive PDT in glioblastoma multiforme: A single centre phase III randomised controlled trial. *Lasers Med. Sci*., 23(4):361–367 (2008).
106. Xia DL, He JS, Zhang YD, Tang JT, Zhang B, Wang SC, and Huang QL. Study of a multivariate statistical 5-ALA-based discrimination method for fluorescence spectra of colonic tissue of SD rats. *Spectrosc. Spectral Anal*., 25(12):2029–2033 (2005).
107. Yang SJ, Shieh MJ, Lin FH, Lou PJ, Peng CL, Wei MF, Yao CJ, Lai PS, and Young TH. Colorectal cancer cell detection by 5-aminolaevulinic acid-loaded chitosan nano-particles. *Cancer Lett.*, 273(2):210–220 (2009).
108. Bakalova R, Zhelev Z, and Gadjeva V. Quantum dots versus organic fluorophores in fluorescent deep-tissue imaging – merits and demerits. *Gen. Physiol. Biophys.*, 27(4):231–242 (2008).
109. Xing Y and Rao JH. Quantum dot bioconjugates for in vitro diagnostics & in vivo imaging. *Cancer Biomarkers*, 4(6):307–319 (2008).
110. Lee H, Lee K, Kim IK, and Park TG. Synthesis, characterization, and in vivo diagnostic applications of hyaluronic acid immobilized gold nanoprobes. *Biomaterials*, 29(35): 4709–4718 (2008).
111. Lin WC, Toms SA, Johnson M, Jansen ED, and Mahadevan-Jansen A. In vivo brain tumor demarcation using optical spectroscopy. *Photochem. Photobiol.*, 73(4):396–402 (2001).
112. Bottiroli G, Croce AC, Locatelli D, Nano R, Giombelli E, Messina A, and Benericetti E. Brain tissue autofluorescence: An aid for intraoperative delineation of tumor resection margins. *Cancer Detect. Prev.*, 22(4):330–339 (1998).
113. Ikeda N, Honda H, Hayashi A, Usuda J, Kato Y, Tsuboi M, Ohira T, Hirano T, Kato H, Serizawa H, and Aoki Y. Early detection of bronchial lesions using newly developed videoendoscopy-based autofluorescence bronchoscopy. *Lung Cancer*, 52(1):21–27 (2006).
114. Fawzy Y and Zeng HZ. Intrinsic fluorescence spectroscopy for endoscopic detection and localization of the endobronchial cancerous lesions. *J. Biomed. Opt*., 13(6):064022 (2008).
115. Ueno K, Kusunoki Y, Imamura F, Yoshimura M, Yamamoto S, Uchida J, and Tsukamoto Y. Clinical experience with autofluorescence imaging system in patients with lung cancers and precancerous lesions. *Respiration*, 74(3):304–308 (2007).
116. Huttenberger D, Gabrecht T, Wagnieres G, Weber B, Linder A, Foth HJ, and Freitag L. Autofluorescence detection of tumors in the human lung – spectroscopical measurements in situ, in an in vivo model and in vitro. *Photodiagn. Photodyn.*, 5(2):139–147 (2008).

117. Aihara H, Sumiyama K, Saito S, Tajiri H, and Ikegami M. Numerical analysis of the autofluorescence intensity of neoplastic and non-neoplastic colorectal lesions by using a novel videoendoscopy system. *Gastrointest. Endosc.*, 69(3):726–733 (2009).
118. DaCosta RS, Wilson BC, and Marcon NE. Optical techniques for the endoscopic detection of dysplastic colonic lesions. *Curr. Opin. Gastroenterol*., 21(1):70–79 (2005).
119. McCallum AL, Jenkins JT, Gillen D, and Molloy RG. Evaluation of autofluorescence colonoscopy for the detection and diagnosis of colonic polyps. *Gastrointest. Endosc.*, 68(2):283–290 (2008).
120. Huh WK, Cestero RM, Garcia FA, Gold MA, Guido RS, McIntyre-Seltman K, Harper DM, Burke L, Sum ST, Flewelling RF, and Alvarez RD. Optical detection of high-grade cervical intraepithelial neoplasia in vivo: Results of a 604-patient study. *Am. J. Obstet. Gynecol.*, 190(5):1249–1257 (2004).
121. Chang SK, Marin N, Follen M, and Richards-Kortum R. Model-based analysis of clinical fluorescence spectroscopy for in vivo detection of cervical intraepithelial dysplasia. *J. Biomed. Opt*., 11(2):024008 (2006).
122. Szygula M, Wojciechowski B, Adamek M, Kawczyk-Krupka A, Cebula W, Zieleznik W, Biniszkiewicz T, Pietrusa A, and Sieron A. Photodynamic vs. autofluorescent diagnosis of urinary bladder using Xillix life system. *Physica Medica*, 20:55–57 (2004).
123. D'Hallewin MA, Bezdetnaya L, and Guillemin F. Fluorescence detection of bladder cancer: A review. *Eur. Urol.*, 42(5):417–425 (2002).
124. Falk GW, Oldenburgh M, Schoenfield L, and Lopez R. Autofluorescence endoscopy is inferior to white light endoscopy for detection of dysplasia/cancer in Barrett's esophagus surveillance. *Gastrointest. Endosc.*, 67(5):Ab174–Ab174 (2008).
125. Borovicka J, Fischer J, Neuweiler J, Netzer P, Gschossmann J, Ehmann T, Bauerfeind P, Dorta G, Zurcher U, Binek J, and Meyenberger C. Autofluorescence endoscopy in surveillance of Barrett's esophagus: a multicenter randomized trial on diagnostic efficacy. *Endoscopy*, 38(9):867–872 (2006).
126. Drakaki E, Kaselouris E, Makropoulou M, Serafetinides AA, Tsenga A, Stratigos AJ, Katsambas AD, and Antoniou C. Laser-induced fluorescence and reflectance spectroscopy for the discrimination of basal cell carcinoma from the surrounding normal skin tissue. *Skin Pharmacol. Physiol*., 22(3):158–165 (2009).
127. Borisova EG, Troyanova PP, and Avramov LA. Fluorescence spectroscopy for early detection and differentiation of cutaneous pigmented lesions. *Optoelectron. Adv. Mater.-Rapid Comm*., 1(8):388–393 (2007).
128. Chowdary MVP, Mahato KK, Kumar KK, Mathew S, Rao L, Krishna CM, and Kurien J. Autofluorescence of breast tissues: Evaluation of discriminating algorithms for diagnosis of normal, benign, and malignant conditions. *Photomed. Laser Surg*., 27(2):241–252 (2009).
129. Breslin TM, Xu FS, Palmer GM, Zhu CF, Gilchrist KW, and Ramanujam N. Autofluorescence and diffuse reflectance properties of malignant and benign breast tissues. *Ann. Surg. Oncol.*, 11(1):65–70 (2004).
130. van de Poll SWE, Dasari RR, and Kramer JR. The role of laser-induced fluorescence spectroscopy in the detection of human atherosclerosis. *Curr. Sci.*, 77(7):934–941 (1999).
131. Ramanujam N, Mitchell MF, MahadevanJansen A, Thomsen SL, Staerkel G, Malpica A, Wright T, Atkinson N, and RichardsKortum R. Cervical precancer detection using a multivariate statistical algorithm based on laser-induced fluorescence spectra at multiple excitation wavelengths. *Photochem. Photobiol.*, 64(4):720–735 (1996).
132. Schomacker KT, Frisoli JK, Compton CC, Flotte TJ, Richter JM, Nishioka NS, and Deutsch TF. Ultraviolet laser-induced fluorescence of colonic tissue – basic biology and diagnostic potential. *Lasers Surg. Med.*, 12(1):63–78 (1992).
133. Hariri LP, Tumlinson AR, Besselsen DG, Utzinger U, Gerner EW, and Barton JK. Endoscopic optical coherence tomography and laser-induced fluorescence spectroscopy in a murine colon cancer model. *Lasers Surg. Med.*, 38(4):305–313 (2006).

134. Palmer GM, Zhu CF, Breslin TM, Xu FS, Gilchrist KW, and Ramanujam N. Comparison of multiexcitation fluorescence and diffuse reflectance spectroscopy for the diagnosis of breast cancer (March 2003). *IEEE Trans. Biomed. Eng.*, 50(11):1233–1242 (2003).
135. Georgakoudi I, Sheets EE, Muller MG, Backman V, Crum CP, Badizadegan K, Dasari RR, and Feld MS. Trimodal spectroscopy for the detection and characterization of cervical precancers in vivo. *Am. J. Obstet. Gynecol.*, 186(3):374–382 (2002).
136. van den Broek FJC, Fockens P, van Eeden S, Kara MA, Hardwick JCH, Reitsma JB, and Dekker E. Clinical evaluation of endoscopic trimodal imaging for the detection and differentiation of colonic polyps. *Clin. Gastroenterol. Hepatol*., 7(3):288–295 (2009).

Chapter 21
Molecular Imaging Using Fluorescence and Bioluminescence to Reveal Tissue Response to Laser-Mediated Thermal Injury

Mark A. Mackanos, E. Duco Jansen, and Christopher H. Contag

21.1 Overview

For decades biological investigation has focused on a reductionist approach, which has greatly advanced our understanding of the biological process, but has also served to move the analysis further and further away from the living body. This was necessary as we sought to identify the cells, genes, mutations and/or etiological agents that were associated with a given process. The information generated through these approaches can now be used to advance more integrative strategies in which specific cellular and molecular events can be studied in the context of the functional circulation and intact organ systems of living animals, and humans. Essential tools for integrative analyses of biology include imaging modalities that enable visualization of structure and function in the living body. The relatively recent development of molecular probes as exogenous contrast agents and reporter genes that encode proteins with unique properties that can be distinguished from tissues and cells has ushered in a new set of approaches that are being called molecular imaging.

The field of molecular imaging has emerged as a convergence of cellular and molecular biology with imaging science, and is aimed at direct interrogation of biological function in living animals. One strategy within the molecular imaging tool box is the use of genes that encode proteins with unique optical properties that make them distinguishable from mammalian cells and tissues. These proteins can be enzymes that convert substrates and (chemical) energy to light or they can be fluorescent, each providing a detectable signature. The enzymatic production of light is referred to as bioluminescence and the enzymes themselves are known as luciferases. This chapter discusses the general concepts of optical molecular imaging and focuses on the use of such optically unique proteins to study the response of tissues to laser exposure.

M.A. Mackanos (✉)
Department of Biomedical Engineering, Vanderbilt University, Nashville, TN, USA
e-mail: mark.mackanos@vanderbilt.edu

A.J. Welch, M.J.C. van Gemert (eds.), *Optical-Thermal Response of Laser-Irradiated Tissue*, 2nd ed., DOI 10.1007/978-90-481-8831-4_21,

21.2 In Vivo Bioluminescence Imaging (BLI)

Luciferases are biological sources of light from organisms such as fireflies and jelly fish [1]. These enzymes have been used for several decades as reporter genes in cell culture and as bioassays due to their extraordinary sensitivity of detection. More recently they have been used in living animals for in vivo bioluminescence imaging (BLI). Since BLI was first described in mammals over a decade ago [2, 3], its utility for revealing biological function has been demonstrated in multiple areas of biomedical research [4] in which human biology and disease can be modeled in small animals. This imaging approach is well-suited for small laboratory animals where optical probing depth requirements are typically no more than 2–3 cm and thus light absorption and scattering in tissue is less of an issue than in larger animals and humans. Luciferases as a class of enzymes require chemical substrates for signal generation, and, as such, systemic administration of these compounds is required for generating an in vivo signal. Thus, translation of BLI to the clinic is not readily feasible and may only occur in small niches of medicine. However, due to its outstanding sensitivity, BLI has developed into a well-accepted tool for preclinical studies where continuous spatiotemporal monitoring of biological processes in living animals is enabled.

BLI has been applied successfully to the study of a wide range of biological processes including: infection, gene expression, host response to infection, hematopoietic reconstitution, graft-versus-host disease, assessment of tumor burden and growth kinetics, neovascularization, response to chemotherapy, tumor immunology, stem cell biology and tissue regeneration [4–6]. BLI is based on the expression of reporter genes that encode bioluminescent proteins – luciferases. Since these enzymes serve as optical reporters that can be built into cultured cells or incorporated as transgenes into animals, they can be used to mark and report on gene function or targeted cells. In this manner, new robust and informative animal models can be developed that can be studied noninvasively with reasonable spatial resolution, very good temporal resolution and excellent sensitivity. BLI can be superimposed on existing animal models without major modifications, and can be used to guide the investigator to the right times and tissues for analyses of more conventional biomolecular assays. Imaging of bioluminescent reporters in animal models of human biology and disease serves to advance the study of mammalian biology by refining the models and accelerating the development of novel therapies with rapid in vivo functional assays.

The use of BLI as a noninvasive assay is increasing with the development of more targeted reporter enzymes and these advances have been recently and comprehensively reviewed elsewhere [4–8]. Thus, we will only briefly review the underlying principles, advantages and limitations of BLI, and then focus the discussion on the use of BLI to assess tissue responses to thermal stress, including in vivo temporal analysis of expression patterns of stress response genes. Advances in BLI, including modified reporter genes, multifunctional reporter proteins, three-dimensional image reconstruction, multispectral imaging, and validated transgenic reporter mice, have been improving the versatility of BLI, and the yield and accuracy of the information

obtained [5, 9–13]. Thus, this imaging modality is becoming a standard approach in biomedical research with broad applications in mammalian development and physiology, gene expression patterns, pathophysiology, and therapeutic responses.

21.2.1 Principles of In Vivo Bioluminescence Imaging

Light-emitting enzymes, such as the luciferases derived from the firefly (Photinus pyralis), the sea pansy (*Renilla reniformis*), the copepod crustacean (*Gaussia princeps*), and the jellyfish (*Aequoria victoria*) emit visible light (λ_{max} 490–612 nm) and have been used as exquisitely sensitive reporters of gene expression, and cell physiology in cultured cells [14–18]. The luciferase enzyme catalyzes the oxidation of the substrate luciferin in the presence of adenosine triphosphate (ATP), Magnesium (Mg^{2+}), and molecular oxygen [19]. The oxidation of luciferin forms oxyluciferin in an electronically excited state [1]. When the oxyluciferin returns to the ground state, broadband light is emitted from 500 to 700 nm with a peak of emission at 612 nm at 37°C [20]. One single photon is emitted for each oxidized luciferin molecule [21]. Given excess luciferin and ATP, the amount of luciferase is rate limiting and thus emitted light is proportional to the concentration of luciferase protein [22].

Since visible light can traverse living mammalian tissues with limited, but predictable penetration based on the optical properties of mammalian tissues [20, 23–26], these enzymes can be used as internal biological sources of light and deployed much the same as they are used in cell culture. Engineered genes encoding luciferases can be introduced into mammals and their emission can be detected externally without removing the overlying tissues using imaging systems that consist of lenses, filters and charged-couple device (CCD) cameras [2–5, 27]. Sensitive CCD cameras initially developed for the field of astronomy that operate in the visible to near infrared regions of the spectrum (wavelength range of 300–1100 nm) have found new uses with the growing field of in vivo optical imaging. To reduce detector noise and increase sensitivity, a number of camera architectures have been developed [28].

Traditionally, intensified charge-coupled device (ICCD) and cooled charge-coupled device (CCD) cameras have been the most effective imaging systems for collection of both fluorescence and bioluminescence in vitro and in vivo [9, 20, 29, 30]. A back illuminated cooled CCD (−90°C) is the current standard for bioluminescent imaging and analysis because it has much greater quantum efficiency, on the order of 85%, at 650 nm when compared to an ICCD system which only has a 1% quantum efficiency at 650 nm [20]. In addition, the signal to noise ratio is also improved with the cooled CCD since it does not need a photon intensifier and the dark-noise is reduced by the CCD cooling [20, 29].

More recently, a new electron multiplying charge-coupled device (EMCCD) has been developed which allows for multiplying the electron counts while still on the chip, thus allowing for increased sensitivity. In addition, this device can allow for increased detection rates; however, the signal to noise is decreased unless the device is cooled, and the collection rate is slowed [20, 29–31]. The largest current back

illuminated EMCCD camera chip is limited to 512 × 512 pixels and is much more expensive when compared to a more traditional cooled CCD on the order of a 2048 × 2048 pixel size [9, 20, 29]. Despite the limited size of the EMCCD and relative expense, the speed of these systems has led to them being incorporated into a number of commercial systems. They are being evaluated in biomedical research, and further improvements may increase their utility.

Imaging systems used for BLI are considerably less expensive than those for other imaging modalities such as magnetic resonance imaging (MRI) or positron emission tomography (PET); however, the resolution that can be obtained with using optical reporters in living animals is inherently poor (owing to light scattering) unless invasive means of removing overlying tissues or inserting a microscope objective into the body are used [32–35]. Therefore, combining low resolution macroscopic images, such as those obtained by BLI to first localize labeled cells, with high magnification microscopic studies can optimize data sets for animal studies. The macroscopic images can be used to guide the placement of the microscopic instrument, or the sampling of tissues for ex vivo histological analyses. In this way, the sequential combination of relatively low resolution, but highly sensitive BLI with high-resolution analyses, results in a powerful and comprehensive optical imaging strategy [5].

The biophotonic signals derived from luciferase reactions are relatively weak, and attenuation by tissue scattering and absorption further reduce the amount of light that is transmitted through the tissues. However, due to a near complete absence of background signal from mammalian tissues, the signal-to-noise ratio (SNR) for BLI is extraordinary relative to other imaging modalities [5, 36, 37]. Sources of background may be 10,000 fold less than the luciferase signals, making BLI one of the most sensitive noninvasive in vivo imaging modalities for studies of laboratory rodents. The sensitivity of detection will vary with reporter gene expression levels in cells, availability of cofactors and substrate, depth of the cells in the body, and composition of tissue overlying the source [20, 26, 38].

The substrates that are used for BLI are either D-luciferin ((*d*-(-)-2-(6-hydroxy-2-benzothiazolyl)thiazoline-4-carboxylic acid)) or coelenterazine, depending on which type of luciferase enzyme is being used as a reporter. The pharmacology of these two compounds can therefore affect the imaging results. Moreover, the bioluminescence reactions catalyzed by firefly luciferase (using D-luciferin as substrate) and renilla luciferase (using coelenterazine as substrate) emit light at different parts of the spectrum with maxima at 590 and 480 nm respectively. If two reporters are to be studied, these substrates are injected sequentially with coelenterazine being given first due to its much shorter circulation time. These substrates are administered exogenously to the study subjects typically via intraperitoneal (IP) injection. Injection of D-luciferin into the peritoneal cavity results in fairly stable levels, in nearly all tissues, from 5 to 15 min after injection. Intravenous (IV) injections lead to rapid distribution to most tissues and rapid clearance with a peak at about 1 min after injection of luciferin. Coelenterazine is typically injected IV and at doses that are likely sub-optimal due to the cost of the compound [39]. Coelenterazine is a high energy molecule that can be catalyzed by serum proteins, and this leads to

higher background signals than when D-luciferin is used [26]. Although not exhaustively studied, neither of the two substrates have to date been associated with any significant toxicities [40].

21.3 Recent Advances

21.3.1 Dual and Triple Reporter Genes

To maximize the amount and quality of data obtained from an animal study, dual and multi-functional reporters can be employed. A number of such reporter genes have been created and these typically encode a luciferase, a fluorescent protein, and a reporter gene for nuclear (PET) medicine imaging (e.g. herpes simplex virus-derived thymidine kinase (HSV TK)) [41–45]. Dual function reporters were first described using fusions of firefly or Renilla luciferase and green fluorescent protein (GFP), which were transferred into cell lines, and these were adoptively transplanted into animals for BLI and fluorescence microscopy at the time of necropsy [46–49]. The power of these reporters is that they can link in vivo macroscopic measurements to high resolution ex vivo assays, such as fluorescence microscopy, and flow cytometry [43, 44, 48–56], and allow cross validation via PET imaging [43, 53]. TK in the triple gene fusions encodes a protein that concentrates radiolabeled gancyclovir and its analogs in cells expressing this protein [44, 57].

Alternative options to gene fusions are coupled enzyme reporting systems that use sequential enzymatic reactions to catalyze a complex substrate. The best demonstration of this is the use of a bacterial enzyme beta-galactosidase (beta-gal), which typically catalyzes the formation of blue and fluorescent dyes for detection by microscopy and flow cytometry. In the sequential reporter-enzyme luminescence strategy these enzymes use a caged D-luciferin-galactoside conjugate (developed by Promega Corp., Madison WI). This substrate can be cleaved by beta-gal to release D-luciferin which subsequently is utilized by luciferases. This enables the indirect visualization of beta-gal activity in vivo if this reporter is co-expressed with luciferase [58].

21.3.2 3D-Image Reconstruction

Macroscopic BLI data obtained with many of the current imaging systems consists of surface-weighted, planar recordings superimposed on grayscale reference images of the subjects. The surface-weighting is due to the fact that superficial signals are more readily detected than those from deep inside the body, due to light absorption and scattering. Resolving distinct but overlapping signals that originate from two or more sites is a unique challenge for BLI since the biological sources of light cannot be modulated externally. However, solutions to this problem are forthcoming from recent advances in understanding and modeling of photon migration through mammalian tissues, and from tremendous improvements in 3D-image

reconstruction and multiview instruments that can enable bioluminescence tomography (BLT). BLT is likened to X-ray computed tomography (CT) in relation to its planar precursor, Roentgen X-rays. BLT has the potential to localize and quantify bioluminescent light sources in small living animals by generating 3D data sets from (a) multiple images taken at several different angles with either a ring of mirrors projecting to a single CCD-camera [10], or a stage for moving the animal coupled with a rotating mirror; and/or (b) spectral imaging data using optical bandpass filters [10, 59–61]. Spectrally resolved imaging of luciferase emission also enables localizing emitters of two different wavelengths in the body [62, 63]. This technology may therefore help diagnose diseases, evaluate therapies, and facilitate drug development with mouse models in the future.

While photons are subject to both scattering and absorption in tissues, in the red part of the spectrum scattering by far dominates absorption. Photon propagation is governed by the radiative transfer equation, which in this instance can be simplified to diffusion theory because of the scattering dominance [64–69]. In the red and near infra-red area of the spectrum, light is thus able to deeply penetrate through tissues and a significant number of photons escape to the outside. Most implementations of BLT involve capturing these photons in at least four separate planar bioluminescence images taken at 90° intervals [64–69]. It is then possible to develop a discrete solution for the bioluminescence data based on the optical properties of the various anatomic tissues and structures such that a finite element or boundary element inverse source problem can be solved. This solution can then be registered with CT, PET, or other methods for a correlation with anatomic structures; however, combination instruments incorporating these modalities have not yet been described.

Currently, many of the advances in BLT have been accomplished in the area of algorithm development using a number of different approaches, including finite-element reconstruction, boundary integral methods, iterative reconstruction, Born-type approximation, spectrally resolved, and temperature-modulated methods [10, 67, 68, 70–72]. To date, the vast majority of BLT studies have been limited to phantoms, however, recently Wang et al. demonstrated for the first time the real feasibility of localizing and quantifying bioluminescent activity in three-dimensional space of living mice with the use of BLT [71]. Similarly, Virostko et al. [72] used a combined surface topography/spectrally resolved approach to resolve the spatial location and biological activity in a diabetic mouse model that hosted transplanted pancreatic islets transfected with the luciferase reporter gene. Characterization of the 3D reconstruction algorithm, which was done using in vitro and in vivo comparisons of tritium luminescent beads with BLT software from the Caliper/Xenogen Living Image 3D imaging package [9], showed that the current accuracy of reconstruction is a function based on the source depth. It was found that two separate light sources could be spatially resolved as long as the distance between them was no less than the depth at which they were located. At shallow depths of 2.5 and 5.0 mm the reconstruction algorithm has a tendency to overestimate the source depth (by 68 and 22% respectively), while at the deeper depths of 7.5–15 mm the depth reconstruction is fairly accurate (less than 3% error on average). A shallow 2.5 mm depth causes

a 31% reconstruction overestimate of photon intensity. A 22% overestimation was seen at 5.0 mm, while only a 3 and 1% overestimation was seen at 7.5 and 10.0 mm in depth respectively [72]. Similar results have been seen by Comsa et al. [73], and it has been found that the limits are likely due to the source depth being shorter than the mean free path.

More recently, Kuo et al. [9] used a technique called diffuse luminescence imaging tomography (DLIT) on a mouse phantom with excitation through an optical fiber and for in vivo analysis using both tritium glass beads and PC-3 M-luc cells for excitation. Analysis was performed using bandpass filters at 560, 580, 600, 620, and 640 nm for multispectral imaging and 3-D reconstruction based on single view imaging. Given the knowledge of the exact location of the tritium glass beads implanted in a mouse, it was possible to determine less than a 6% error in source intensity and a 0.5 mm error in depth [9]. These developments therefore herald a new phase in bioluminescence imaging where three-dimensionally resolved data can provide highly sensitive and anatomically accurate information for biomedical investigators and thus save time and resources.

21.4 Fluorescence Imaging

Fluorescence has been used extensively for visualizing cell biology at many levels, from molecules to intact organisms, but is most well-suited for small and relatively transparent organisms like fruit flies and worms [74, 75]. Fluorescence imaging has mostly used small organic dyes attached to targeting agents, such as antibodies, to assess the presence of a target protein on the cell surface. Antibody-mediated staining of intracellular proteins normally requires cell fixation and permeabilization and is typically not performed in vivo [76]. Later advances allowed fluorophores to directly recognize organelles, nucleic acids**,** and certain important ions in living cells. In the past decade, fluorescent proteins have allowed for the noninvasive imaging in living cells and organisms using the genes that encode these proteins as reporter genes [76]. Research with small organic dyes, fluorescent proteins, quantum dots and even self-illuminating quantum dots [77] are advancing the field of in vivo fluorescent imaging.

In vivo fluorescent imaging differs from bioluminescent imaging in that the absorption of excitation light by a fluorophore is required to cause emission of fluorescence photons at a wavelength that is longer than the excitation wavelength. In contrast, BLI photons are generated de novo from a chemiluminescent reaction. The most common fluorescent marker of cellular function is green fluorescent protein (GFP), the discovery of which yielded the Nobel prize in chemistry to Drs. Shimomura, Shalfie, and Tsien in 2008. However, this marker, and others that are excited by, and emit, visible light of short (<600 nm) wavelengths, have limitations for in vivo studies [78]. While luciferases emit broadband light with up to 60% of the spectrum being greater than 600 nm, GFP is excited by blue light and emits green light. The absorption of visible light by hemoglobin – in this part of the spectrum the most significant absorber in the body – is greatest below 600 nm and this limits

the use of these reporters to superficial sites [76]. Longer wavelengths of excitation and emission in fluorescence (650–850 nm) allow for improved depths of imaging, and these are under development for in vivo imaging [76, 79–81].

Fluorescent proteins are stable and can accumulate over time, which compromises quantitative assessment of dynamic processes. Luciferases tend to be short-lived and thus the use of bioluminescence imaging for quantifying dynamic processes is possible [5, 20, 82]. Fluorescent proteins remain stable in ex vivo tissues, and, therefore, offer the benefit of analysis in tissue sections after biopsy or necropsy [20, 80–82]. Obtaining relevant data from fluorescent proteins can be improved by imaging labeled cells in a live mouse using a traditional fluorescent or confocal microscope [83]. High-resolution imaging in scattering tissue at depths of up to ~1 mm is improved with multi-photon excitation with ultrashort pulsed infrared light because the collected emission photons must have originated from the illumination focus, even if they have suffered significant tissue scattering [84, 85]. Moreover, the two-photon approach reduces the potential for photo bleaching and minimizes background signal from out of focus tissue since fluorescence is only generated at the focus where the intensity is sufficiently high to allow the two-photon excitation to occur. Fluorescence imaging at greater depths is possible using novel serial reconstruction techniques in fixed tissue [86] or by tomography in live tissues, albeit with limited resolution and sensitivity [60, 87].

Bioluminescence differs from fluorescence in that a simple chemical reaction provides the emitted light, rather than the absorption of excitation light. Bioluminescence provides many advantages compared to fluorescent markers, especially in vivo. Luciferase does not require optical excitation and therefore does not cause un-wanted autofluorescence. Moreover, in BLI, the light transport problem is significantly reduced since light only has to traverse the tissue from its point of origin to the surface.

Fluorescent proteins have a number of requirements for use in in vivo assays. Fluorescent proteins need to express efficiently without toxicity in the chosen system and must be bright enough to allow sufficient signal above autofluorescence to be reliably detected and imaged [80]. The fluorescent protein must have sufficient photostability to be imaged for the duration of the experiment [80]. If the fluorescent protein needs to be expressed as a fusion to another protein of interest, it should not oligomerize [80]. The fluorescent protein should be insensitive to environmental effects that could complicate quantitative interpretation of experimental results [80]. Last, in multiple-labeling experiments, the set of fluorescent proteins should have minimal crosstalk in their excitation and emission channels [80].

The brightness of a fluorescent protein is determined by the intrinsic brightness of the protein which is controlled by its maturation speed and efficiency, extinction coefficient, quantum yield, and photostability [80]. In addition to the fluorescent protein limits, the optical properties of the imaging setup with regards to illumination wavelengths and intensity, spectra of filters and dichroic mirrors, and camera sensitivity to the emission spectra are also important [80]. For photostability, it is essential to have the optimal filter sets to obtain the best performance from the fluorescent protein.

21.4.1 New Fluorescent Proteins

While fluorescent imaging has been limited for in vivo analysis, Tsien et al. have been developing new monomeric red, orange and yellow fluorescent proteins derived from the Discosoma sp. red fluorescent protein. These novel fluorescent proteins show a great deal of improved options of in vivo analysis in mice [76, 79–81]. By using proteins with excitation and emission in the 550–600 nm range, improved data from deeper tissue as well as lower levels of autofluorescence can be obtained [79–81]. Tsien et al. have found several new colors of proteins that have increased tolerance of N- and C-terminal fusions, and improvements in extinction coefficients, quantum yields and photostability [79–81]. However, no single variant is optimal for all criteria of interest. Purified proteins named after fruits (mHoneydew, mBanana, mOrange, tdTomato, mTangerine, mStrawberry, and mCherry) have all been tested and have shown improved excitation (487–587 nm) and emission (537–610 nm) wavelengths for use in vivo [79–81]. The greatest brightness (extinction coefficient, 138,000 $M^{-1}cm^{-1}$, and quantum yield, 0.69) was found in tdTomato; however, the molecular weight was also doubled due to dimerization requirements [79]. The monomer, mCherry, offers the longest wavelengths, highest photostability, fastest maturation and excellent pH resistance of all of the newest proteins [79]. The excitation and emission maxima are 3 nm longer than those of mRFP1, where it is the closest upgrade. While its quantum efficiency is slightly lower (0.22 compared to 0.25 for mRFP1), its higher extinction coefficient (due to near-complete maturation), tolerance of N-terminal fusions and photostability have made mRFP1 obsolete [79]. New, far-red, fluorescent proteins can produce cleaner spectral separation and are currently being evaluated [79, 80].

Monomeric proteins with longer excitation and emission, greater brightness and improved photostability are being sought to enable more effective in vivo imaging experiments [79–81]. Efficiently folding monomeric photoconvertible proteins will improve the ability to perform the photolabeling of fusion proteins [80]. In addition, the long-wavelength end of the fluorescent protein spectrum will continue to expand, which will allow for more sensitive and efficient imaging in thick tissue and whole animals in the future.

21.5 Bioluminescent Imaging of Heat Shock Protein 70 with Laser Thermal Stress/Injury

An improved understanding of the cellular mechanisms and activated pathways will help improve the selection of laser parameters that can enhance laser treatment by reducing undesirable side effects and can provide new pathways for therapy [88]. For most laser interaction, the optical energy is converted to thermal energy, and cells and tissues can be thermally damaged either reversibly or irreversibly. Traditionally, biophysical markers such as vacuolization, hyperchromasia and protein denaturation have been typical signs of thermal damage [89–92]. Subtle thermal

effects are not as visible. Heat shock proteins (hsp) are among the most sensitive biological indicators of thermal stress. Hsp70 proteins are found in all organisms and show high levels of similarity across species as evidenced by a high level of amino acid conservation [93]. The high conservation levels indicate that hsp70 is crucial for survival due to withstanding evolutionary change over time. The hsp family of proteins is normally constitutively expressed with 5–10% of total protein content in a given cell, but can be upregulated to 15% of the cellular content in the presence of cellular stress [94].

Most mammalian cells respond to a temperature increase of at least 5–6°C from the normal 37°C body temperature [95, 96]. Hsp70 plays an important role in thermotolerance, apoptosis, and necrosis in all cells. Hsp70 levels in cells have commonly been determined by methods including Western blotting and in situ hybridization. Both of these techniques involve time and labor intensive procedures as well as requiring tissue excision and lysis for analyses. In addition, dynamic (temporal) information cannot be obtained using this approach since only single time points can be assessed from a given tissue. By contrast, the use of an optically active reporter gene would provide a convenient dynamic readout when used with BLI [97, 98].

In these studies the reporter gene is expressed from the promoter sequence from one of the hsp genes, murine hsp70a1. This genetic construct then acts as a light with a switch where control of the transcription is linked to the cellular control of hsp70 expression. In these studies luciferase (luc) from the North American firefly was used. Therefore, by tracking the amount of luciferase by virtue of BLI, hsp70 transcription can be quantitatively assessed. In addition, the native hsp70 promoters present in the genome continue to produce hsp70 normally so that cell physiology is not perturbed. However, in cases of extreme hyperthermal stress, the factors that activate hsp gene expression may be limiting. At this level with the promoter-reporter fusion, the quantification is based on hsp70 transcription and not protein levels. Changes are measured by imaging after cellular stress. This stress can be from a number of sources including laser irradiation, and hsp70 gene activation has been established using different laser protocols. As a result, responses leading to increases in thermotolerance or proceeding cell death can be quantified with BLI [99].

21.5.1 In Vitro Imaging of hsp70-Luc-NIH-3T3 Cells Treated with a Ho:YAG Laser

Hsp70 transcription has been monitored after laser-induced injury in a stable cell line (NIH-3T3) containing the firefly luciferase (luc) reporter gene attached to the hsp promoter (murine hsp70a1) [97, 99]. Thermal injury of the cells was generated with a pulsed holmium-yttrium aluminum garnet (Ho:YAG) laser ($\lambda = 2.1\ \mu m, \tau_p = 250\ \mu s, 3$ Hz). The pulse energy was changed from 10 s, 30 pulses and 108 mJ/pulse; 20 s, 60 pulses and 68 mJ/pulse; 30 s, 90 pulses and 51 mJ/pulse; and 40 s, 120 pulses and 43 mJ/pulse. The laser energy was applied to the bottom of a 96-well black plastic plate through an optical fiber in a 37°C incubator which

allowed 27% of the total energy to reach the cells. It was reported that a minimum pulse energy (65 mJ/pulse [total energy 1.95 J; total radiant exposure = 6 J/cm^2]) was needed to activate the hsp70 response, and a higher energy (103 mJ/pulse [total energy 3.09 J; total radiant exposure = 9.6 J/cm^2]) was associated with a reduction in hsp70 response and cell death.

The bioluminescence levels correlated well with actual hsp70 protein concentration as determined by enzyme-linked immunosorbent assays (ELISA). Within a small range between a lower activation threshold and an upper threshold that leads to cell death, the hsp70 response followed an Arrhenius relationship when constant temperature laser experiments were carried out. This stable cell line with the hsp70a1-luc construct was able to provide an effective and convenient method to noninvasively quantify hsp70 transcription as a function of time after moderate-temperature heat shock. Moderate temperature-induced damage phenomena in tissue have been difficult to examine with conventional methods of detection such as light microscopy [89]. Monitoring transcription using a genetic engineering approach has provided a method sensitive enough to detect the subtle biochemical changes in cells in a noninvasive way. The method of BLI of these genetically engineered cells provided an accurate measure for the production of hsp70 in response to laser-induced heat damage. In addition, this method allowed for multiple time point measurements from the same sample which is critical when there are time variations in the response of the cells [89].

21.5.2 In Vitro Imaging of Hsp70-Luc-Organotypic Raft Cultures Treated with a Carbon Dioxide Laser

In a study, by Wilmink et al. [100], a tissue model that emulates human tissue was used to assess sub-lethal laser thermal damage noninvasively, quantitatively, and sequentially. The model employs an organotypic raft culture (in essence, tissue-engineered skin) equipped with the same heat-inducible reporter gene system (hsp70) as described above to assess the heat shock response using BLI [101–103]. The model was developed, characterized, and validated using a heated water bath and a CO_2 laser using parameters similar to those used in laser skin resurfacing (LSR) [100].

It was found that an organotypic raft culture equipped with a bioluminescent reporter gene, under control of a heat activated promoter (hsp70a1), could be used to monitor thermally modulated gene expression which provides noninvasive detection of subtle biochemical changes of biological processes demanding sequential analysis including laser-tissue damage studies. Laser irradiation was performed with a continuous wave (CW) CO_2 laser with a 4.5 mm radius selected to ensure full raft irradiation. A Gaussian beam profile caused the center of the raft culture to experience the highest irradiance. Laser irradiance with 1 min of exposure with 0.43, 0.72, 1.0, and 1.43 W of power corresponding to 0.697, 1.131, 1.584, and 2.262 W/cm^2 was used with n=4 for each raft culture [100].

BLI was performed with an IVIS 100 imaging system at 0, 2, 4, 6, 8, 12, 24, and 48 h after heating with a 3 min integration time (Fig. 21.1) [100]. These studies

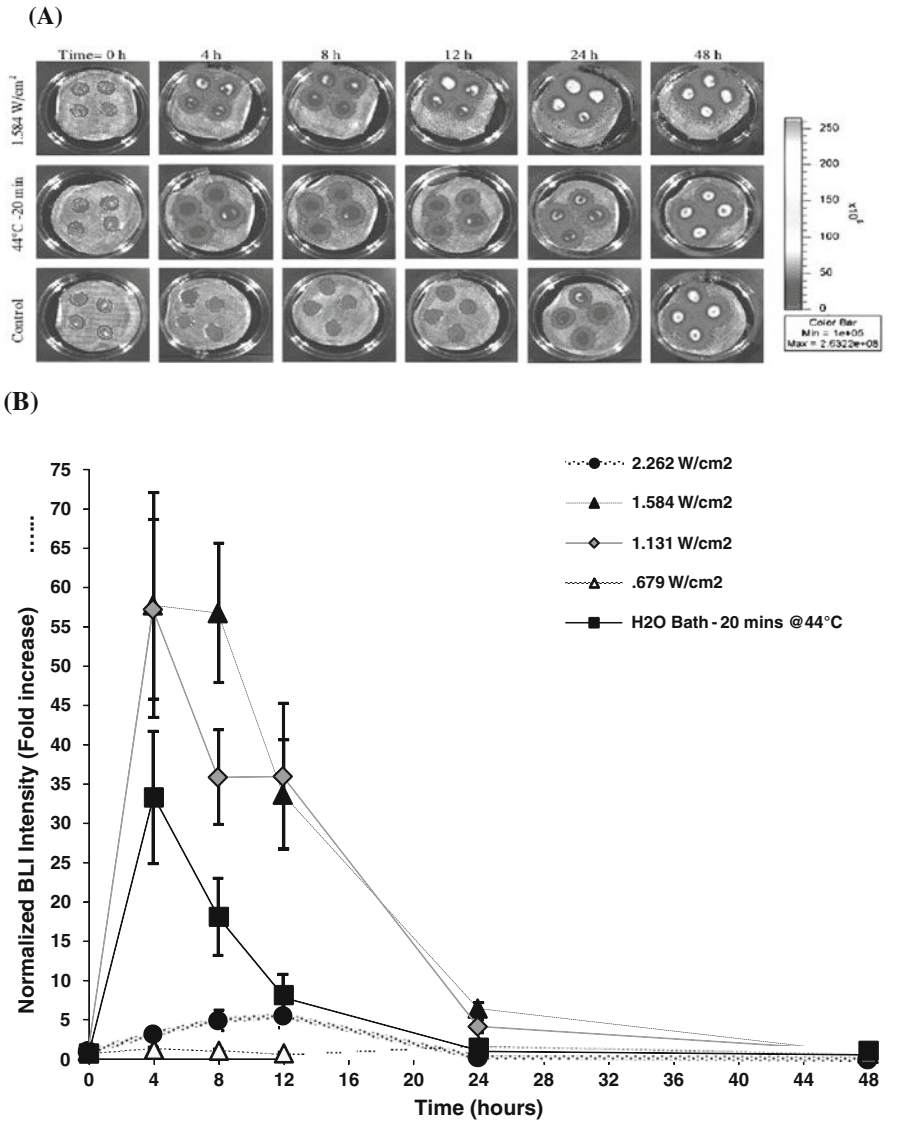


Fig. 21.1 (**a**) Visualization of hsp70 induction in organotypic raft cultures. Raft cultures were exposed to heat shock conditions via a 44°C water bath for 20 min, a CO_2 laser irradiation of 1.584 W/cm^2 for 1 min, and a control with no heat shock. Rafts were supplied with 200 μl of 0.94 mg/ml of luciferin (Biosynth AG) 1 min prior to imaging. (**b**) A plot of the intensity of the bioluminescent signal over time in adenovirally transfected raft cultures. For all laser radiated samples, the maximal expression occurred at $t = 4$–8 h after exposure. Maximal expression occurred at $t = 4$ h for the rafts heated in a water bath. The rafts exposed to 0.72 Watts (1.584 W/cm^2) had the highest bioluminescence expression. The data normalized to the bioluminescent intensity of the control rafts (not heated), and the means and standard deviations are plotted. Values are reported as mean +/– standard errors; error bars smaller than the symbol are not shown

indicated that hsp70 expression is the highest at 4–12 h following laser irradiation. CO_2 laser irradiated rafts revealed that a minimum radiation of 0.679 W/cm^2 was needed to activate the hsp70 response, and a maximum irradiation of 2.262 W/cm^2 was associated with tissue ablation and significant cell death. Performing reverse transcription polymerase chain reactions (RT-PCR) showed that hsp70 mRNA levels increased with prolonged heating exposure. Enzyme-linked immunosorbent protein assays (ELISA) confirmed that luciferase was an accurate quantitative surrogate for hsp70 intracellular protein levels. Immunohistochemical analyses confirmed that maximal hsp70 expression occurred at a depth of 150 μm. The results indicated that BLI in engineered tissue equivalents provides a powerful model that enables sequential gene expression studies. Such a model can be used as a high throughput screening platform for laser-tissue interactions [100].

21.5.3 In Vivo Imaging of Hsp70-Luc-FVB Mice Treated with a Carbon Dioxide Laser on the Skin and Organs

BLI was used to analyze the approximate levels of hsp70 transcription activation following laser injury to the skin with a CO_2 laser by O'Connell-Rodwell et al. [104]. This approach enabled macroscopic analysis of individual lesions in live mice with a temporal resolution that is relevant to Hsp70 induction. The short half life of the luciferase enzyme (~2 h [105]) results in the assay being one of transcriptional activity [97], and, as such, both the activation and cessation of transcription can be assessed [97]. Transgenic (Tg) mice, with the Hsp70 promoter and a reporter gene, have been used in analysis of thermal stress [106]. To facilitate analysis of thermal stresses in response to laser exposure, a transgenic mouse based on a modified reporter comprised of firefly luciferase and enhanced green fluorescent protein (eGFP) joined by the 2A "ribosome slippage site" from the foot and mouth disease virus (Hsp70A1-luc-2A-GFP), was used [97, 99, 107]. This dual reporter construct allows for the macroscopic analysis of Hsp70 expression in skin in vivo via luciferase expression using BLI and microscopic analysis of Hsp70 expression after necropsy via GFP imaging. This mouse model was used to assess laser thermal stress and the resultant spatiotemporal patterns of Hsp70 expression in the skin in vivo after laser irradiation.

Tissue thermal stress was generated with a CO_2 laser (PLX-100, Parallax Technology Incorporated, Waltham, Massachusetts) and evaluated in the Hsp70A1-L2G mouse. Experiments were performed with a 1 s pulse duration. Four irradiation spots were generated at separate locations on the shaved back of each mouse in a grid pattern at 1.8, 2.7, 3.5, and 4.4 J/cm^2. A 6 mm diameter flat-top spot, created by using a mask, was used to generate an affected area large enough to provide adequate resolution to determine the expression profile using an IVIS 200 imaging system. Imaging was conducted using high resolution binning (4) and a 3.9 cm field of view to obtain a 60 μm resolution across each treated location. BLI was performed 10 min post IP injection of 50 μl /10 g mouse of 30 mg/ml stock solution of D-luciferin with a 1 min integration time (Fig. 21.2a) [104].

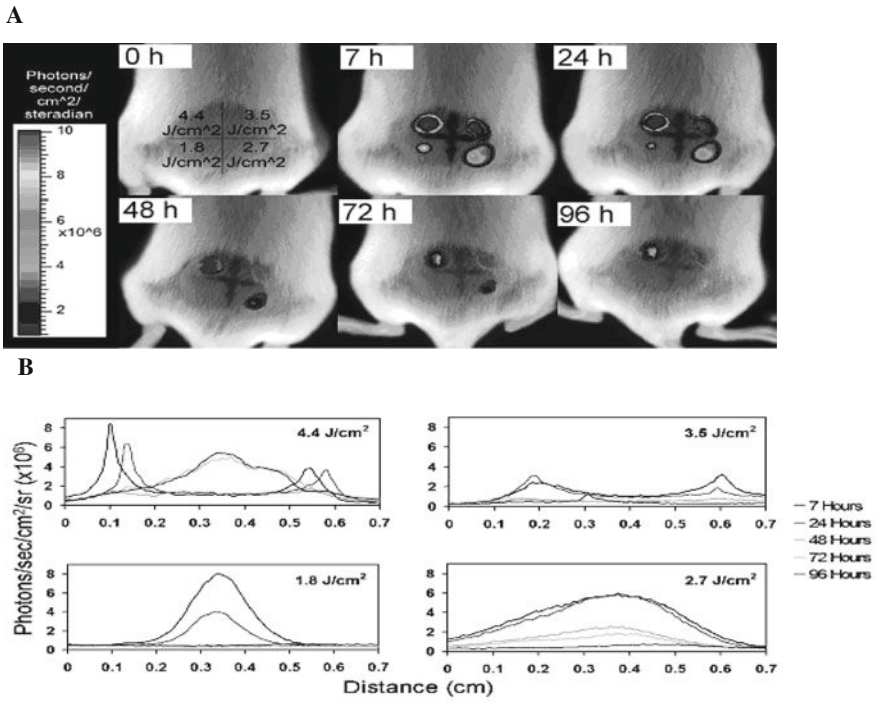


Fig. 21.2 (**a**) Bioluminescent signals indicative of the in vivo thermal response of mouse skin following laser injury. One mouse is shown with images taken at 7, 23, 48, 72, and 96 h post laser heating. (**b**) Each panel shows a series of 7-mm-long profile plots across the 6-mm laser lesion at each of the five time points and four radiant exposures for the 1-s CO_2 laser pulse. A Gaussian pattern of expression is observed for the lowest radiant exposure. At the highest radiant exposure, the center region of the pulse does not emit light at the early time points, but responds by later time points

Profile plots of each lesion generated by the laser were obtained at five imaging times for comparison by placing a 7 mm line across the center of each lesion using ImageJ (National Institutes of Health) and examining the profile of expression across this line. Analysis shows a qualitative peak of expression at 7 h post irradiation, which is apparent up to 96 h (Fig. 21.2b) [104]. Antibody staining of GFP in tissue sections from samples subjected to similar laser parameters indicated that GFP and Hsp70 expression correlate spatially throughout the laser spot.

In addition to bioluminescent imaging of hsp70 seen after laser irradiation of the skin, further research was performed on the same transgenic mouse to analyze the effect of laser irradiation on organs of the mouse in unpublished research by Mackanos et al. (Mackanos MA and Contag CH (2007) "personal communication"). The first set of experiments involved opening three mice through the peritoneum which allowed access to the heart through the diaphragm. A 1 s, 1.8 J/cm^2, 6 mm diameter CO_2 laser pulse was applied to the heart. Another set of three mice was

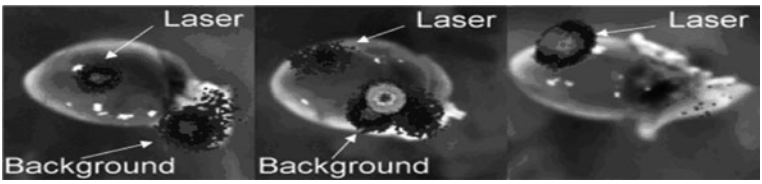


Fig. 21.3 After the peritoneum was opened, a 6 mm diameter 1 s, 1.8 J/cm^2 laser pulse was applied to the heart of three mice through the diaphragm. Removed hearts were imaged 6 h after irradiation. The laser thermal spots are seen. In addition, two of the hearts showed a large background signal near the aorta

irradiated with the same energy on one of the two kidneys as well as on the liver. The contralateral kidney was used as a negative control. The mice were sutured and kept alive for 6 h to allow for hsp70 expression.

Intact live mice were imaged for bioluminescence in the same way that was performed for the skin treatment alone. The organs were then removed and placed in an eight well plate with PBS and luciferin within minutes of BLI of the intact mouse for organ BLI imaging. Images for the 1 s pulse duration are shown in Fig. 21.3. A bright spot located at the same laser irradiation spot is seen on all three hearts suggesting increased expression of hsp70. In addition, two of the three mice showed a high level of background from the aorta. The results of the kidney and liver laser exposure are shown in Fig. 21.4. The first row shows the lack of bioluminescence from the unexposed control kidneys. The second row shows a small increase in bioluminescence from the laser treated kidneys from the same mouse. The laser treated bioluminescence is seen from the laser treated liver in the third row. The background signal from the skin was 1–2 orders of magnitude stronger than the signal from the organs, thus preventing analysis from an intact mouse. While the heart, kidney, and liver all show an inducible hsp70 signal after laser treatment, further analysis of laser treatment of organs would need a different transgenic mouse to be developed to allow for in vivo imaging (Mackanos MA and Contag CH (2007) "personal communication").

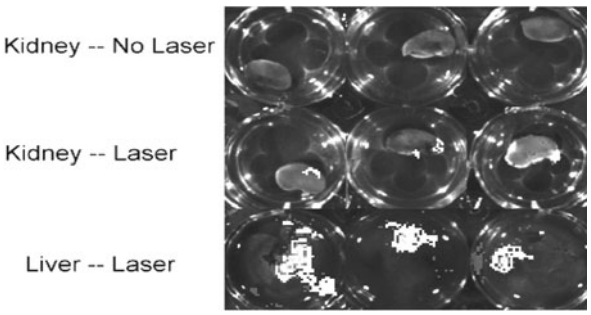


Fig. 21.4 After the peritoneum was opened, a 6 mm diameter 1 s, 1.8 J/cm^2 laser pulse was applied to one kidney and the liver of three mice. One kidney in each mouse was used as a control. Removed organs were imaged 6 h after irradiation. The laser thermal spots are seen in the laser irradiated kidneys and livers for all three mice

21.6 Photodynamic Therapy (PDT)

21.6.1 In Vivo Imaging of E. Coli-*Luc Injected Balb/c Mice Treated with a 660 nm Diode Laser*

The worldwide rise in multi-antibiotic resistance among bacteria that infect wounds and burns [108] has led to the search for alternative methods of selectively destroying bacteria without harming the host tissue [109]. Because the infection is initially localized to the wound, one method of selectively killing bacteria may be the combination of dyes and visible light, known as photodynamic therapy (PDT) [110]. PDT is a therapy method for cancer and other diseases and has received regulatory approvals for several indications in many countries [111]. It uses certain nontoxic dyes know as photosensitizers (PS) that should preferentially accumulate at the target tissue or cell-type. Subsequent illumination with visible light of the appropriate wavelength excites the PS molecule which ultimately gives rise to the formation of reactive oxygen species [110].

Genetically engineered bacteria that emit bioluminescence can be detected in vivo using a sensitive imaging camera [2] and thus the extent of infection can be determined in real time in living animals, providing both temporal and spatial information about the labeled pathogen [112, 113]. The work described here, by Hamblin et al. [110], shows a proof-of-principle of PDT for infection using a topical application of PS followed by illumination with red laser light to destroy bacteria infecting excision wounds in mice without harming the host tissue.

The bioluminescent strain of *E. coli* DH5α was constructed by transformation with the plasmid pCGLS1 which is an expression vector that contains a complete bacterial luciferase operon [114]. Four full-thickness excision wounds were made in a line along the dorsal surface using surgical scissors and forceps in Balb/c mice. A suspension of *E. coli* was inoculated into each wound. Following the topical application, a targeted polycationic photosensitizer conjugate between poly-L-lysine and chlorine6 was applied and was allowed to penetrate the Gram (–) outer bacterial membrane. Subsequent activation was performed with a 665 nm, 1 W diode laser coupled with a 200 μm optical fiber that delivered light with a uniform circular spot of 3 cm in diameter to the *E. coli* infected excision wounds in mice. Two wounds on each mouse were irradiated with 100 mW/cm^2. Mice were given a total radiant exposure of 160 J/cm^2 in four 40 J/cm^2 aliquots with imaging taking place after each aliquot of light. The total illumination time was 27 min.

The *E. coli* emit a bioluminescent signal that allows the infection to be rapidly quantified, using a low-light imaging system. BLI was performed with a low-light imaging system containing an intensified CCD camera which revealed that there was a light-dose dependent loss of luminescence in the wound treated with the conjugate and light, not seen in untreated wounds, showing that the photodynamic treatment did not damage the host tissue. It was demonstrated for the first time that rapid eradication of bacteria infecting a wound in living mice, using targeted PDT, showed the strength of an in vivo assay for infection in living animals. This study points to the possible use of this methodology in the rapid control of wounds and other localized infection [110].

21.7 Low-Level Laser Therapy (LLLT)

21.7.1 In Vivo Imaging of iNOS-Luc-FVB Mice Treated with a Four Channel Diode Laser

Low-level laser therapy (LLLT) is used clinically, among other indications, for its proposed anti-inflammatory effects, pain relief and acceleration of the regeneration of damaged tissue [115, 116]. In addition, many studies have been performed to elucidate the mechanisms of biostimulatory effects of low-level laser irradiation in contact hypersensitivity reactions [116] and induced edema [117]; however, little is know about the effects of LLLT on the acute inflammatory process at the cellular and molecular levels.

BLI methods offer the opportunity for the visualization of biological phenomena at the molecular level using reporter gene technology in small animals at a temporal resolution comparable to gene expression and protein life times in cells. Recently, Zhang et al. [118] developed a transgenic murine model, FVB/N-Tg(iNOS-luc)Xen, (iNOS-luc), which permits in vivo screening for the expression of the iNOS (inducible nitric oxide synthase) gene during the inflammatory process. In work by Moriyama et al. [119], a murine model was selected due to the small diameter of the mouse knee joint, which permits adequate detection of the bioluminescence signal. The joint size is comparable with the penetration depth of optical radiation in the wavelength window of 635–905 nm used for LLLT, allowing light penetration for different wavelengths without presenting a confounding factor for data analysis. In this study, BLI was used as a noninvasive method for monitoring iNOS gene expression after LLLT, mediated by four different wavelengths during the acute inflammatory process.

Thirty transgenic mice with the luciferase gene under control of the iNOS promoter were allocated randomly with different wavelengths (λ = 635, 785, 808, and 905 nm) and compared to a non-treated control group. In addition, twelve mice of the parental strain FVB were used for histological studies to investigate mononuclear leukocyte and neutrophil recruitment following inflammation induced by injection of xymosan A and LLLT. Laser treatment (25 mW/cm^2, 200 s, 5 J/cm^2) was applied to the knees 15 min after inflammation induction with a four-channel diode laser unit (PRO, Ontario, Canada) in a continuous wave emission mode. The knees were bent at 90° to expose the synovial space below the patella, which defined the point of light incidence.

Measurements of iNOS expression were performed at 0, 3, 5, 7, 9, and 24 h by measuring the bioluminescence signal using an IVIS Imaging System. The LLLT effects were measured by monitoring iNOS expression using photon counts per second emitted with in an ROI and compared with no-LLLT-treated animals. Experiments were performed over a three month period to ensure that observed BLI differences were not a result of the animal's age. The results showed a significant increase in BLI signal after irradiation with a 635 nm laser when compared to the non-irradiated animals and the other LLLT-treated groups, with a peak at 5 h after induction. Other treated groups were not significantly different from the control group; however, a significant difference was seen between the 635 and 785 nm

group. This indicated wavelength dependence of the LLLT effect on iNOS expression during the inflammatory process, and thus demonstrates an action spectrum of iNOS gene expression following LLLT in vivo that can be detected by BLI. Histology demonstrated the presence of fewer inflammatory cells in the synovial joints of mice irradiated with 635 nm compared with non-irradiated knee joints. This non-invasive technique permitted resolution of the changes in gene expression in the same animals monitored at different time points after LLLT treatment and showed an increase in BLI signal from 635 nm laser irradiated animals when compared with the other wavelengths and the control mice. This research found that fewer inflammatory cells were present after 635 nm-mediated LLLT when compared with the non-irradiated joints 5 h after injection of xymosan A and indicates that 635 nm-mediated LLLT results in higher expression of iNOS per cell after LLLT.

21.8 Recent Advances in Bioluminescent and Fluorescent Imaging

Developments in this field can be divided into technological advances aimed at improving the imaging technology itself (novel detectors, imaging systems) and the development of novel biological agents (red-shifted mutants of luciferases and fluorescence proteins, multi-reporter probes, engineered cells and animals, etc.). On the instrumentation side, the main focus at the present time appears to be on the development of systems capable of three-dimensional molecular imaging (i.e. bioluminescence tomography (BLT)). Work is under development by several different research groups to reveal molecular and cellular signatures with bioluminescent probes for in vivo small animals that allows for localization and quantification of a bioluminescent or fluorescent source in vivo in three dimensions [9, 11, 60, 120–129]. Currently four bioluminescent images are generally collected in ninety degree increments for analysis and are either treated with mathematical analysis alone, or are combined with computed tomography (CT) or MRI techniques for comparison. Further work on integrated multi-modality imaging systems is expected to yield a number of novel capabilities in the near future.

Current numerical BLT algorithms being used include boundary element based analysis, EM reconstruction techniques, multispectral imaging, and combined analysis with positron emitted tomography (PET) or optical-PET [121]. Future multiple angle imaging will require new imaging systems, while single view BLT reconstruction can be done with existing commercially available cooled CCD cameras. For the latter, BLT advancements will be enhanced by the addition of fluorescent tomographic principles since bioluminescence can be modeled as fluorescence in a steady-state domain [129]. Further advances with BLT algorithms will need the solution of a mathematically difficult and ill-posed inverse problem due to a lack of source-detector pairs [60]. Further work based on instrumentation as well as modeling of photon propagation and incorporation of optical heterogeneity is necessary for further advancements with bioluminescent tomography.

21.9 Conclusions

Linking reporter gene expression to genetic elements that respond to tissue stress has enabled the study of stress physiology in vivo. As additional transgenic mice are created using different promoter elements, studies of the tissue response will be further enabled and cascades of tissue signaling events can be studied. The hope is that this information will lead to the development of approaches that can prevent tissue injury and promote healing. The study of these approaches will be accelerated by the use of these transgenic mice as reporter lines that provide real-time, sensitive and high throughput data on patterns of tissue injury and the healing response.

References

1. Hastings JW. Chemistries and colors of bioluminescent reactions: A review. *Gene*, 173:5–11 (1996).
2. Contag CH, Contag PR, Mullins JI, Spilman SD, Stevenson DK, and Benaron DA. Photonic detection of bacterial pathogens in living hosts. *Mol. Microbiol*., 18:593–603 (1995).
3. Contag CH, Spilman SD, Contag PR, Oshiro M, Eames B, Dennery P, Stevenson DK, and Benaron DA. Visualizing gene expression in living mammals using a bioluminescent reporter. *Photochem. Photobiol*., 66:523–531 (1997).
4. Contag PR, Olomu IN, Stevenson DK, and Contag CH. Bioluminescent indicators in living mammals. *Nat. Med*., 4:245–247 (1998).
5. Contag CH and Bachmann MH. Advances in vivo bioluminescence imaging of gene expression. *Annu. Rev. Biomed. Eng*., 4:235–260 (2002).
6. Negrin RS and Contag CH. Innovation – In vivo imaging using bioluminescence: A tool for probing graft-versus-host disease. *Nat. Rev. Immunol*., 6:484–490 (2006).
7. Sadikot RT and Blackwell TS. Bioluminescence imaging. *Proc. Am. Thorac. Soc*., 2:511–532, 537–540 (2005).
8. Helms MW, Brandt BH, and Contag CH. Options for visualizing metastatic disease in the living body. *Infect. Inflammation: Impacts Oncogenesis*, 13:209–231 (2006).
9. Kuo C, Coquoz O, Troy TL, Xu H, and Rice BW. Three-dimensional reconstruction of in vivo bioluminescent sources based on multispectral imaging. *J. Biomed. Opt*., 12:024007 (2007).
10. Chaudhari AJ, Darvas F, Bading JR, Moats RA, Conti PS, Smith DJ, Cherry SR, and Leahy RM. Hyperspectral and multispectral bioluminescence optical tomography for small animal imaging. *Phys. Med. Biol*., 50:5421–5441 (2005).
11. Wang G, Shen H, Liu Y, Cong A, Cong WX, Wang Y, and Dubey P. Digital spectral separation methods and systems for bioluminescence imaging. *Opt. Express*, 16:1719–1732 (2008).
12. Villalobos V, Naik S, and Piwnica-Worms D. Current state of imaging protein-protein interactions in vivo with genetically encoded reporters. *Annu. Rev. Biomed. Eng*., 9:321–349 (2007).
13. Contag CH and Ross BD. It's not just about anatomy: In vivo bioluminescence imaging as an eyepiece into biology. *J. Magn. Reson. Imaging*, 16:378–387 (2002).
14. Dewet JR, Wood KV, Helinski DR, and Deluca M. Cloning of firefly luciferase Cdna and the expression of active luciferase in *Escherichia-Coli*. *Proc. Natl. Acad. Sci. U.S.A*., 82:7870–7873 (1985).
15. Ow DW, Wood KV, Deluca M, Dewet JR, Helinski DR, and Howell SH. Transient and stable expression of the firefly luciferase gene in plant-cells and transgenic plants. *Science*, 234:856–859 (1986).

16. Ow DW, Jacobs JD, and Howell SH. Functional regions of the cauliflower mosaic virus-35s Rna promoter determined by use of the firefly luciferase gene as a reporter of promoter activity. *Proc. Natl. Acad. Sci. U.S.A.*, 84:4870–4874 (1987).
17. Wilson T and Hastings JW. Bioluminescence. *Annu. Rev. Cell Develop. Biol.*, 14:197–230 (1998).
18. Wood KV and Deluca M. Photographic detection of luminescence in *Escherichia-Coli* containing the gene for firefly luciferase. *Anal. Biochem.*, 161:501–507 (1987).
19. Gould SJ and Subramani S. Firefly luciferase as a tool in molecular and cell biology. *Anal. Biochem.*, 175:5–13 (1988).
20. Rice BW, Cable MD, and Nelson MB. In vivo imaging of light-emitting probes. *J. Biomed. Opt.*, 6:432–440 (2001).
21. Nguyen VT, Morange M, and Bensaude O. Firefly luciferase luminescence assays using scintillation-counters for quantitation in transfected mammalian-cells. *Anal. Biochem.*, 171:404–408 (1988).
22. Brasier AR, Tate JE, and Habener JF. Optimized use of the firefly luciferase assay as a reporter gene in mammalian-cell lines. *Biotechniques*, 7:1116–1122 (1989).
23. Tromberg BJ, Coquoz O, Fishkin JB, Butler J, Svaasand LO, Wyss P, Tadir Y, and Haller U. Noninvasive characterization of tissue optical properties using frequency domain photon migration. *Photomed. Gynecol. Reprod.*, 116–132 (2000).
24. Tuchin PJ, Pollard H, and Bonello R. A randomized controlled trial of chiropractic spinal manipulative therapy for migraine. *J. Manipulative Physiol. Ther.*, 23:91–95 (2000).
25. Cheong WF, Prahl SA, and Welch AJ. A review of the optical-properties of biological tissues. *IEEE J. Quantum Electron.*, 26:2166–2185 (1990).
26. Zhao H, Doyle TC, Coquoz O, Kalish F, Rice BW, and Contag CH. Emission spectra of bioluminescent reporters and interaction with mammalian tissue determine the sensitivity of detection in vivo. *J. Biomed. Opt.*, 10:041210 (2005).
27. Contag PR. Whole-animal cellular and molecular imaging to accelerate drug development. *Drug Discovery Today*, 7:555–562 (2002).
28. Oshiro M and Moomaw B. Cooled vs. intensified vs. electron bombardment CCD cameras–applications and relative advantages. *Methods Cell Biol.*, 72:133–156 (2003).
29. Coates CG, Denvir DJ, McHale NG, Thornbury KD, and Hollywood MA. Optimizing low-light microscopy with back-illuminated electron multiplying charge-coupled device: Enhanced sensitivity, speed, and resolution. *J. Biomed. Opt.*, 9:1244–1252 (2004).
30. Sinclair MB, Haaland DM, Timlin JA, and Jones, HDT. Hyperspectral confocal microscope. *Appl. Opt.*, 45:6283–6291 (2006).
31. Heemskerk JWT, Westra AH, Linotte PM, Ligtvoet KM, Zbijewski W, and Beekman FJ. Front-illuminated versus back-illuminated photon-counting CCD-based gamma camera: Important consequences for spatial resolution and energy resolution. *Phys. Med. Biol.*, 52:N149–N162 (2007).
32. Egen JG, Rothfuchs AG, Feng CG, Winter N, Sher A, and Germain RN. Macrophage and T cell dynamics during the development and disintegration of mycobacterial Granulomas. *Immunity*, 28:271–284 (2008).
33. Halin C, Mora JR, Sumen C, and von Andrian UH. In vivo imaging of lymphocyte trafficking. *Annu. Rev. Cell Develop. Biol.*, 21:581–603 (2005).
34. Stoll S, Delon J, Brotz TM, and Germain RN. Dynamic imaging of T cell-dendritic cell interactions in lymph nodes. *Science*, 296:1873–1876 (2002).
35. Sumen C, Mempel TR, Mazo IB, and von Andrian UH. Intravital microscopy: Visualizing immunity in context. *Immunity*, 21:315–329 (2004).
36. Troy T, Jekic-McMullen D, Sambucetti L, and Rice B. Quantitative comparison of the sensitivity of detection of fluorescent and bioluminescent reporters in animal models. *Mol. Imaging*, 3:9–23 (2004).
37. Welsh DK and Kay SA. Bioluminescence imaging in living organisms. *Curr. Opin. Biotechnol.*, 16:73–78 (2005).

38. Lipshutz GS, Gruber CA, Cao YA, Hardy J, Contag CH, and Gaensler KML. In utero delivery of adeno-associated viral vectors: Intraperitoneal gene transfer produces long-term expression. *Mol. Ther*., 3:284–292 (2001).
39. Bhaumik S and Gambhir SS. Optical imaging of Renilla luciferase reporter gene expression in living mice. *Proc. Natl. Acad. Sci. U.S.A*., 99:377–382 (2002).
40. Schipper ML, Patel MR, and Gambhir SS. Evaluation of firefly luciferase biolurninescence mediated photodynarnic toxicity in cancer cells. *Mol. Imaging Biol*., 8:218–225 (2006).
41. Ray P, De A, Min JJ, Tsien RY, and Gambhir SS. Imaging tri-fusion multimodality reporter gene expression in living subjects. *Cancer Res*., 64:1323–1330 (2004).
42. Day RN, Nordeen SK, and Wan YH. Visualizing protein-protein interactions in the nucleus of the living cell. *Mol. Endocrinol*., 13:517–526 (1999).
43. Edinger M, Hoffmann P, Contag CH, and Negrin RS. Evaluation of effector cell fate and function by in vivo bioluminescence imaging. *Methods*, 31:172–179 (2003).
44. Ray S, Paulmurugan R, Hildebrandt I, Iyer M, Wu L, Carey M, and Gambhir SS. Novel bidirectional vector strategy for amplification of therapeutic and reporter gene expression. *Human Gene Ther*., 15:681–690 (2004).
45. Ponomarev V, Doubrovin M, Serganova I, Vider J, Shavrin A, Beresten T, Ivanova A, Ageyeva L, Tourkova V, Balatoni J, Bornmann W, Blasberg R, and Tjuvajev JG. A novel triple-modality reporter gene for whole-body fluorescent, bioluminescent, and nuclear noninvasive imaging. *Eur. J. Nucl. Med. Mol. Imaging*, 31:740–751 (2004).
46. Day RN, Kawecki M, and Berry D. Dual-function reporter protein for analysis of gene expression in living cells. *Biotechniques*, 25:848–856 (1998).
47. Liu JX, Wang YB, Szalay AA, and Escher A. Visualizing and quantifying protein secretion using a Renilla luciferase-GFP fusion protein. *Luminescence*, 15:45–49 (2000).
48. Nakajima Y, Fujiwara S, Sawai H, Imashimizu M, and Tsuzuk, M. A phycocyanin-deficient mutant of Synechocystis PCC 6714 with a single-base substitution upstream of the cpc operon. *Plant Cell Physiol*., 42:992–998 (2001).
49. Costa GL, Sandora MR, Nakajima A, Nguyen EV, Taylor-Edwards C, Slavin AJ, Contag CH, Fathman CG, and Benson JM. Adoptive immunotherapy of experimental autoimmune encephalomyelitis via T cell delivery of the IL-12 p40 subunit. *J. Immunol*., 167:2379–2387 (2001).
50. Hardy J, Edinger M, Bachmann MH, Negrin RS, Fathman CG, and Contag CH. Bioluminescence imaging of lymphocyte trafficking in vivo. *Exp. Hematol*., 29:1353–1360 (2001).
51. Edinger, M, Sweeney, TJ, Tucker AA, Olomu AB, Negrin RS, and Contag CH. Noninvasive assessment of tumor cell proliferation in animal models. *Neoplasia (New York)*, 1:303–310 (1999).
52. Edinger M, Cao YA, Hornig YS, Jenkins DE, Verneris MR, Bachmann MH, Negrin RS, and Contag CH. Advancing animal models of neoplasia through in vivo bioluminescence imaging. *Eur. J. Cancer*, 38:2128–2136 (2002).
53. Edinger M, Cao YA, Verneris MR, Bachmann MH, Contag CH, and Negrin RS. Revealing lymphoma growth and the efficacy of immune cell therapies using in vivo bioluminescence imaging. *Blood*, 101:640–648 (2003).
54. Kim JG, Choi YS, Kim KY, Bae JS, Kim, I, Sohn HD, and Jin BR. Genomic structure and phylogenetic analysis of the luciferase gene of the firefly, Luciola lateralis (Coleoptera: Lampyridae). *Eur. J. Entomol*., 101:1–11 (2004).
55. Mandl S, Schimmelpfennig C, Edinger M, Negrin RS, and Contag CH. Understanding immune cell trafficking patterns via in vivo bioluminescence imaging. *J. Cell. Biochem.*, 87(S39):239–248 (2002).
56. Mandl SJ, Mari C, Edinger M, Negrin RS, Tait JF, Contag CH, and Blankenberg FG. Multi-modality imaging identifies key times for annexin V imaging as an early predictor of therapeutic outcome. *Mol. Imaging*, 3:1–8 (2004).

57. Kim YJ, Dubey P, Ray P, Gambhir SS, and Witte ON. Multimodality imaging of lymphocytic migration using lentiviral-based transduction of a tri-fusion reporter gene. *Mol. Imaging Biol*., 6:331–340 (2004).
58. Wehrman TS, von Degenfeld G, Krutzik P, Nolan GP, and Blau HM. Luminescent imaging of beta-galactosidase activity in living subjects using sequential reporter-enzyme luminescence. *Nat. Methods*, 3:295–301 (2006).
59. Ntziachristos V, Bremer C, and Weissleder, R. Fluorescence imaging with near-infrared light: New technological advances that enable in vivo molecular imaging. *Eur. Radiol*., 13:195–208 (2003).
60. Ntziachristos V, Ripoll J, Wang LHV, and Weissleder R. Looking and listening to light: The evolution of whole-body photonic imaging. *Nat. Biotechnol*., 23:313–320 (2005).
61. Tonary AM and Pezacki JP. Simultaneous quantitative measurement of luciferase reporter activity and cell number in two- and three-dimensional cultures of hepatitis C virus replicons. *Anal. Biochem*., 350:239–248 (2006).
62. Kadurugamuwa JL, Modi K, Coquoz O, Rice B, Smith S, Contag PR, and Purchio T. Reduction of astrogliosis by early treatment of pneumococcal meningitis measured by simultaneous imaging, in vivo, of the pathogen and host response. *Infect. Immun*., 73:7836–7843 (2005).
63. Hoffman EA, Conley SM, Stamer WD, and McKay BS. Barriers to productive transfection of trabecular meshwork cells. *Mol. Vision*, 11:869–875 (2005).
64. Alexandrakis G, Rannou FR, and Chatziioannou AF. Tomographic bioluminescence imaging by use of a combined optical-PET (OPET) system: A computer simulation feasibility study. *Phys. Med. Biol*., 50:4225–4241 (2005).
65. Cong AX and Wang G. A finite-element-based reconstruction method for, 3D fluorescence tomography. *Opt. Express*, 13:9847–9857 (2005).
66. Cong WX, Wang G, Kumar D, Liu Y, Jiang M, Wang LV, Hoffman EA, McLennan G, McCray PB, Zabner J, and Cong A. Practical reconstruction method for bioluminescence tomography. *Opt. Express*, 13:6756–6771 (2005).
67. Cong W and Wang G. Boundary integral method for bioluminescence tomography. *J. Biomed. Opt*., 11:020503 (2006).
68. Cong WX, Durairaj K, Wang LV, and Wang G. A Born-type approximation method for bioluminescence tomography. *Med. Phys*., 33:679–686 (2006).
69. Slavine NV, Lewis MA, Richer E, and Antich pp. Iterative reconstruction method for light emitting sources based on the diffusion equation. *Med. Phys*., 33:61–68 (2006).
70. Dehghani H, Davis SC, Jiang SD, Pogue BW, Paulsen KD, and Patterson MS. Spectrally resolved bioluminescence optical tomography. *Opt. Lett*., 31:365–367 (2006).
71. Wang G, Cong W, Durairaj K, Qian X, Shen H, Sinn P, Hoffman EA, McLennan G, and Henry M. In vivo mouse studies with bioluminescence tomography. *Opt. Express*, 14:7801–7809 (2006).
72. Virostko J, Powers AC, and Jansen ED. Validation of luminescent source reconstruction using single-view spectrally resolved bioluminescence images. *Appl. Opt*., 46:2540–2547 (2007).
73. Comsa DC, Farrell TJ, and Patterson MS. Bioluminescence imaging of point sources implanted in small animals post mortem: Evaluation of a method for estimating source strength and depth. *Phys. Med. Biol*., 52:5415–5428 (2007).
74. Chalfie M, Tu Y, Euskirchen,G, Ward WW, and Prasher DC. Green fluorescent protein as a marker for gene-expression. *Science*, 263:802–805 (1994).
75. Plautz JD, Day RN, Dailey GM, Welsh SB, Hall JC, Halpain S, and Kay SA Green fluorescent protein and its derivatives as versatile markers for gene expression in living Drosophila melanogaster, plant and mammalian cells. *Gene*, 173:83–87 (1996).
76. Giepmans BNG, Adams SR, Ellisman MH, and Tsien RY. Review – the fluorescent toolbox for assessing protein location and function. *Science*, 312:217–224 (2006).
77. So MK, Xu CJ, Loening AM, Gambhir SS, and Rao JH. Self-illuminating quantum dot conjugates for in vivo imaging. *Nat. Biotechnol*., 24:339–343 (2006).

78. Tsien RY. The green fluorescent protein. *Annu. Rev. Biochem*., 67:509–544 (1998).
79. Shaner NC, Campbell RE, Steinbach PA, Giepmans BNG, Palmer AE, and Tsien RY. Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nat. Biotechnol*., 22:1567–1572 (2004).
80. Shaner NC, Steinbach PA, and Tsien RY. A guide to choosing fluorescent proteins. *Nat. Methods*, 2:905–909 (2005).
81. Shu XK, Shaner NC, Yarbrough CA, Tsien RY, and Remington SJ. Novel chromophores and buried charges control color in mFruits. *Biochemistry*, 45:9639–9647 (2006).
82. Contag CH, Jenkins D, Contag FR, and Negrin RS. Use of reporter genes for optical measurements of neoplastic disease in vivo. *Neoplasia*, 2:41–52 (2000).
83. Izzo AD, Mackanos MA, Beckham JT, and Jansen ED. In vivo optical imaging of expression of vascular endothelial growth factor following laser incision in skin. *Lasers Surg. Med*., 29:343–350 (2001).
84. Zipfel WR, Williams RM, and Webb WW. Nonlinear magic: Multiphoton microscopy in the biosciences. *Nat. Biotechnol*., 21:1368–1376 (2003).
85. Helmchen F and Denk W. Deep tissue two-photon microscopy. *Nat. Methods*, 2:932–940 (2005).
86. Tsai PS, Friedman B, Ifarraguerri AI, Thompson BD, Lev-Ram V, Schaffer CB, Xiong C, Tsien RY, Squier JA, and Kleinfeld D. All-optical histology using ultrashort laser pulses. *Neuron*, 39:27–41 (2003).
87. Keller PJ, Pampaloni F, and Stelzer EHK. Life sciences require the third dimension. *Curr. Opin. Cell Biol*., 18:117–124 (2006).
88. Welch AJ, Motamedi M, Rastegar S, LeCarpentier GL, and Jansen D. Laser thermal ablation. *Photochem. Photobiol*., 53:815–823 (1991).
89. Thomsen S. Pathological analysis of photothermal and photomechanical effects of laser-tissue interactions. *Photochem. Photobiol*., 53:825–835 (1991).
90. Thomsen SL. Identification of lethal thermal injury at the time of photothermal treatment. In: GJ Mueller and A Roggan (eds) *Laser-Induced Interstitial Thermotherapy*, Vol. PM25P. SPIE Press, Bellingham, WA, 459–467 (1995).
91. Thomsen SL. Mapping of thermal injury in biologic tissues using quantitative pathologic techniques. In: TP Ryan and TZ Wong (eds) *Thermal Treatment of Tissues with Image Guidance*, Vol. 3594. SPIE, Bellingham, WA, 82–95 (1999).
92. Thomsen SL. Qualitative and quantitative pathology of clinically relevant thermal lesions. In: TP Ryan (ed) *Matching the Energy Source to the Clinical Need*, Vol. CR75. SPIE, Bellingham, WA, 425–459 (1999).
93. Kiang JG and Tsokos GC. Heat shock protein, 70 kDa: Molecular biology, biochemistry, and physiology. *Pharmacol. Ther*., 80, 183–201 (1998).
94. Pockley AG. Heat shock proteins, inflammation, and cardiovascular disease. *Circulation*, 105:1012–1017 (2002).
95. Mizuno S, Ishii A, Murakami Y, and Akagawa H. Stress dose-dependent suppression of heat shock protein gene expression by inhibiting protein synthesis during heat shock treatment. *Cell Struct. Funct.,* 22:7–13 (1997).
96. Huang L, Mivechi NF, and Moskophidis D. Insights into regulation and function of the major stress-induced hsp70 molecular chaperone in vivo: Analysis of mice with targeted gene disruption of the hsp70.1 or hsp70.3 gene. *Mol. Cell. Biol*., 21:8575–8591 (2001).
97. O'Connell-Rodwell CE, Shriver D, Simanovskii DM, McClur, C, Cao YA, Zhang WS, Bachmann MH, Beckham JT, Jansen ED, Palanker D, Schwettman HA, and Contag CH. A genetic reporter of thermal stress defines physiologic zones over a defined temperature range. *Faseb J*., 18:264–271 (2004).
98. Beckham JT, Baran JA, Mackanos MA, Crooke C, Takahashi T, O'Connell-Rodwell C, Contag CH, and Jansen ED. Bioluminescence imaging as a marker for cellular Hsp70 response to thermal laser injury. In: SL Jacques (ed) *SPIE Laser-Tissue Interaction XIV*, Vol. 4961. SPIE, San Jose, CA, pp. 153–164 (2003).

99. Beckham JT, Mackanos MA, Crooke C, Takahashi T, O'Connell-Rodwell C, Contag CH, and Jansen ED. Assessment of cellular response to thermal laser injury through bioluminescence imaging of heat shock protein, 70. *Photochem. Photobiol*., 79:76–85 (2004).
100. Wilmink GJ, Opalenik SR, Beckham JT, Davidson JM, and Jansen ED. Assessing laser-tissue damage with bioluminescent imaging. *J. Biomed. Opt*., 11(4):041114 (2006).
101. Kao B, Kelly KM, Majaron B, and Nelson JS. Novel model for evaluation of epidermal preservation and dermal collagen remodeling following photorejuvenation of human skin. *Lasers Surg. Med*., 32:115–119 (2003).
102. Kao BS, Kelly KM, Aguilar G, Hosaka Y, Barr RJ, and Nelson JS. Evaluation of cryogen spray cooling exposure on in vitro model human skin. *Lasers Surg. Med*., 34:146–154 (2004).
103. Viehoever AR, Anderson D, Jansen ED, and Mahadevan-Jansen A. Organotypic raft cultures as an effective in vitro tool for understanding Raman spectral analysis of tissue. *Photochem. Photobiol*., 78:517–524 (2003).
104. O'Connell-Rodwell CE, Mackanos MA, Simanovskii D, Cao Y-A, Bachmann MH, Schwettman HA, and Contag CH. In vivo analysis of heat-shock-protein-70 induction following pulsed laser irradiation in a transgenic reporter mouse. *J. Biomed. Opt*., 13:030501 (2008).
105. Leclerc GM., Boockfor FR, Faught WJ, and Frawley LS. Development of a destabilized firefly luciferase enzyme for measurement of gene expression. *Biotechniques*, 29:590–601 (2000).
106. Christians ES and Benjamin IJ. The stress or heat shock (HS) response: Insights from transgenic mouse models. *Methods*, 35:170–175 (2005).
107. Donnelly MLL, Hughes LE, Luke G, Mendoza H, ten Dam E, Gani D, and Ryan MD. The 'cleavage' activities of foot-and-mouth disease virus, 2A site-directed mutants and naturally occurring '2A-like' sequences. *J. Gen. Virol*., 82:1027–1041 (2001).
108. Albrich WC, Angstwurm M, Bader L, and Gartner R. Drug resistance in intensive care units. *Infection*, 27:S19–S23 (1999).
109. Weinstein RA. Nosocomial infection update. *Emerging Infect. Dis*., 4:416–420 (1998).
110. Hamblin MR, O'Donnell DA, Murthy N, Contag CH, and Hasan T. Rapid control of wound infections by targeted photodynamic therapy monitored by in vivo bioluminescence imaging. *Photochem. Photobiol*., 75:51–57 (2002).
111. Dougherty TJ, Gomer CJ, Henderson BW, Jori G, Kessel D, Korbelik M, Moan J, and Peng Q. Photodynamic therapy. *J. Natl. Cancer Inst*., 90:889–905 (1998).
112. Rocchetta HL, Boylan CJ, Foley JW, Iversen PW, Letourneau DL, McMillian CL, Contag PR, Jenkins DE, and Parr TR. Validation of a noninvasive, real-time imaging technology using bioluminescent *Escherichia coli* in the neutropenic mouse thigh model of infection. *Antimicrobial Agents Chemother*., 45:129–137 (2001).
113. Francis KP, Joh D, Bellinger-Kawahara C, Hawkinson MJ, Purchio TF, and Contag PR. Monitoring bioluminescent Staphylococcus aureus infections in living mice using a novel luxABCDE construct. *Infect. Immun*., 68:3594–3600 (2000).
114. Siragusa GR, Nawotka K, Spilman SD, Contag PR, and Contag CH. Real-time monitoring of *Escherichia coli* O157: H7 adherence to beef carcass surface tissues with a bioluminescent reporter. *Appl. Environ. Microbiol*., 65:1738–1745 (1999).
115. Vladimirov YA, Osipov AN, and Klebanov GI. Photobiological principles of therapeutic applications of laser radiation. *Biochem.-Moscow*, 69:81–90 (2004).
116. Kandolf-Sekulovic L, Kataranovski M, and Pavlovic MD. Immunomodulatory effects of low-intensity near-infrared laser irradiation on contact hypersensitivity reaction. *Photodermatol. Photoimmunol. Photomed*., 19:203–212 (2003).
117. Albertini R., Aimbire FSC, Correa FI, Ribeiro W, Cogo JC, Antunes E, Teixeira SA, De Nucci G, Castro-Faria-Neto HC, Zangaro RA, and Lopes-Martins RAB. Effects of different protocol doses of low power gallium-aluminum-arsenate (Ga-Al-As) laser radiation (650 nm) on carrageenan induced rat paw ooedema. *J. Photochem. Photobiol. B-Biol*., 74:101–107 (2004).

118. Zhang N, Weber A, Li B, Lyons R, Contag PR, Purchio AF, and West DB. An inducible nitric oxide synthase-luciferase reporter system for in vivo testing of anti-inflammatory compounds in transgenic mice. *J. Immunol*., 170:6307–6319 (2003).
119. Moriyama Y, Moriyama EH, Blackmore K, Akens MK, and Lilge L. In vivo study of the inflammatory modulating effects of low-level laser therapy on iNOS expression using bioluminescence imaging. *Photochem. Photobiol*., 81:1351–1355 (2005).
120. Wang G, Shen HC, and Cong W. Temperature-modulated bioluminescence tomography. *Opt. Express*, 14:7852–7871 (2006).
121. Wang G, Cong WX, Durairaj K, Qian X, Shen H, Sinn P, Hoffman E, McLennan G, and Henry M. In vivo mouse studies with bioluminescence tomography. *Opt. Express*, 14:7801–7809 (2006).
122. Wang G, Cong WX. Shen HO, Qian X, Henry M, and Wang Y. Overview of bioluminescence tomography-a new molecular imaging modality. *Front. Biosci*., 13:1281–1293 (2008).
123. Han WM, Kazmi K, Cong WX, and Wang G. Bioluminescence tomography with optimized optical parameters. *Inverse Probl*., 23:1215–1228 (2007).
124. Lv YJ, Tian J, Cong WX, Wang G, Yang W, Qin CG, and Xu M. Spectrally resolved bioluminescence tomography with adaptive finite element analysis: Methodology and simulation. *Phys. Med. Biol*., 52:4497–4512 (2007).
125. Paroo Z, Bollinger RA, Braasch DA, Richer E, Corey DR, Antich PP, and Mason RP. Validating bioluminescence imaging as a high-throughput, quantitative modality for assessing tumor burden. *Mol. Imaging*, 3(2):117–124 (2004).
126. Soloviev VY. Tomographic bioluminescence imaging with varying boundary conditions. *Appl. Opt*., 46:2778–2784 (2007).
127. Unlu MB and Gulsen G. Effects of the time dependence of a bioluminescent source on the tomographic reconstruction. *Appl. Opt*., 47:799–806 (2008).
128. Zavattini G, Vecchi S, Mitchell G, Weisser U, Leahy RM, Pichler BJ, Smith DJ, and Cherry SR. A hyperspectral fluorescence system for, 3D in vivo optical imaging. *Phys. Med. Biol*., 51:2029–2043 (2006).
129. Hielscher AH. Optical tomographic imaging of small animals. *Curr. Opin Biotechnol*., 16:79–88 (2005).

Chapter 22
The Optics of Bruising

Lise Lyngsnes Randeberg, Bjørn Skallerud, Neil E.I. Langlois, Olav Anton Haugen, and Lars Othar Svaasand

22.1 Introduction

Forensic medicine is a field of medicine where technology plays an increasingly important role in securing and evaluating evidence in, for example, child abuse cases and cases of domestic violence. Methods from chemistry and biological sciences have found a wide application within forensic medicine. Optical technologies like microscopy are also widely used. Despite this, in vivo or post mortem optical diagnostics by spectroscopy have traditionally not had an important role in clinical or forensic examinations. Forensic medical optics as a field might include all kinds of optical analysis for use within forensic science. This includes everything from microscopic techniques to methods for examination of evidence from a crime scene. This chapter will, however, focus on the use of optical diagnostics for examining skin, with a focus on identification, characterization and age determination of minor traumatic injuries like skin bruises.

Minor traumatic injuries like bruises have received less attention than larger and often lethal trauma like bullet wounds or injuries resulting from car crashes. The most common methods to evaluate bruises have been direct visual inspection or visual inspection of photos [1–3]. This method is highly subjective and has been found to be inaccurate [2–5]. Visual methods will depend on the experience of the examiner and also the age of the person, since the ability to see the yellow color of old bruises declines with age [6]. An error rate of up to 50% has been found in controlled experiments where experienced forensic pathologists were asked to sort bruises in three categories: recent (0–3 days), intermediate (3–7 days), and old (more than 7 days) [3, 7]. Maguire et al. [5] concluded in their review of age determination of bruises in children that visual age determination of such injuries had no

L.L. Randeberg (✉)
Department of Electronics and Telecommunications, Norwegian University of Science and Technology, Trondheim, Norway
e-mail: Lise.Randeberg@iet.ntnu.no

A.J. Welch, M.J.C. van Gemert (eds.), *Optical-Thermal Response of Laser-Irradiated Tissue*, 2nd ed., DOI 10.1007/978-90-481-8831-4_22,

scientific basis and should be avoided. There is, therefore, a need for more objective techniques for evaluation of such injuries. Optical spectroscopy combined with mathematical modeling of the bruising process and the photon transport in tissue has increased the current knowledge about the physics of bruising and has provided the fundamentals for developing more accurate optical diagnostic methods for use in forensic science [8].

22.2 Physiology

Skin is composed of epidermis and dermis overlying a layer of fat that can be termed the hypodermis. The epidermis can be subdivided into layers. At its base it has the proliferating squamous epithelial cells that are shed into the layers of the epidermis where they die, and then pass upwards, towards the surface. Melanocytes, the cells that produce pigment are also found in the basal layer. The surface layer of the epidermis, the stratum corneum, comprises anucleate, dead flattened cells containing the protein keratin. In the dermis there are blood vessels and adnexal structures, such as hair follicles and sweat glands. The thickness of the various layers will depend on the location on the body, and also on age, gender and life style [9]. Smoking and sun exposure will influence the aging of the skin [9] and thereby also the way people bruise. Bruises are caused by an impact that doesn't tear or split the surface of the skin, but is strong enough to tear blood vessels within or underneath the skin [10]. Bruising can occur anywhere in the skin as long as it is a vascular region. Bruises frequently occur in the junction between the dermis and the subcutaneous fat, since the connective tissue in this region is loose and gives less protection against tearing [11, 12]. Bruises might also result from medical procedures or examinations, especially in elderly people who are treated with anticoagulants. Several medical conditions are associated with increased bleeding [11]. These conditions include both hereditary diseases like hemophilia and acquired diseases, such as cirrhosis of the liver and autoimmune diseases. The use of anticoagulant drugs will influence strongly the tendency for bruising. Deep hematomas in the muscular tissue might first develop as palpable lumps within the tissue. Such injuries will not necessarily create visible bruises, since only blood above depths of approximately 600 microns into the skin will contribute to the perceivable skin color [13].

22.2.1 Chromophores, Biochemistry and Immunological Reaction

Oxy- and deoxyhemoglobin are two of the most important chromophores in skin, and exhibit a characteristic and well known absorption spectrum (see Fig. 22.1). When a hemorrhage occurs within the skin, the hemoglobin molecules escaping the damaged vessels are considered as alien to the body, and the immune system initiates an immediate response to the hemorrhage. The biochemistry and tissue responses to mechanical irritation are nicely described by e.g. Bauer et al.

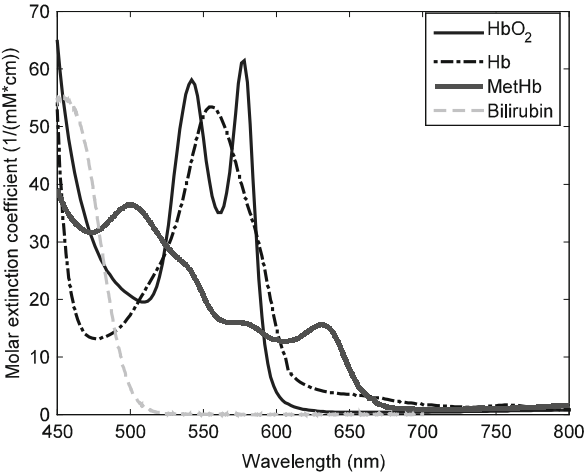


Fig. 22.1 The molar extinction coefficient of hemoglobin, bilirubin and methemoglobin. Data from Ziljstra [27] and Du [28]

[14, 15]. Mechanical irritation causes pain receptor stimulation and histamine release in the injured region. Nervous signals induced by pain receptors cause release of neuropeptides in a larger area surrounding the stimulated site. Histamine and neuropeptides cause vasodilation by relaxation of smooth muscle in arterioles with the effect occurring after 20–30 s. Szczesny et al. [16] showed that mechanical trauma caused extravasation, leakage of macromolecules, increased passage of leukocytes and increased lymph flow. Neutrophils and macrophages are recruited to the damaged area, and start to remove the hemoglobin within hours after the injury.

Macrophages are equipped with an enzyme system to break down hemoglobin [17, 18]. This enzyme, heme oxidase, breaks the porphyrin ring in the heme group and converts the heme to a green chromophore, biliverdin (see Fig. 22.2). Biliverdin is then rapidly converted to yellow bilirubin, causing the yellow color of old bruises [18–20]. This last transition from biliverdin to bilirubin is so rapid that true green color of a bruise is rarely seen. Most often the perception of green color in a bruise is caused by the optical properties of the surrounding tissue, and not by the presence of a green chromophore. Brown color in bruises might be caused by hemosiderin [11, 21, 22] or methemoglobin [1, 23]. Hemosiderin is a brown pigment that is formed by macrophages when they phagocytize erythrocytes. In animals hemosiderin has been found in macrophages as early as 24–48 h after injury [11]. However, in bruises of human skin, the literature indicates hemosiderin is not seen before 72 h after injury [24, 25]. See also Langlois [26] for a review and discussion of the literature on this topic.

Elevated methemoglobin levels may be found in individuals with a genetic deficiency of methemoglobin reductase. However, in these individuals elevated levels of methemoglobin will also be expected in normal skin, and not only in the bruises.

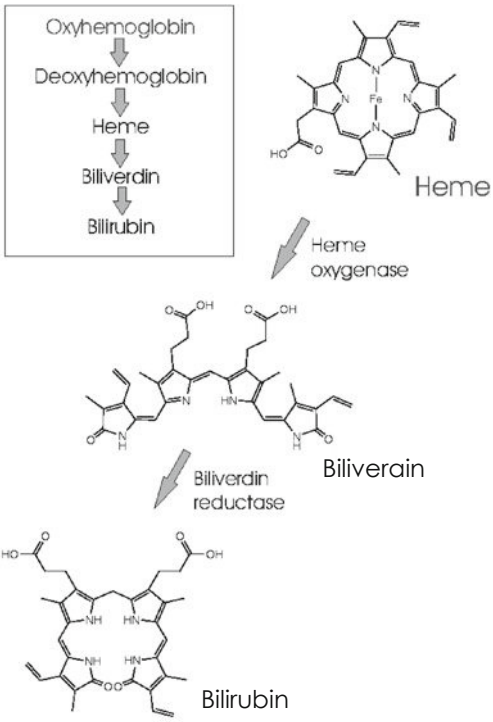


Fig. 22.2 Heme catabolism. Figure from [29]

22.2.2 Temporal Development

22.2.2.1 Immediate Skin Reactions and Resulting Injury

The macroscopic skin changes during the first phase after injury is not well described in the literature. This is probably due to the fact that practically no one sees their clinician within 30 min after injury. Randeberg et al. [30] have published data on the immediate skin reactions following minor trauma and the temporal development of bruises in a porcine model. In this study the immediate development of bruises was studied as a function of the weight and speed of the impacting object. They concluded that superficial bruises occur as a result of high speed, light weight objects, whereas deeper injuries required a heavier object and a longer contact time between the object and the surface. Paintballs create donut shaped bruises that are fully established within 15–20 min after injury. Low speed, blunt objects did not cause persistent skin changes, but gave deep muscular bleeding in most cases (see Fig. 22.3). The first reaction to a high speed trauma was a wheal and flare reaction that developed within 20–30 s after injury. A white zone could be seen in the central part of the injury surrounded by erythema, the white zone gradually changed color with time and the bruises were fully developed within 15–20 min

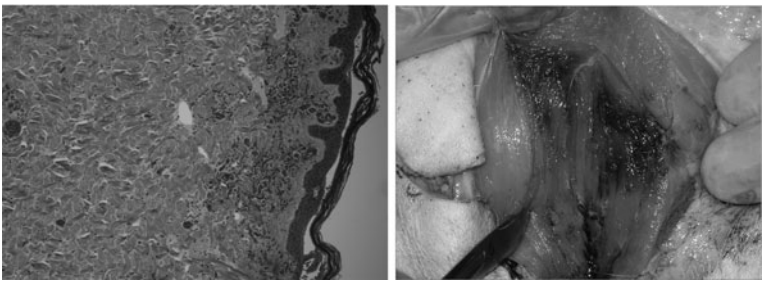


Fig. 22.3 Skin injuries in a porcine model. *Left panel*: Histology of a biopsy collected at the red edge of a paintball injury. The histology shows dilated capillaries with red blood cells and interstitial bleeding at a depth of about 200–300 micron. *Right panel*: Deep muscular hemorrhage resulting from a blunt trauma with a 2.5 kg pendulum device with a maximum speed of 4.1 m/s at impact. The hemorrhage occurred at a depth of approximately 3 cm into the tissue. Photo from [30]

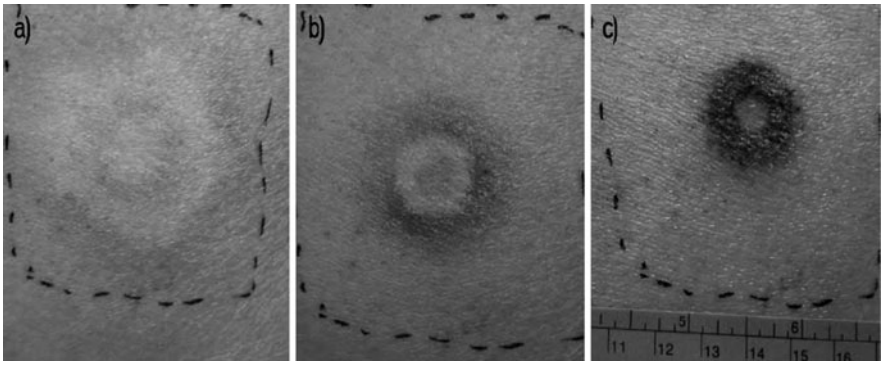


Fig. 22.4 Temporal development of a paintball injury. The paintball had a diameter of 1.7 cm, and weighed 3.4 g. The paintball hit the skin with a speed of 100 m/s. The photos were taken after: (**a**) 30–60 s; (**b**) 5 min; (**c**) 2 h and 5 min. (Reprinted with permission from [30].)

(see Fig. 22.4). The spectroscopic results in this work showed hemoglobin deoxygenation, and an increase in dermal blood volume fraction which reached a maximum after 10–15 min. No deep muscular hemorrhages were observed in those cases. In real injuries the injury mechanisms can be expected to be any mixture of superficial and deep hemorrhages, creating mixed injuries.

22.2.2.2 Temporal Development of Bilirubin Concentration

The development of bilirubin in bruises was first described by Virchow in 1847 [31]; since then, the appearance of bilirubin has been described frequently in the literature. The description of the development of yellow color in bruises [1, 3, 4, 11, 32] can be attributed to the local accumulation of bilirubin, which has been demonstrated by authors applying spectroscopic examination of the bruised skin to describe the color [12, 19, 23, 33, 34]. The bilirubin concentration at a given time

after injury will depend on the size of the hemorrhages, and the individual immune response. The bilirubin production will start when the macrophages have arrived at the injury site, and their enzymatic breakdown system has been initialized. This might occur as early as a few hours after the incident, and in most cases within 24–48 h after injury in animals [11, 30], and 72 h after injury in humans [24, 25]. The amount of bilirubin formed in a bruise is limited by the amount of blood leaking out of damaged vessels. Visible yellow color has been perceived as early as 18 h after injury [1]. The bilirubin formation is believed to peak at different times in different individuals. It is expected that the peak occurs earlier in children and later in elderly people as compared to young adults. The pattern of bilirubin formation is probably a function of the immune system in each person. The temporal development of bilirubin formation has been investigated in a spectroscopic study including 18 bruises in 10 volunteers, and in a hyperspectral study including 8 bruises in 8 persons. All volunteers were healthy, fit students in the age range 21–26 years. All bruises were acquired during sports activities. In most injuries the bilirubin formation was found to peak at four or five days after injury. For larger injuries the measured bilirubin content in skin in some cases showed a double peak with a first maximum at 2–4 days and a later maximum at approximately day 6. This first maximum is consistent with the findings by Yajima et al. [34], who observed a peak in yellow color measured by the change in CIELab *b value at day 4 after injury. The second peak in the results presented here can be explained by several causes: a low measurement frequency combined with a large variance of the data hiding the real peak, a repeated injury at the same location (not uncommon in martial arts), or the observation might also be explained by optical shielding caused by hemoglobin. This effect is further described in Section 22.3.1.1. For multiple bruises in the same individual, the absolute level of the bilirubin peak will depend on the size of the hemorrhage and the site of injury. However, preliminary results indicate that the relative, normalized bilirubin peak in one individual will appear at the same time after injury regardless of the site of injury and the size of the hemorrhage. Typical temporal development of the bilirubin index in bruised areas is shown in Fig. 22.5.

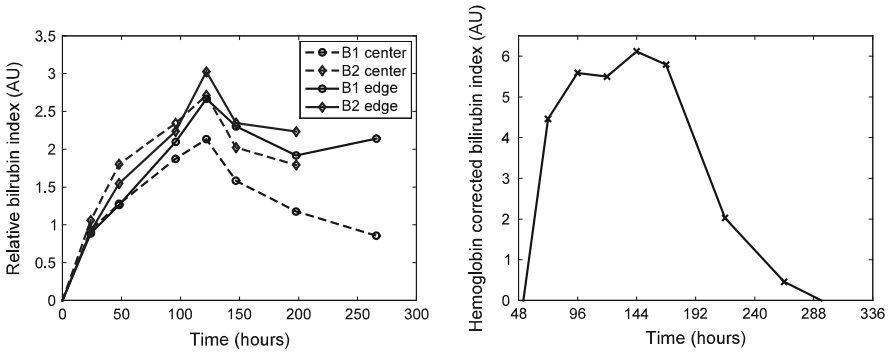


Fig. 22.5 Bilirubin index as a function of time after injury. *Left panel*: Two bruises B1 and B2, at different locations on the body. Measurements performed at the center and at the edge of the bruises. *Right panel*: Bilirubin index exhibiting a double peak, one at 96 h and one at 144 h after injury

22.2.2.3 The Central White Spot

A phenomenon that has received little attention is the central area of pallor that develops in many bruises. This whitish spot emerges one or two days after injury in many bruises, but not in all [8, 29]. The white spot has been observed to form in regions that experienced a substantial localized trauma resulting in damage to larger vessels in the skin. The spot might emerge after a venopuncture to collect blood samples, after a fall onto a pointed rock, or after a blow from a blunt object. The white zone will have approximately the shape and size of the object that created the injury. The spot will gradually increase in size with time until it fills the central area of the bruise. This effect is hypothesized to be connected to the immediate immune reaction to the trauma. If larger vessels are punctured, the amount of neutrophils and macrophages will increase rapidly in the zone surrounding the damaged vessels. This locally increased immune reaction will cause a faster breakdown of hemoglobin in this area, which leads to the whitening central zone. This effect can be simulated using the model presented in Section 22.3.3. Figure 22.6 shows the temporal development of a bruise with a white central zone. The bruise was a result of a fall from a height of approximately 1 m onto a square metal object with a size of approximately 3 cm times 10 cm. The bruise was located at the upper thigh, and the subject was a male at the age of 35. In elderly people the development of the whitish spot will be delayed as compared to young adults.

22.2.2.4 Other Chromophores

Beta-carotene, water and lipids are chromophores that also can be found in skin, and that may influence measurements on bruised skin. Just after injury the water content in the skin will often be increased due to formation of an edema in connection with the trauma. The changes in water absorption may affect the overall scattering

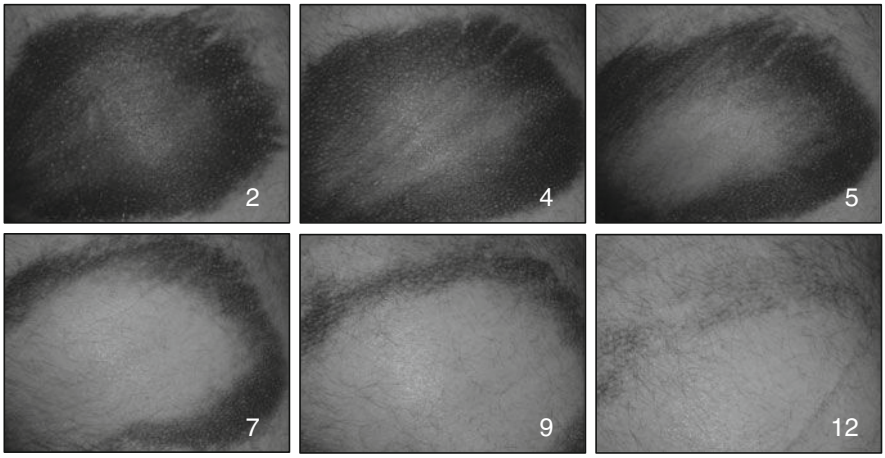


Fig. 22.6 Temporal development of a bruise. Photos taken 2, 4, 5, 7, 9, and 12 days after injury

in the tissue, and thereby the measured skin reflectance and the perceived skin color. The effect of edema on perceivable skin color is illustrated in Fig. 22.4a. The water content of the skin can to some degree be monitored by monitoring the absorption peak at 980 nm, or by monitoring the spectrum at even longer wavelengths. Lipids are found both in the stratum corneum of the skin and in the subcutaneous fat layer. The lipid absorption can in some cases be seen as a local minimum of the skin reflectance at 930 nm. This feature is especially seen in fair skinned individuals. Some lipids absorb also at 480 nm, mostly due to accumulation of beta-carotene, which is lipophilic. Beta-carotene may cause problems due to the similarity with bilirubin. High levels of beta-carotene in the skin may be confused with bilirubin, misleading the observer to estimate the bruise to be older than the actual age. However, beta-carotene will appear also in the normal skin and can therefore be corrected for. Bilirubin absorbs at 460 nm when dissolved in chloroform, and beta-carotene exhibits a double peak in absorption at approximately 440 and 480 nm [28] (see also Fig. 22.7). Beta-carotene accumulates in skin in individuals who ingest a diet rich in red and orange vegetables. The effect of beta-carotene in normal skin can be seen in Fig. 22.7, right panel.

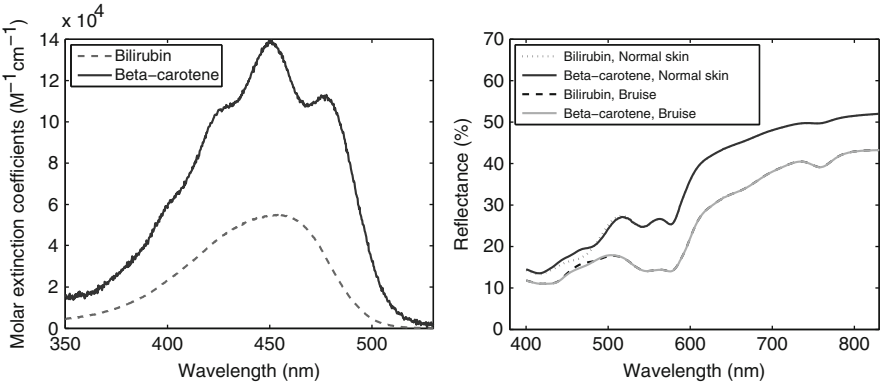


Fig. 22.7 Bilirubin and betacarotene absorption and influence on reflectance spectra. *Left panel*: Molar extinction coefficients, data from [28]. *Right panel*: The effect of beta-carotene and bilirubin on skin reflectance spectra. Simulated skin reflectance spectra of normal skin and a 5 day old medium sized bruise. Bilirubin or beta-carotene was added to the spectra manually to show the similar effect the absorbers exhibit on the reflectance spectra. Simulated using the models presented in Sections 22.3.3 and 22.3.4.1

22.3 The Physics of Minor Trauma

22.3.1 The Optics of Traumatic Injuries

Spectroscopic examination of bruises utilizes the same principle as visual inspection, and as a stand-alone technique to identify spectral changes; spectroscopy would probably only do slightly better than visual inspection. Several authors have tried to use spectroscopy to characterize and determine the age of skin bruises

[12, 19, 23, 33, 34]. The key to success in optical diagnostics of bruises is not only detection and evaluation of chromophore concentrations, but the understanding of tissue optics and how the optical properties of the tissue affect the measurements. This section will review how the perception and optical properties affect the diagnostics, and describe some of the pitfalls that should be avoided in optical diagnostics of bruises.

22.3.1.1 Observation of Yellow Color, Perception and Shielding

The first scientific study of yellow color in bruises was written by Virchow in 1847 [31]. He noted the yellow color of old bruises and believed it to be caused by a compound he called haematoidin, which later was proven to be bilirubin. The finding of yellow color in bruises has since then been associated with old bruises. The point in time when the yellow color first can be observed is frequently discussed in the literature, and everything between 18 and 40 h have been indicated as the limit for observable yellow color, see e.g. [1, 11, 32, 35]. The observed variations in time may be due to several reasons, and one should be careful when using bilirubin as a single marker to age bruises. In humans the ability to perceive yellow color declines with age and might influence the observer's accuracy [6]. The brain will compensate for the change in transmission of the lens in such a way that it is difficult for the person to detect the changes. However, this correction by the brain becomes visible in the case where the eye lens is replaced by an acrylic lens due to cataract. In this case the world will turn purple just after the surgery, and remain so for weeks. This phenomenon of chromatic adaptation, illustrates how difficult it is for a person to detect slow changes in the visual system, although this affects the ability to perceive certain colors.

All the chromophores present in skin will affect the total absorption spectrum and the perceived color. In larger bruises, the hemoglobin pool in the tissue might last for more than a week. During this period it will be difficult to visually separate yellow from dark blue. The bilirubin absorption will be masked by high hemoglobin absorption in skin. Such a shielding effect has been described for blood vessels in port-wine stains by Svaasand et al. [36]. In bruises, bilirubin will be more easily detected using a spectrometer than by visual inspection, although shielding of the chromophore due to hemoglobin absorption will make it difficult to determine the true chromophore concentration even when using spectroscopy. The simulations shown in Fig. 22.8 illustrate this effect. A bruise was simulated at five days after injury using the model presented in Section 22.3.3 using two different sizes of the hemorrhage. The bruise in the left panel was simulated with twice as large a hemorrhage as the bruise in the right panel. All other parameters were kept equal. The bilirubin content was not simulated by the equations given in Section 22.3.3.2 but added manually by adding a given bilirubin absorption to the absorption coefficient in the model. The same amount of bilirubin was added to both simulations. These graphs clearly show the shielding effect of the hemoglobin. The use of an inverse photon transport model is required to be able to reconstruct the true absorption coefficients in the bruised area.

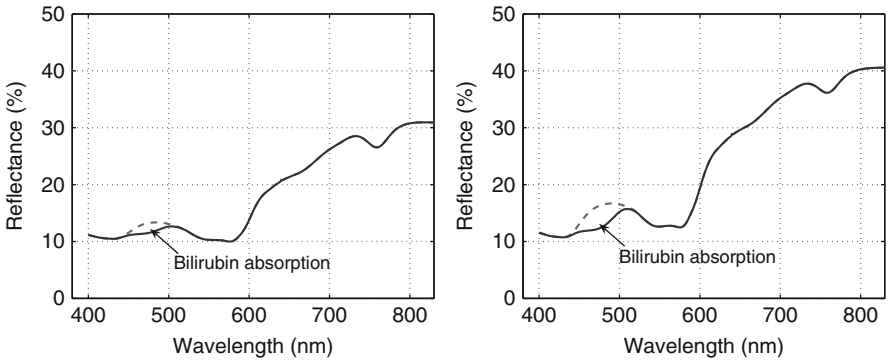


Fig. 22.8 Shielding of bilirubin absorption in bruises. Simulations of skin reflectance in 5 day old bruises. The *dashed lines* indicate the reflectance spectrum of the bruise with no bilirubin absorption added. All parameters, including the bilirubin absorption, are kept equal except the starting blood fraction of the bruise. *Left panel*: High hemoglobin absorption (Starting blood fraction in bruise: 0.9); *Right panel*: Reduced hemoglobin absorption (Starting blood fraction in bruise: 0.45)

22.3.1.2 Photobleaching of Bilirubin in Skin

In the fifties an observant nurse discovered that jaundiced newborns got better if they were exposed to sun shining in through the windows. Soon after that, light treatment was established as a therapy for neonatal jaundice [37, 38]. Neonatal jaundice is caused by accumulation of hydrophobic bilirubin in the body after birth causing a yellow skin tone and in the more severe cases a risk of brain damage. Light in the blue/green part of the spectrum induces a configuration change in the bilirubin molecules that make them more water soluble [38]. In this context it can be assumed that bilirubin in bruises in sun exposed skin areas will behave differently from bilirubin in bruises in non-exposed areas of the body. This hypothesis is partly supported by the data shown in the left panel of Fig. 22.5. This graph shows the temporal development of the bilirubin index of two bruises of comparable size at different locations of the body. The bilirubin index in the same person showed similar shape and peaked at the same point in time regardless of the position on the body, but the absolute level varied. This hypothesis needs to be further investigated by controlled experiments to be fully confirmed. In addition to this effect, sun exposed skin will also have increased pigmentation compared to non-exposed skin, such changes will also contribute to an altered appearance of bruises in sun exposed areas of the body.

22.3.1.3 Reflectance Spectra, Observed Bruise Color and Depth of the Injury

Although several authors have described the biochemical changes in skin following mechanical irritation, findings have traditionally not been related to the macroscopic skin and tissue changes that can be observed following such injuries, and even fewer

authors have considered the connection between optical changes and macroscopic tissue changes. Bohnert et al. [12], published a relation between reflectance spectra, observed bruise color and depth of the injury. They found that superficial hemorrhages cause bright red bruises, while deep subcutaneous bruises appear blue or purple in post mortem examinations of bruises. The observed blue color of deep bruises is caused by a combination of the optical properties of skin and a visual illusion. The effect is similar to the appearance of blue veins in skin explained so excellently by Kienle et al. [39]. The effect is caused by the difference in penetration depth of red and blue light, due to differences in both absorption and scattering. Red light penetrates deeper into the skin and is readily absorbed by deoxyhemoglobin in deep hematomas. The eye detects less reflected red light from the bruised areas compared to the adjacent skin, and the brain decides that the skin looks blue. If the bruised skin is shielded from the surrounding healthy skin, it will appear to have a more grayish hue. In darker skinned individuals, bruises are more difficult to identify, and as for the vessels, they may have a more greenish appearance than in lighter skinned individuals. This is due to the melanin absorption shifting the reflectance spectrum toward longer wavelengths.

The results presented in [30] also illustrate the relation between the depth of the injury and the perceived color of the skin. Paintballs are light weight projectiles that hit the skin with high speed (approx 100 m/s). Paintball injuries develop rapidly, and have been found to be fully established within 20 min after injury [30]. The energy from the paintballs is dissipated close to the skin surface due to the light weight, high speed, and short interaction time with the skin. Superficial oscillations are set up and cause tearing of smaller vessels at depths of approximately 200–300 microns. This causes a predominantly red color of the bruise (see Fig. 22.4). Heavier instruments will cause deeper damage to the skin, and the bruises may not even be visible until some days after injury. In this case the skin will be compressed and not torn, and will therefore not be damaged by the impact. Most real bruises are caused by tearing of vessels due to shear forces, and the damage may appear at all depths, depending on the nature of the impact. Very severe damage is associated with both deep hemorrhage, and small, circular, red petechial hemorrhages at the skin surface.

22.3.1.4 Color Coordinates

The CIE Lab color coordinate system [40] has been successfully applied in dermatology and medicine to describe skin color and color changes. This system can also be utilized in forensic medical optics to achieve a consistent and objective description of skin color. This system is adjusted to fit the response of the human visual system and consists of three parameters that describe the color of an object: L describes the lightness (100 = white, 0 = black), a describes the axis from green to red with negative values for green colors, and b describes the axis from blue to yellow with negative values for blue color. Several authors have explored the use of color coordinates to describe the color of bruises in either living persons or post mortem [12, 19, 33, 34, 41–43]. The conclusion of these studies is that colorimetry

is a good tool to describe skin color in an objective manner, and to assess changes in skin color. The ratio of red and blue color is related to the depth of the bruise. Color coordinates can therefore, at least in principle, be used to tell if the bruise is deep or shallow. In most cases such evaluation can be done visually or, even better, by examining the full reflectance spectrum of the bruise. The color coordinate *b* can be used to determine whether the bruise contains yellow color, although it is not a suitable method for accurate determination of the age of a bruise. Variations in other skin optical properties might also affect the skin color and thereby the color coordinates. As an example, a large dietary intake of carrots increases beta-carotene content in the skin which increases the *b* coordinate, which may be interpreted as bilirubin. This is especially a risk if only the bruise and not the adjacent normal skin is measured.

In fresh bruises, the Lab color coordinates are closely related to physical parameters like the dermal blood volume fraction, the tissue oxygenation and the erythema index. This has been shown by an in vivo study on porcine skin (see [30] for methodology). The lightness parameter *L*, is negatively correlated to the dermal blood volume fraction, *Bd*, and to the erythema index, *Ec*, and positively correlated to the oxygenation of the bruised area. The *a* parameter is not correlated with the oxygenation but is positively correlated with the dermal blood volume fraction and the erythema index, which means that the color of the skin gets redder with increasing blood content, as expected. The *b* parameter is strongly correlated to the oxygenation of the bruised area since the skin color is closely related to the ratio of oxy- and deoxyhemoglobin in the bruised area. It is important to keep in mind that these parameters also will be influenced by the local skin pigmentation [44]. The correlation coefficients are given in Table 22.1, and a graphical illustration of the relation between the dermal blood volume fraction, *Bd*, and the color coordinate *L* is shown in Fig. 22.9.

Table 22.1 Correlation between CIE Lab color coordinates and physical skin parameters measured on fresh in vivo porcine bruises

	Oxy	*Bd*	*Ec*
L	0.72	−0.91	−0.81
a	–	0.74	0.92
b	0.85	−0.48	–

Oxy = tissue oxygenation estimated at 660/805 nm, *Bd* = dermal blood volume fraction, *Ec* = erythema index. Only correlations with *p*-values less than $p <= 0.002$ are shown

22.3.2 Biomechanical Modeling of Bruised Skin

Tissue is a mixture of solid and fluid phases, which makes the modeling and analysis of the mechanics of skin bruises a challenging task. The problem of fitting biological material model parameters is present, and the in vivo versus in vitro testing needs

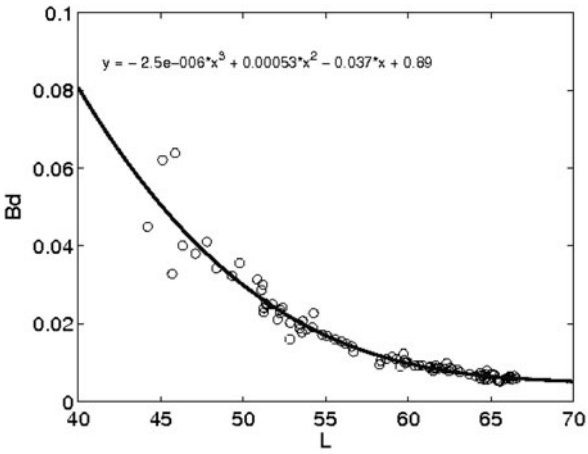


Fig. 22.9 Relation between the dermal blood volume fraction, *Bd*, and the CIELab color coordinate *L* in bruised porcine skin. A description of the experimental methods used in this study is published in [30]

careful consideration. In the case of bruising, a full dynamic analysis is required due to fast impact trauma. One approach is to analyze the elastic response of skin by means of hyperelastic and viscoelastic theories. One should note that in this case a single phase solid is assumed, and the viscoelasticity accounts for the fluid transport in the tissue.

Wu et al. [45] found that in a hyperelastic-viscoelastic approach the bulk modulus was high, although the (single phase) material was compressible. The viscoelasticity may be modeled by separating the stress tensor into a term due to pure hyperelasticity and one viscous term. The hyperelastic contribution describes the instantaneous response, whereas the viscous term describes the delayed tissue response.

Another approach to this problem can be bi-phasic modeling which employs a solid phase modeled by means of hyperelasticity (the solid soft tissue matrix), and a fluid phase modeled by some permeability function. This approach has a long tradition in geomechanics, but has also been used extensively in biomechanics related to cartilage [46]. Oomens et al. [47] have extended this approach to skin. A numerical hyperelastic-viscoelastic analysis of the impact of a ball fired onto the skin from a paint-ball gun is shown in Fig. 22.10.

The location of the highest shear stresses given in this figure corresponds very well with the observed radius of the red ring which can be observed at the skin surface after being hit by a paintball (see Fig. 22.4). Thus, this analysis gives reasonable estimates of the locations of regions with the highest damage to the vasculature (vessel ruptures).

22.3.3 Transport of Hemoglobin Species in Tissue by Convection and Diffusion

The models presented in this paragraph have previously been published in refs. [8] and [29].

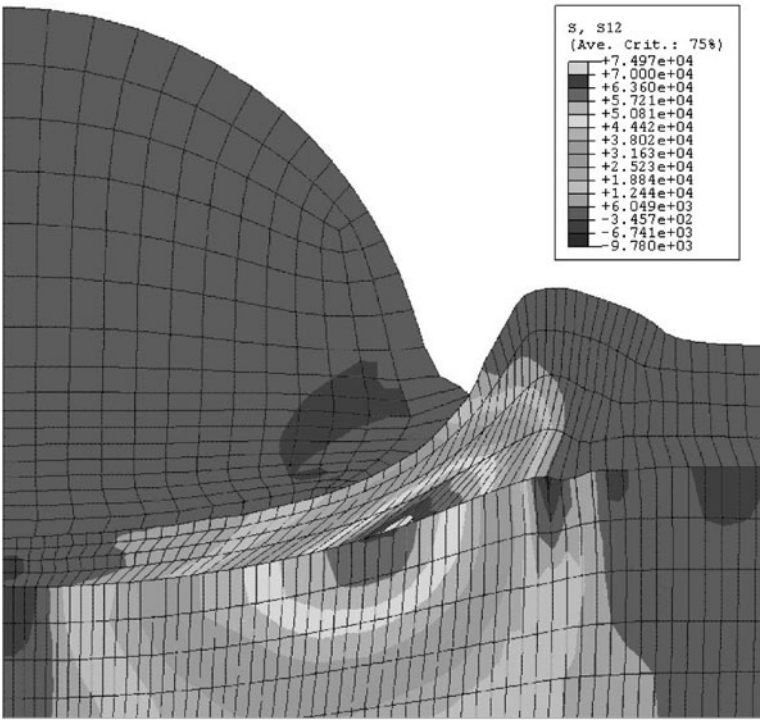


Fig. 22.10 Contour plot of shear stresses on the deformed mesh at end of simulation. Data from [48]

Minor trauma will often result in a hemorrhage due to tearing of smaller vessels such as arterioles and venules. Such localized damage may result in pooling of blood in subcutaneous tissues, and eventually transport of blood away from the damaged area. Blood can easily flow along muscle bundles, membranes, as well as along and across muscle fibers. The driving force of this flow will be the pressure of the blood from the damaged veins or arteries together with gravitational effects.

This transport can be described by Darcy's law

$$j_i = -K_{ik} \frac{\partial p}{\partial x_k} \tag{22.1}$$

where j_i is the components of the blood flux vector, p is the local pressure, x_i is the spatial coordinate of a Cartesian system and K_{ik} is the hydraulic conductivity tensor. In isotropic tissues $K_{ik} = K\delta_{ik}$, where δ_{ik} is the Kronecker delta, the law can then be expressed

$$\vec{j} = -K \nabla p \tag{22.2}$$

where K is the hydraulic conductivity of whole blood. The blood can be considered as an incompressible liquid, thus the continuity of the blood flux vector can be expressed

$$\nabla \cdot \vec{j} = 0 \tag{22.3}$$

In a homogeneous isotropic tissue with a lesion of spherical geometry the flow outside the lesion can be expressed

$$\vec{j} = \frac{Q}{4\pi r^2} \cdot \frac{\vec{r}}{r} \tag{22.4}$$

where Q is the total discharge rate of blood from the lesion and r is the distance from its center. The local pressure p at an arbitrary position r from source can be expressed by Eqs. (22.2) and (22.4)

$$p = \int_r^\infty \frac{Q}{4\pi K r^2} dr = \frac{Q}{4\pi K r} \tag{22.5}$$

where the pressure at infinity is set equal to zero. The time Δt required for blood to saturate a volume of radius c is given by

$$\Delta t = \frac{4\pi}{3Q}(c^3 - a^3)\rho \tag{22.6}$$

where ρ is the porosity of the tissue, i.e., the fractional volume of tissue that is filled with extravascular blood. Thus, the time can be expressed, by Eqs. (22.5) and (22.6)

$$\Delta t = \frac{(c^3 - a^3)\rho}{3aK\Delta p} \tag{22.7}$$

where Δp is the pressure difference between the blood-pool and the region of radius c. The radius of the blood pool is a. Equation (22.7) is also valid for flow from a sub-dermal hemispheric pool of blood perfusing radially into subcutaneous tissues. The corresponding expression in the case of linear flow along a channel of constant cross-section becomes, by Eq. (22.2)

$$\Delta t = \frac{L\rho}{K\Delta p} \tag{22.8}$$

where L is the length of the channel and Δp is the pressure drop along the channel.

Values of the hydraulic conductivity across the abdominal muscle have been reported to increase with abdominal pressure, and are in the range $K = 0.1 - 0.6 \cdot 10^{-12}$ m^4/Ns for pressures from 2–20 mmHg [49]. Based on the present patient data the hydraulic conductivity was determined to be in the

region $K = 0.1 - 0.5 \cdot 10^{-10}$ m^4/Ns. It is, however, not unexpected that the component of the hydraulic conductivity along the muscle fiber and membranes can be significantly larger than the corresponding value across these structures. The value is still small as compared with other media, e.g., the hydraulic conductivity of sandstone aquifer is $K = 1 \cdot 10^{-9}$ m^4/Ns [50]. The hydraulic conductivity is set to $K = 1 \cdot 10^{-11}$ m^4/Ns, the porosity $\rho = 0.1$ and $\Delta p = 2.6$ kPa (20 mmHg). The pressure, which here corresponds to a typical venous pressure, can be much higher if arteries/arterioles are damaged. Further on, if perfusion over larger distances is considered, gravitational contribution to the pressure gradient must be considered.

The results give the time required for blood to perfuse from a pressurized region of 5 mm diameter to a distance of up to 30 mm. The diameters of typical bruises are in the range 20–60 mm, and it can thus be concluded that the subcutaneous perfusion of blood will take place within a few hours.

The use of Darcy's law, which is based on quasi-static conditions, is a coarse simplification of the dynamics. However, it is expected to be good enough to predict a proper order of magnitude of the time scale.

22.3.3.1 Transport of Hemoglobin in Dermis

The extravascular transport of hemoglobin in tissue can be expressed by convection and diffusion

$$j_i = -D_{ik}\frac{\partial N}{\partial x_k} - \xi K_{ik}\frac{\partial p}{\partial x_k} \tag{22.9}$$

where j_i is the component of the hemoglobin flux vector, D_{ik} is the hemoglobin diffusivity tensor, K_{ik} is the Darcy constant for whole blood and ξ is the volume fraction of hemoglobin in whole blood, i.e. $\xi = 0.14$–0.15 corresponding to a hemoglobin concentration of 140–150 g/l. Hemoglobin will be transported into dermis by diffusion from the subcutaneous layer of whole and hemolyzed blood. The transport mechanism is in the present model assumes diffusion. When the Darcy term in Eq. (22.9) is neglected, the transport flux vector in an isotropic medium becomes

$$\vec{j}_H = -D_H \nabla N_H \tag{22.10}$$

where D_H is the hemoglobin diffusivity and N_H is the density of molecular hemoglobin. The continuity of hemoglobin in a medium with loss can be expressed

$$\nabla \cdot \vec{j}_H = -\frac{\partial N_H}{\partial t} - \frac{N_H}{\tau_H} \tag{22.11}$$

where t is time and τ_H is the hemoglobin relaxation time. This lifetime is determined by several mechanisms such as drainage by the lymphatic system, conversion to biliverdin/bilirubin or consumption by macrophages. The relaxation time can be expressed

$$\frac{1}{\tau_H} = \frac{1}{\tau_L} + \frac{1}{\tau_B} + \frac{1}{\tau_M} + \ldots\ldots \tag{22.12}$$

where τ_L, τ_B, and τ_M are the relaxation times due to lymphatic drainage, conversion to biliverdin/bilirubin and to macrophage activity, respectively. This model is a simplification of the real, complex situation found in living tissue. This model will need to be refined as the knowledge about the interaction and the relative importance of the physiological processes increase.

There is very limited information on in vivo hemoglobin diffusivity in the literature. However, the diffusivity of myoglobin in skeletal muscle cell is reported to be $1.2 \cdot 10^{-11}$ m^2/s [51]. The molecular weight of molecular hemoglobin (66,000 Dalton) is approximately four times as large as that of myoglobin (17,000 Dalton). Therefore, the diffusivity of hemoglobin in skin is expected to be less than this value, and fitting the model with patient data indicate a diffusivity in the range of $D_H = 3 \cdot 10^{-12}$ m^2/s.

The relaxation time can be determined from experimental data. The data available at the time suggests values in the range of $\tau_H = 1.7 - 2.6 \cdot 10^5$ s (2–3 days). For elderly people, this value can be expected to be somewhat longer [11, 52]. The observed value will depend on factors like the immune response; and based on observations, it is expected to be shorter in children compared to adults. The relaxation time is also observed to be significantly less in a region of 5–10 mm around damaged vessels, causing a white spot in the most severely damaged regions as described in Section 22.2.2.3. This whitish region forms about 0.5–1 days after injury and expands with time. This phenomenon, which is assumed to be caused by a stronger immune response in vicinity of damaged vessels, reduces the hemoglobin relaxation time to $\tau_H = 0.2 - 0.4 \cdot 10^5$ s (0.25–0.5 days). The temporal development of the hemoglobin and bilirubin content in skin is shown in Fig. 22.11 for a

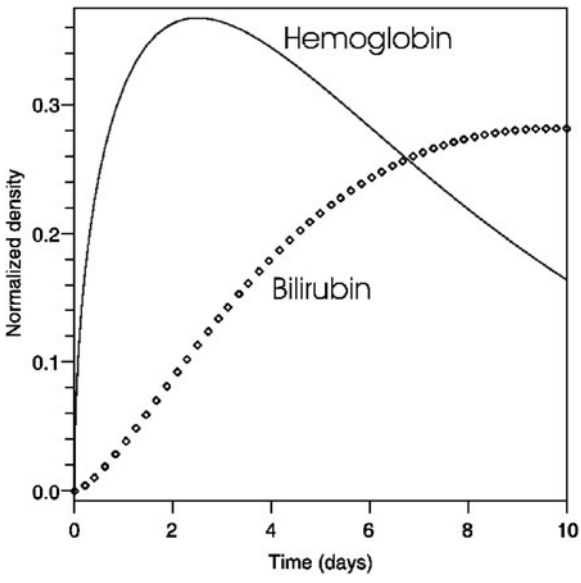


Fig. 22.11 Hemoglobin and bilirubin densities in dermis predicted by the mathematical model

relaxation time of 5 days for hemoglobin, and 6 days for bilirubin. These values are probably most suitable for elderly people, since the numbers were based on patients enrolled into hospital for cardiac surgery [8].

22.3.3.2 Transport and Generation of Bilirubin in Dermis

Diffusion of bilirubin in dermis can be expressed by a transport flux vector $\vec{j}_B$

$$\vec{j}_B = -D_B \nabla N_B \tag{22.13}$$

where D_B is the bilirubin diffusivity and N_B is the density of molecular bilirubin. The transition from biliverdin to bilirubin is fast compared to other timescales of the system. Therefore, the generation of bilirubin can be characterized by the time τ_B, which characterizes the hemoglobin to bilirubin/biliverdin transition.

The continuity of bilirubin can be expressed

$$\nabla \cdot \vec{j}_B = -\frac{\partial N_B}{\partial t} - \frac{N_B}{\tau_{BR}} + \frac{N_H}{\tau_B} \tag{22.14}$$

where τ_{BR} is the bilirubin relaxation time. If diffusion of bilirubin can be neglected Eq. (22.14) is simplified to

$$\frac{\partial N_B}{\partial t} = -\frac{N_B}{\tau_{BR}} + \frac{N_H}{\tau_B} \tag{22.15}$$

22.3.3.3 Planar Skin Model

The diffusion of hemoglobin into dermis can be modeled as flow in a planar model. Epidermis is composed of very densely packed keratinocytes that form a diffusion barrier. The boundary condition at the basal layer can therefore be taken as zero flux. The sub-dermal hemoglobin source can be taken to be instantaneously established at time t=0 since the diffusion process is much slower than the process for distribution of blood in subcutaneous tissue. The spatial and temporal distribution of hemoglobin in dermis $N_H(x,t)$ can be expressed, by Eqs. (22.10) and (22.11)

$$\begin{aligned} N_H(x,t) = N_{H0}\Bigg\{ \sum\nolimits_{n=0}^{\infty} \Bigg[(-1)^n erfc\left(\frac{1}{2}\frac{(2n+1)d - x}{\sqrt{D_H t}}\right) \\ + (-1)^n erfc\left(\frac{1}{2}\frac{(2n+1)d + x}{\sqrt{D_H t}}\right) \Bigg] \Bigg\} e^{-\frac{t}{\tau_H}} \end{aligned} \tag{22.16}$$

where N_{H0} is the subcutaneous density of hemoglobin, x is the distance from the basal layer, and d is the dermal thickness. In simulations of real bruises, the parameter N_{H0} is estimated based upon the size of the bruise. As a rule of thumb, large bruises (diameter > 5 cm), are simulated using $N_{H0} = 0.9$, medium sized bruises (diameter 5–2 cm) are simulated using $N_{H0} = 0.4$, and small bruises

(diameter <2 cm) are simulated using $N_{H0} = 0.2$. Variations of these numbers may be required for complex injuries. The relaxation time of dermal hemoglobin and hemoglobin in the subcutaneous region are both assumed to be τ_H. This assumption is made in order to derive a reasonably simple analytic solution. If a larger complexity is introduced in the model, it would be reasonable to switch to numerical methods to solve the problem. The corresponding value for the average hemoglobin concentration $\bar{N}_H(t)$ in dermis can be expressed by integration of Eq. (22.16) over dermis

$$\bar{N}_H(t) = 2N_{H0}\sqrt{\frac{D_H t}{d^2}} \times \left\{ \frac{1}{\sqrt{\pi}} - 2\sum_{n=1}^{\infty}(-1)^n \frac{dn\left(erfc(\frac{dn}{\sqrt{D_H t}})\sqrt{\pi} - e^{-\frac{d^2n^2}{D_H t}}\sqrt{D_H t}\right)}{\sqrt{\pi D_H t}} \right\} e^{-\frac{t}{\tau_H}} \tag{22.17}$$

The series in Eq. (22.16) converges rapidly, and for time scales where $t < 4d^2/D_H$ the first term gives satisfactory accuracy

$$N_H(x,t) = N_{H0}\left\{ erfc\left(\frac{1}{2}\frac{d-x}{\sqrt{D_H t}}\right) + erfc\left(\frac{1}{2}\frac{d+x}{\sqrt{D_H t}}\right)\right\} e^{-\frac{t}{\tau_H}} \tag{22.18}$$

The first of these terms represents diffusion into a semi-infinite medium and the second term represents a mirror source ensuring zero flux at the basal layer. This expression represents the same solution as in Eq. (22.16), where the contribution from an infinite number of mirror sources is given by the summation. For shorter times the series can be truncated, resulting in Eq. (22.18). The corresponding approximation for the average dermal hemoglobin distribution follows from integration of Eq. (22.17)

$$\bar{N}_H(t) = \frac{2N_{H0}}{d\sqrt{\pi}}\left\{ \sqrt{D_H t} - \sqrt{D_H t}e^{-\frac{d^2}{D_H t}} + d\sqrt{\pi}\,erfc\left(\frac{d}{\sqrt{D_H t}}\right)\right\} e^{-\frac{t}{\tau_H}} \tag{22.19}$$

The corresponding solution for bilirubin distribution in presence of diffusion in dermis can be established by forming a Green's function from the hemoglobin distribution, together with a mirror-distribution that establishes the criterion of no bilirubin transport into the epidermis

$$N_B(x,t) = \int_{\tau=0}^{t}\int_{\xi=0}^{d} N_H(\xi,\tau)\frac{\left(e^{-\frac{(x-\xi)^2}{4D_B(t-\tau)}} + e^{-\frac{(x+\xi)^2}{4D_B(t-\tau)}}\right)e^{-\frac{t-\tau}{\tau_{BR}}}}{2\tau_B\sqrt{\pi D_B(t-\tau)}}d\xi\,d\tau \tag{22.20}$$

where $N_H(\xi,\tau) = N_H(x \Rightarrow \xi, t \Rightarrow \tau)$ from Eq. (22.16) or (22.18)

Diffusion into dermis of bilirubin formed from subcutaneous hemoglobin can be found by adding a corresponding term to Eq. (22.20),

$$N_B(x,t) = \int_{\tau=0}^{t} \int_{\xi=0}^{\infty} [N_H(\xi,\tau)H(d-\xi) + N_{H0}e^{-\frac{\tau}{\tau_H}}H(\xi-d)] \times \frac{\left(e^{-\frac{(x-\xi)^2}{4D_B(t-\tau)}} + e^{-\frac{(x+\xi)^2}{4D_B(t-\tau)}}\right)e^{-\frac{t-\tau}{\tau_{BR}}}}{2\tau_B\sqrt{\pi D_B(t-\tau)}} d\xi\, d\tau \tag{22.21}$$

where *H(x)* is the Heaviside unit step function. If all diffusion of bilirubin can be neglected, the solution of Eqs. (22.20) and (22.21) can be expressed

$$N_B(x,t) = e^{-\frac{t}{\tau_{BR}}} \int_{\tau=0}^{t} \frac{N_H(x,\tau)e^{\frac{\tau}{\tau_{BR}}}}{\tau_B} d\tau \tag{22.22}$$

where $N_H(x,\tau) = N_H(x, t \Rightarrow \tau)$ from Eqs. (22.16) or (22.18).

When diffusion within the dermis is present but diffusion across the dermal/subcutaneous tissue border layer is neglected, the average bilirubin concentration can be expressed

$$\bar{N}_B(t) = e^{-\frac{t}{\tau_{BR}}} \int_{\tau=0}^{t} \frac{\bar{N}_H(\tau)e^{\frac{\tau}{\tau_{BR}}}}{\tau_B} d\tau \tag{22.23}$$

where $\bar{N}_H(\tau) = \bar{N}_H(t \Rightarrow \tau)$ from Eq. (22.19). For small values of time, i.e. $t < d^2/D_H$, Eq. (22.19) can be approximated

$$\bar{N}_H(t) = \frac{2N_{H0}}{d}\sqrt{\frac{D_H t}{\pi}}e^{-\frac{t}{\tau_H}} \tag{22.24}$$

and the corresponding value for the average bilirubin concentration becomes Eqs. (22.23) and (22.24)

$$\bar{N}_B(t) = \frac{N_{H0}\tau_{BR}\tau_H}{\tau_B(\tau_H-\tau_{BR})d}\left\{2\sqrt{\frac{D_H t}{\pi}}e^{-\frac{t}{\tau_H}} - erf\left(\sqrt{\frac{(\tau_H-\tau_{BR})t}{\tau_{BR}\tau_H}}\right)\sqrt{\frac{D_H\tau_{BR}\tau_H}{\tau_H-\tau_{BR}}}e^{-\frac{t}{\tau_{BR}}}\right\} \tag{22.25}$$

22.3.4 Photon Transport in Bruised Tissue

The photon transport in bruises can be described by a variety of techniques, e.g. numerical methods like Monte Carlo simulations or finite element models, or analytical methods like diffusion theory. This section will briefly describe an analytic approach based on the diffusion approximation [13, 29, 53–55].

22.3.4.1 Analytic Photon Transport Model for Skin

This reflectance model was originally developed for analysis of port-wine stain birthmarks and has also been applied to other skin conditions like neonatal jaundice [13, 53, 54, 56], and has been proven to agree well with Monte Carlo simulations [54]. The model has also been applied to bruises [8, 23].

The model describes two plane layers on a region of semi-infinite extent. The first layer simulates the epidermis and the upper part of papillary dermis, the second layer represents the upper part of dermis and the last semi-infinite region represents the deeper parts of dermis. The chromophores are assumed to be uniformly distributed within each layer. Melanin absorption is calculated using the model given by Svaasand et al. [13], and the hemoglobin absorption coefficient is based on spectra given by W.G. Zijlstra et al. [27]. The scattering properties are assumed to be known and were modeled using the data from Saidi et al. [57] scaled to fit adult skin [53].

Volume fractions of the present chromophores, blood oxygenation, scattering properties and skin thickness, are required parameters for the reflectance model. The required optical properties of skin, i.e. pigmentation, blood volume and blood oxygenation can either be set directly in the model or be determined iteratively from measured spectra using the reflectance model in inverse mode. A thorough description of modeling of the individual optical parameters, and how they can be derived from measured spectra, can be found in [29].

Monte Carlo techniques and finite element modeling may be applied to model more complex geometries.

22.4 Optical Identification and Characterization of Bruises

22.4.1 Reflection Spectroscopy

Reflection spectroscopy is an old but still convenient method for optical characterization of skin and skin lesions. Modern technology has provided the market with small, easy manageable fiber connected spectrometers that can be carried to the patient's bedside without much effort. This makes reflectance spectroscopy an appropriate technique for measurements where the scientist visits the subjects on site. In characterization of bruises this is an important advantage, since the measurements can be performed at the location where the subjects are found, for example in hospitals. One of the largest disadvantages of the method is that, in most configurations, contact is required between the reflectance probe and the skin. Several probe designs can be used to acquire reflectance spectra, and the only required items in addition to a spectrometer and a laptop computer are a light source, some sort of sensor to collect the reflected light, and a reliable method of calibrating the measurement against a reflectance standard. The largest problem with the collection optics has been to collect enough of the backscattered red light. Blue light is more readily scattered than red light, thus increasing the probability of collecting

backscattered blue light compared to scattered red light. The red light might be scattered out of the sampling volume. Integrating spheres are well suited to give high quality reflectance spectra, since the large aperture of the sphere also collects the backscattered red light. However, the spheres are a bit cumbersome to use, due to the often bulky shape and their physical size. They also make it difficult to control the measured area accurately. Fiber probes are easy to place correctly, but have traditionally collected less backscattered red light. However, new probe designs with properly spaced illumination and detection fibers have become a good alternative to the integrating spheres. The small size of the fiber probes facilitates the correct placement of the probe during measurement. Since both kinds of probes require some sort of contact between the skin and the probe, it is important to keep in mind that measurements should always be conducted in the same way. A change in the pressure with which the probe is applied to the skin will affect the measurements significantly (Fig. 22.12). This measurement shows that if pressure is applied to the probe, the area will be blanched and the features in the reflectance spectra will be altered with respect to the natural condition of the skin. This might be a good thing if you are looking for something other than blood; but in measurements of bruises, the blood absorption is considered to be one of the key features. In measurements of bruises one must be careful to place the probe in the right position in the bruise and in the normal skin. The normal skin measurement must be collected at a sufficient distance to avoid contributions from, e.g., bilirubin diffusing from the injured zone. The measurement of the bruise itself should be made in a homogenous region of the bruise, and the whitish central zone forming in many bruises should be avoided.

Reflectance spectroscopy is a volumetric technique which gives a picture of the average optical properties of the skin in the sampled volume. Best results are obtained when reflectance data is analyzed with an optical model. Another large drawback of the technique is the lack of spatial resolution, which may be overcome by switching to imaging techniques like hyperspectral imaging.

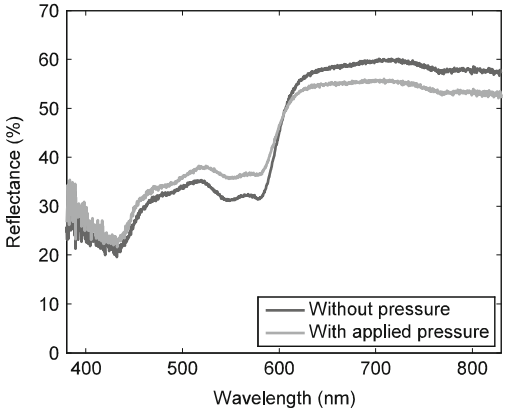


Fig. 22.12 The effect of applying pressure to the integrating sphere during reflectance measurements. Data collected at the volar side of the forearm of a female volunteer

22.4.2 Hyperspectral Imaging

Hyperspectral imaging is a modality which combines imaging and spectroscopy. In short, the technique provides an image of the object, with the possibility of collecting full spectra in each pixel of the image. Most cameras detect the radiance from the scene and must be calibrated by the use of a reflectance standard to give the reflectance of the object. The technique comes in many varieties, and different names are used for the same technique. Chemical imaging, hyperspectral imaging and multi-spectral imaging are terms that have been used for related techniques. The imaging principle for hyperspectral imaging is shown in Fig. 22.13. The technical realization of this method varies as well. Some cameras apply a tunable filter to collect data from selected wavelength bands, and collect data from the whole scene for each setting of the filter. Other instruments are scanning cameras that collect all the needed wavelengths in one go using a diffraction grating and a detector array configuration, but must be physically scanned to collect data from the whole scene. These are non-contact techniques, and can be used for collecting data from larger skin areas. This gives the possibility of scanning whole bodies to detect bruised areas. Payne et al. [58] have shown that hyperspectral imaging is equal to or better than traditional spectroscopic techniques to detect bilirubin in a mixture with hemoglobin. They applied hyperspectral imaging and spectroscopy to detect hemoglobin and bilirubin mixtures on cloth and in pieces of skin. Randeberg et al. [59] applied hyperspectral imaging to fresh bruises in a porcine model, and used traditional spectroscopic analysis to monitor the oxygenation of the bruised area following the injury.

Hyperspectral data can be analyzed using several methods since these data consist of both full spectral information and images giving spatial resolution, giving the possibility of applying both traditional spectroscopic analysis as mentioned above, and image analysis techniques like, for example, principal component analysis (PCA) [59, 60]. Image analysis can be used to segment and enhance bruises

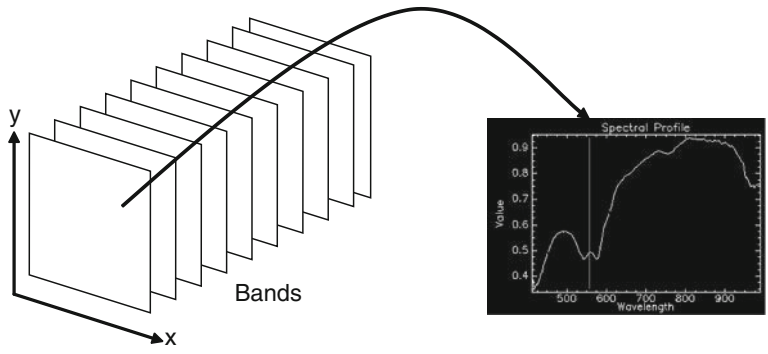


Fig. 22.13 Imaging principle for hyperspectral imaging. The data cube consists of two spatial dimensions and one dimension holding the spectral information

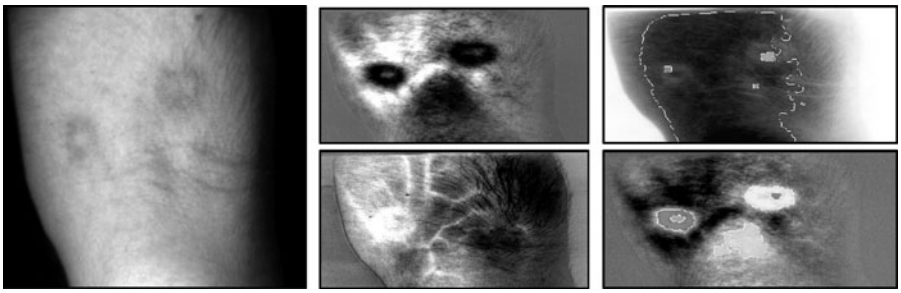


Fig. 22.14 Enhancement and segmentation of bruised skin. *Left panel*: Original image (573 nm); Middle panel: Results from PCA analysis. *Right panel*: Watershed segmentation of the bruised areas based on the original image (*top*), and the PCA analysis (*lower panel*). Data from [60]

in the images before application of spectroscopic techniques for characterization of the identified areas. The data shown in the left panel of Fig. 22.5 was collected from hyperspectral images. An example of segmentation of bruises from hyperspectral images is shown in Fig. 22.14.

22.4.3 Post Mortem Bruises

The examination for possible bruises may be performed at autopsy, with authors having reported the use of spectroscopy for characterization of bruises and differentiation from other colored marks (such as lividity) [12, 41, 61, 62]. Bruises might appear different after death due to re-oxygenation of hemoglobin during refrigerated storage [12, 61]. Furthermore, after death bruises will not undergo the enzymatic breakdown processes of hemoglobin that occur in the living. The reflectance spectra measured on skin post mortem will be affected by the site of the bruise since blood will drain from the upper surfaces of the body. Discolorations due to post mortem degradation processes will make it more difficult to identify bruising in a body [12]. A brownish color that may be due to methemoglobin is often observed in bruised skin after death. An example of a reflectance spectrum collected post mortem from a bruise in the upper arm of a male is shown in Fig. 22.15 together with a spectrum of the adjacent normal skin. The oxygenation of the hemoglobin was found to be approximately 10% in the bruise, and quite high in the normal skin (approximately 90%). However, most of the blood in the normal skin had been drained from the area due to the position of the body. Due to this fact, the measured spectra may also contain information about other chromophores such as cytochromes and myoglobin. The present bruise was measured approximately 24 h after death. At this point in time, no oxidation of hemoglobin to methemoglobin can be observed from the spectra. The blood that was drained from the upper parts of the body accumulated in the lower parts of the body, causing the typical blue marks of post mortem hypostasis. Such marks on the skin might appear similar to bruises.

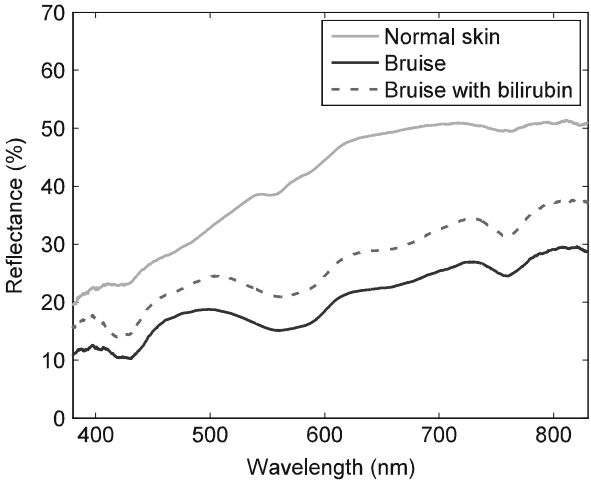


Fig. 22.15 Post mortem bruise and normal skin. The bruise was located on the dorsal side of the over arm

22.4.4 Bruises in Living People

22.4.4.1 Bruises in Children

Several studies have concluded that it is impossible to age bruises in children accurately using traditional methods like visual inspection [5, 7, 63]. A recent study has launched optical spectroscopy as a method to characterize and date bruises in adults objectively [8]. Some authors have suggested looking for patterns in bruising rather than for the color of the bruises in children [64]. It is not given that the methods that are developed for adults will work in children. In general, there are several reasons for developing dedicated methods for bruises in children. This section will review some of the main differences between bruises in adults and discuss some of the challenges regarding accurate optical methods for detection and characterization of bruises in children.

The skin optical properties are different in children compared to adults. Newborn skin has been studied in relation to optical diagnosis of neonatal jaundice [53, 57]. The properties of newborn skin will depend on gestational age, and after birth the skin will continue to develop. The reddish appearance seen in many Caucasian newborns is probably related to less connective tissue and little subcutaneous fat, compared to older children. Young children also tend to have thinner and more delicate skin than adults. Skin pigmentation is also less pronounced at birth than later in life, and develops according to hereditary patterns and sun exposure. Figure 22.16 shows the temporal development of the skin reflectance in the same individual as a function of time after birth. The left panel show a rapid increase of the reflectance during the first four months after birth. This child was born in the late spring, but was shielded against strong direct sun exposure during the summer months. The increase in the reflectance is probably caused by maturation of the collagen fibers in the skin and a buildup of a thicker subcutaneous fat layer. The right panel shows

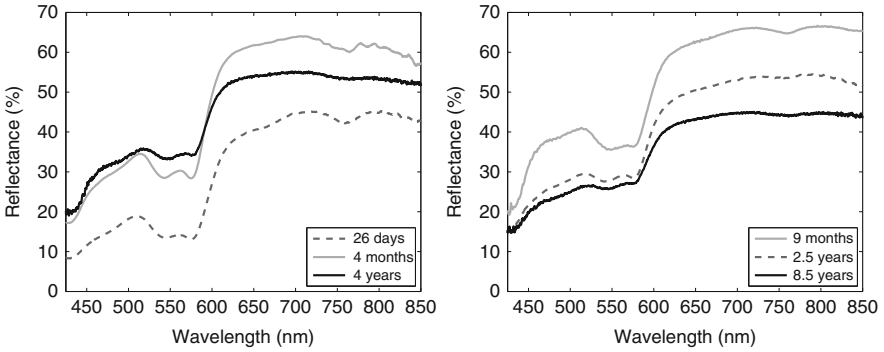


Fig. 22.16 Measured skin reflectance as a function of age. *Left panel*: Caucasian boy, *Right panel*: Caucasian girl

the development of the skin throughout childhood, where the development of the skin pigmentation with age is prominent, even in Caucasian individuals. This girl is a Fitzpatrick skin type III. The lowering of the reflectance is probably also due to less scattering caused by a thicker epidermis and a thinner subcutaneous fat layer after infancy. These examples illustrate large changes in optical properties even in the same children over time. The task of determining the age of bruises in children therefore requires special attention.

Methods developed for adult skin are thus not directly applicable for bruises in children due to differences in skin optical properties and a hypothesized difference in physiological behavior of bruises in children and adults. Bruises in children seem to develop and fade faster than in adults. Faster development of bruises in children compared to adults would be expected due to the thinner and more delicate skin resulting in shorter diffusion times for the blood from the hemorrhage to reach visibly detectable skin depths. Vanezis [11] claims that children exhibit an increased tendency to bruise early in life, due to looser connective tissue and delicate skin quality. It would be expected that bruises in children fade faster than in adults, and preliminary data on children suggests that bruises in children might fade as much as 2–3 times faster than in adults. Figure 22.17 shows measured spectra from a bruise in a nine year old girl, and simulated spectra trying to predict the temporal development of a bruise. The input parameters for the simulations were found using the analytical model presented in Section 22.3.4.1 in inverse mode.

The build up phase of the bruise in Fig. 22.17 is quite similar to the simulated values, suggesting that this girl is older to have a thicker, more adult like skin with a longer build-up time than in a younger child. This assumption is supported by the spectroscopic measurements which show normal skin spectra similar to spectra from young, adult skin. The build-up phase is mainly determined by the thickness of the skin, the pressure in the leaking vessels and the diffusion of hemoglobin in skin. The only variable that depends strongly on age in this case is the skin thickness. The simulations are based on time constants for breakdown of bruises in adult skin, and they predict the breakdown phase to be longer for this bruise than indicated by the

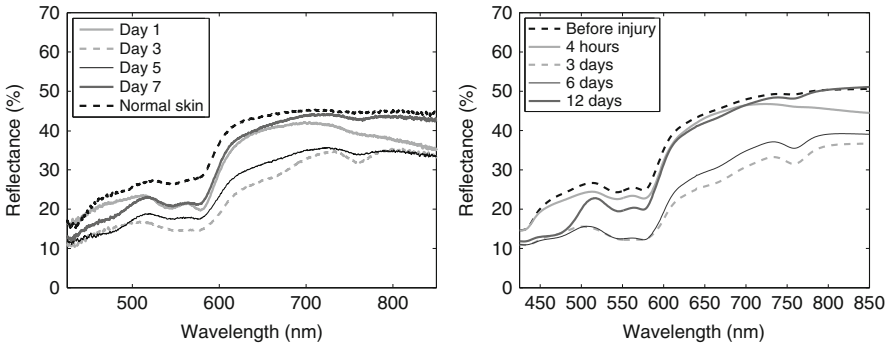


Fig. 22.17 Temporal development of a bruise in a child compared to simulated spectra. *Left panel*: Measured spectra from a bruise (diameter >5 cm) at the upper arm of a 9 year old girl. *Right panel*: Simulated temporal development of a bruise simulated using the models presented in Sections 22.3.3 and 22.3.4.1 using parameters for adult skin

measured spectra. This is probably related to a more efficient breakdown process and removal of the waste products in children than in adults. These results support the need for specific methods developed especially for children. The temporal development of bruises in children will also probably depend on the age of the child. In conclusion, more substantial data are needed to be able to develop a model that works for all ages.

22.4.4.2 Bruises in Adults

Most available spectroscopic data from bruises on living persons are collected from young adults. The tendency to bruise is related to specific parameters such as skin thickness and gender [11]. Women tend to bruise more easily than men due to a larger subcutaneous fat layer and less muscular tissue [11]. However, there will also be large differences between different groups regardless of gender. It has been claimed that redheads bruise more than dark haired or blonde individuals. It might be related to the fact that redheads tend to have less skin pigmentation than people with other hair colors. This makes it easier to identify bruises on fair skin than in more pigmented individuals. The tendency to bruise differs among the population. However, the temporal development is more similar between individuals than one should suspect. The temporal development of a bruise is determined by the skin thickness, the diffusivity of hemoglobin in tissue and the breakdown processes after the hemorrhage [8]. The diffusivity is assumed to be similar for all individuals, and the skin thickness can be measured and corrected for. The most important parameter is thus the breakdown process and the clearing of the hemoglobin. It has been shown that the temporal development of the bilirubin index in one individual is equal for multiple bruises at different locations in the same individual (see Section 22.2.2.2). Slight differences have been found between individuals. These differences are probably linked to individual differences in immune response after the injury.

22.4.4.3 Bruises in Elderly People

Child abuse is a strong motivation for developing new methods for identification, characterization and dating of bruises. Another, more neglected reason is the abuse of elderly people. It is predicted that by 2025 the population of people aged over 60 will be about 1.2 billion [65]. The World Health Organization states in their report on abuse of elderly people [65] that 4–6% of elderly people are abused in developing countries, while numbers of about 7% have been reported from countries like Finland and Canada, with more women being abused than men. Even if only a part of those cases are linked to physical abuse, it is a major problem for the society which need to be addressed; and quantitative, objective diagnostic tools for evaluating skin bruises will be an important contribution to fight this problem.

So far, bruises in elderly people have received little attention in the literature. Mosqueda et al. [52] published a study on the life cycle of bruises on older adults. They investigated the pattern of accidental bruising in 101 elderly people (over the age of 65), and found that most of the bruises occurred in the extremities, and that the mean diameter of the bruises were 3.42 cm. Most of the bruises were resolved on day 11, but some of the persisted as long as 41 days after injury. Yellow color was observed to increase during the first 6 days after injury. The skin in elderly people is different from skin in younger persons [9]. Sun damaged skin will become thinner, with less elasticity caused by a loss of collagen and elastin [66]. The skin of long time smokers will also be negatively affected by this habit [9]. Skin with less support from subcutaneous fat and connective tissue will bruise more easily [11, 66]. Many elderly people also suffer from diseases that either cause increased tendency to bruise or require treatment with anticoagulants.

If spectroscopy is used to characterize bruises in elderly people, there are some facts to be aware of. The perfusion in the skin of elderly people may be reduced compared to younger individuals. This is especially seen in the extremities, and even more frequent in persons suffering from heart or lung diseases. The spectra from normal skin will thus be dominated by deoxyhemoglobin. Oxygenations as low as 20–30% have been seen in normal skin in the legs of patients admitted for bypass surgery.

The loss of connective tissue can be observed in the reflectance spectra as a significantly reduced reflectance for shorter wavelengths, due to less tissue scattering. This loss of tissue scattering will influence the observed absorption in the tissue as well, and should be corrected for in the optical photon transport model used in the analysis of the data. An illustration of these phenomena is given in Figs. 22.18 and 22.19. The left panel in Fig. 22.18 shows a comparison between normal skin spectra measured in the volar side of the arm of a 30 year old woman and a 76 year old man. The right panel shows the spectra from both bruised and normal skin in a male heart surgery patient at the age of 76. These spectra were collected in the volar side of the elbow. The loss of reflection that can be seen for shorter wavelengths are due to age related changes in the skin and not due to pigmentation. This can be deduced from the shape of the curve for wavelengths above 550 nm, high melanin absorption would have caused a flatter curve in this region. The left panel of Fig. 22.19 shows a

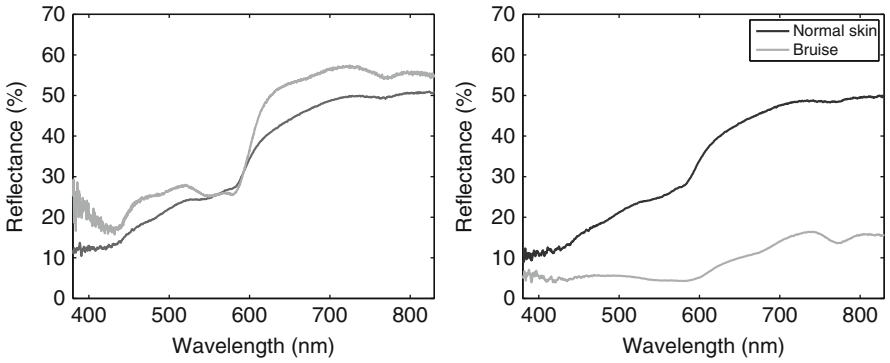


Fig. 22.18 Reflectance spectra of normal and bruised skin. *Left panel*: Normal skin of the volar side of the arm in a 30 year old woman (*top line*), and a 76 year old man (*bottom line*). The sites had not been recently exposed to the sun; *Right panel*: Bruised and normal skin in the volar side of the arm in the same 76 years old male. The bruise was 5 days old

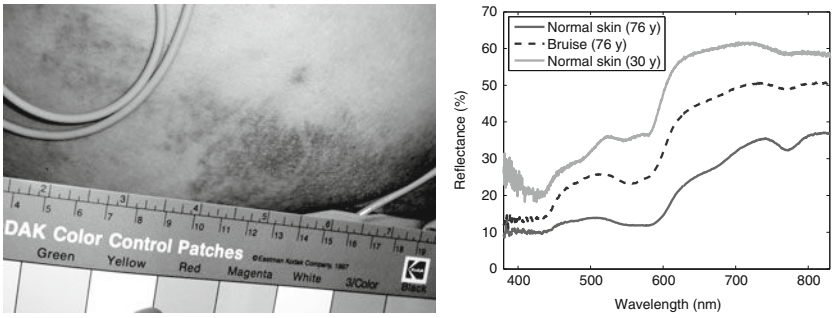


Fig. 22.19 Reflectance spectra and photo of a bruise in the abdominal region of a 76 year old male. The reflectance spectra is compared to a spectrum of the same region in a 30 year old female. The local minimum at 480 nm in the green spectrum is due to beta-carotene as a result of a vegetable rich diet

photo of a bruise in the abdominal region of a 76 year old male, and the right panel shows the reflectance spectra of the same bruise compared to normal skin and the normal skin in the same region of a 30 year old female. The scattering in this region is not as reduced as in the previous example, probably due to less sun exposure and a thicker subcutaneous fat layer. The oxygenation in the skin will differ with depth into the skin. Red light penetrates deeper into the skin, and the reflectance in the red part of the spectrum therefore reflects the oxygenation in deeper skin layers. The deoxyhemoglobin absorption peak at 760 nm is not pronounced in the normal skin measurement. This indicates a good oxygenation and perfusion in deeper skin layers. In the blue/green range of the spectrum the light penetrates shorter into the skin, and the shape of the spectrum reflects the oxygenation at more shallow skin depths. In this case the oxygenation is very poor, indicating a reduced perfusion in the upper

skin layers. The bruises in both examples are influencing the skin in the whole spectral range, indicating substantial influence of the hemorrhage in all skin layers. Both these examples show spectra without a strong contribution from bilirubin.

22.4.5 Age Determination of Bruises

Reliable and reproducible age determination of bruises is the ultimate goal of forensic medical optics. Visual inspection has been proven to be inaccurate, and color coordinates have also been tested without large success. Examination of full reflectance spectra seems to be a path to success, but requires mathematical modeling to be reliable. It is quite easy to determine if a bruise is recent or old based on the reflectance spectrum, but accurate age determination is not trivial. Reflectance spectra from old and recent injuries are presented together with scans of normal skin in Fig. 22.20. The skin reflectance from fresh bruises show features consistent with increased oxyhemoglobin absorption shown by the prominent local minima at 542 and 576 nm, and increased blood content compared to normal skin. Older bruises show more distinct deoxyhemoglobin features with local minimum at 760 nm and a flattened appearance at the spectral region around 550 nm, and a local minimum at approximately 460 nm due to bilirubin.

The mathematical methods presented in Sections 22.3.3 and 22.3.4.1 have been tested for age determination of bruises in a relatively homogeneous group of cardiac surgery patients [8]. In this study it was found that for recent injuries, i.e. less than one day old, the average error in the estimated age was 1 ± 3 h. For injuries between one and ten days old, the estimated error was 0 ± 0.5 days. For older

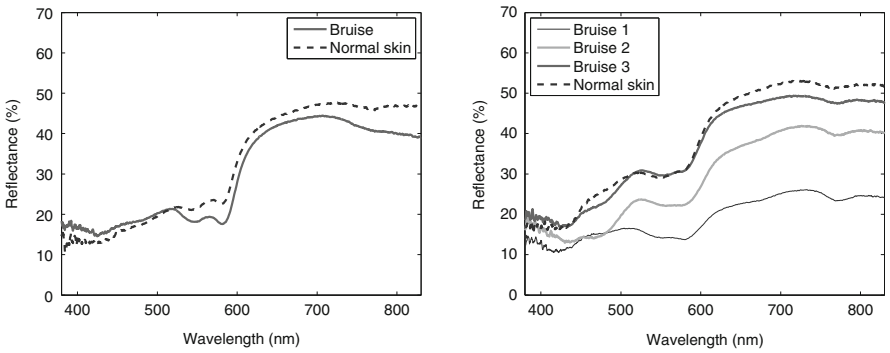


Fig. 22.20 Recent and old bruises. *Left panel*: Recent bruise. A black eye in a female (age above 70 years), measured at 6–12 h after injury. *Right panel*: Old bruise in the abdominal region of a female (age above 70 years). The subject had several bruises of different age in the abdominal region after Klexane injections prior to surgery. The true age of the injuries was unknown, but the time period was known. Bruise 1 is most recent (approximately 3–4 days), and bruise 3 is oldest (approximately 12 days)

injuries the errors were larger. The sample included very few injuries that were older than 10 days. These results are interesting, but the method is still not confirmed to be accurate for other age groups, and the model parameters will probably have to be adjusted to fit those groups. The model itself consists of a combination of the mathematical model for hemoglobin transport in skin, combined with an analytic photon transport model. A measurement of the skin reflectance in normal skin is used to derive the parameters for the normal skin (oxygenation, pigmentation, and dermal blood volume) which is fed into the model as base line information. The size of the hemorrhage is roughly estimated on the basis of the diameter of the bruise, and the oxygenation is estimated from the measured reflectance spectrum of the bruise and fed into the model. The model can then simulate the temporal development of the bruise, and the age of the injury is found by matching the measured reflectance of the bruise to the reflectance simulated by the model.

Much can still be done to improve this method, both on the modeling side and on the collection of data. Hyperspectral imaging will ease the collection of data by providing spatial resolution. The size of the bruise can easily be determined from the image, and one can also choose a good spot for collecting the reflectance spectra or one can choose to use the average spectra of larger regions of the bruise as a basis for age determination. The mathematical models can probably be made more reliable by more detailed numerical models, and a better scientific basis for estimating the input parameters. There is definitely a need for more research on the physiological and biological changes in connection with bruising, to be able to understand the process fully and to improve the age determination algorithms.

22.5 The Future of the Optics of Bruises

Age determination of bruises is a very difficult task. It can be questioned whether it is possible at all. Optical characterization and optical diagnostic techniques provide good tools to achieve a more objective evaluation of bruises in general. Specific and correct age determination can be realized, but probably not with a general model for the whole population. A combination of optical diagnostics and mathematical modeling seems to be the most promising path to accurate dating of bruises. So far, the knowledge of the biological processes involved in bruising is limited. When the knowledge of these processes increases, the quality of the models will also improve. Age, gender and ethnicity may require specific adaptations of the mathematical models. If such modular models are constructed and used with care, the individual legal protection in connection with violence and abuse cases might be increased due to the objectivity of such models.

So far, most of the work within this field has been done on Caucasian skin, and further studies are required to adapt the methods to more pigmented skin types. Bilirubin is difficult to detect in dark skin, and it will be of great importance to identify markers that absorbs in either the NIR or the IR part of the spectrum to achieve methods that will work for all skin types.

The future of forensic medical optics will depend on basic research within both optics and forensic sciences and the combination of these areas. Imaging techniques with increased resolution will increase the spatial resolution and probably also contribute to new achievements within this field.

Acknowledgements Rolf Tore Randeberg, Rune Haaverstad, Jørgen Bru, Andreas Winnem, Henrik Gundersen, Bjørn Fossan Rasmussen, Sandra Blindheim.

References

1. Langlois NE and Gresham GA. The ageing of bruises: a review and study of the colour changes with time. *Forensic Sci. Int.*, 50(2):227–238 (1991).
2. Schwart AJ and Ricci LR. How accurately can bruises be aged in abused children? Literature review and synthesis. *Pediatrics*, 97(2):254–257 (1996).
3. Stephenson T and Bialas Y. Estimation of the age of bruising. *Arch. Dis. Child.*, 74(1):53–55 (1996).
4. Altemeier WA 3rd. A pediatrician's view. Interpreting bruises in children. *Pediatr. Ann.*, 30(9):517–518, 520 (2001).
5. Maguire S, Mann MK, Sibert J, and Kemp A. Can you age bruises accurately in children? A systematic review. *Arch. Dis. Child.*, 90(2):187–189 (2005).
6. Hughes VK, Ellis PS, and Langlois NE. The perception of yellow in bruises. *J. Clin. Forensic Med.*, 11(5):257–259 (2004).
7. Munang LA, Leonard PA, and Mok JYQ. Lack of agreement on colour description between clinicians examining childhood bruising. *J. Clin. Forensic Med.*, 9(4):171–174 (2002).
8. Randeberg LL, Haugen OA, Haaverstad R, and Svaasand LO. A novel approach to age determination of traumatic injuries by reflectance spectroscopy. *Lasers Surg. Med.*, 38(4):277–289 (2006).
9. Sandby-Moller J, Poulsen T, and Wulf HC. Epidermal thickness at different body sites: relationship to age, gender, pigmentation, blood content, skin type and smoking habits. *Acta Derm. Venereol.*, 83(6):410–413 (2003).
10. Capper C. The language of forensic medicine: The meaning of some terms employed. *Med. Sci. Law*, 41(3):256–259 (2001).
11. Vanezis P. Interpreting bruises at necropsy. *J. Clin. Pathol.*, 54(5):348–355 (2001).
12. Bohnert M, Baumgartner R, and Pollak S. Spectrophotometric evaluation of the colour of intra- and subcutaneous bruises. *Int. J. Legal Med.*, 113(6):343–348 (2000).
13. Svaasand LO, Norvang LT, Fiskerstrand EJ, Stopps EKS, Berns MW, and Nelson JS. Tissue parameters determining the visual appearance of normal skin and port-wine stains. *Lasers Med. Sci.*, 10(1):55–65 (1995).
14. Bauer D, Grebe R, and Ehrlacher A. A three-layer continuous model of porous media to describe the first phase of skin irritation. *J. Theor. Biol.*, 232(3):347–362 (2005).
15. Bauer D, Grebe R, and Ehrlacher A. First phase microcirculatory reaction to mechanical skin irritation: a three layer model of a compliant vascular tree. *J. Theor. Biol.*, 232(2):249–260 (2005).
16. Szczesny, G, Viehelmann, A, Nolte D, and Messmer K. Changes in the local blood and lymph microcirculation in response to direct mechanical trauma applied to leg: In vivo study in an animal model. *J. Trauma*, 51(3):508–517 (2001).
17. Gemsa D and Schmid R. Haemoglobin degradation and bilirubin formation. *Klin Wochenschr,* 52(13):609–616 (1974).
18. Pimstone NR, Tenhunen R, Seitz PT, Marver HS, and Schmid R. The enzymatic degradation of hemoglobin to bile pigments by macrophages. *J. Exp. Med.*, 133(6):1264–1281 (1971).
19. Hughes VK, Ellis PS, Burt T, and Langlois NEI. The practical application of reflectance spectrophotometry for the demonstration of haemoglobin and its degradation in bruises. *J. Clin. Pathol.*, 57(4):355–359 (2004).

20. Maines MD and Cohn J. Bile pigment formation by skin heme oxygenase: studies on the response of the enzyme to heme compounds and tissue injury. *J. Exp. Med*., 145(4):1054–1059 (1977).
21. Custer G, Balcerzak S, and Rinehart J. Human macrophage hemoglobin-iron metabolism in vitro. *Am. J. Hematol*., 13(1):23–36 (1982).
22. Thornton RN and Jolly RD. The objective interpretation of histopathological data: an application to the ageing of ovine bruises. *Forensic Sci. Int.*, 31(4):225–239 (1986).
23. Randeberg LL, Winnem AM, Blindheim S, Haugen OA, and Svaasand LO. Optical classification of bruises. *Proc. SPIE*, 5312:53–64 (2004).
24. Betz P. Histological and enzyme histochemical parameters for the age estimation of human skin wounds. *Int. J. Legal Med*., 107(2):60–68 (1994).
25. Betz P and Eisenmenger W. Morphometrical analysis of hemosiderin deposits in relation to wound age. *Int. J. Legal Med*., 108(5):262–264 (1996).
26. Langlois NEI. The science behind the quest to determine the age of bruises—a review of the English language literature. *Forensic Sci. Med. Pathol*., 3:241–251 (2007).
27. Zijlstra WG, Buursma A, and Van Assendelft OW. *Visible and near infrared absorption spectra of human and animal haemoglobin*. VSP Books, Utrecht (2000).
28. Du H, Fuh RA, Li J, Corkan A, and Lindsey JS. PhotochemCAD^{++}: A computer-aided design and research tool in photochemistry. *Photochem. Photobiol*., 68(2):141–142 (1998).
29. Randeberg LL. *Diagnostic applications of diffuse reflectance spectroscopy*. Doctoral theses at NTNU 2005:100. Norwegian University of Science and Technology, Faculty of Information Technology, Mathematics and Electrical Engineering, Trondheim (2005).
30. Randeberg LL, Winnem AM, Langlois NE, Larsen ELP, Haaverstad R, Skallerud B, Haugen OA, and Svaasand LO. Skin changes following minor trauma. *Lasers Surg. Med.*, 39(5):403–413 (2007).
31. Virchow, R. Die pathologischen Pigmente. *Virchows Arch. Pathol. Anat.*, 1:379–486 (1847).
32. Wilson EF. Estimations of the age of cutaneous contusions in child abuse. *Pediatrics*, 60(5):750–752 (1977).
33. Trujillo O, Vanezis P, and Cermignani M. Photometric assessment of skin colour and lightness using a tristimulus colorimeter: Reliability of inter and intra-investigator observations in healthy adult volunteers. *Forensic Sci. Int.*, 81(1):1–10 (1996).
34. Yajima Y and Funayama M. Spectrophotometric and tristimulus analysis of the colors of subcutaneous bleeding in living persons. *Forensic Sci. Int*., 156(2–3):131–137 (2006).
35. Randeberg LL, Larsen ELP, and Svaasand LO. Optical classification of bruises. In *Progress in biomedical optics and imaging – proceedings of SPIE*. San Jose, CA (2004).
36. Svaasand LO, Fiskerstrand EJ, Kopstad G, Norvang LT, Svaasand EK, Nelson JA, and Berns MW. Therapeutic response during pulsed laser treatment of port-wine stains: Dependence on vessel diameter and depth in dermis. *Lasers Med. Sci*., 10(4):235–243 (1995).
37. Cremer RJ, Perryman PW, and Richards DH, Influence of light on the hyperbilirubinemia of infants. *Lancet*, 1:1094–1097 (1958).
38. McDonagh AF and Lightner DA. 'Like a shrivelled blood orange' – bilirubin, jaundice, and phototherapy. *Pediatrics*, 75(3):443–455 (1985).
39. Kienle A, Lilge L, Vitkin IA, Patterson MS, Wilson BC, Hibst R, Steiner R. Why do veins appear blue? A new look at an old question. *App. Opt.* 35:1151–1160 (1996).
40. *Colorimetry*. 2nd ed. Publication CIE 15.2., Vienna. 74 s (1986).
41. Bohnert M, Weinmann W, and S. Pollak S. Spectrophotometric evaluation of postmortem lividity. *Forensic Sci. Int*., 99(2):149–158 (1999).
42. Yajima Y, Nata M, and Funayama M. Spectrophotometric and tristimulus analysis of the colors of subcutaneous bleeding in living persons. *Leg. Med.* (Tokyo), 5(Suppl 1): S342–S343 (2003).
43. Klein A, et al. Determining times of haematomas in living humans by means of spectroreflectometric measurements (in German). *Beitr-Gerichtl-Med.*, 50:235–240 (1992).
44. Stamatas GN and Kollias N. Blood stasis contributions to the perception of skin pigmentation. *J. Biomed. Opt*., 9(2):315–322 (2004).

45. Wu JZ, Dong RG, Smutz WP, and Schopper AW. Nonlinear and viscoelastic characteristics of skin under compression: experiment and analysis. *Biomed. Mater. Eng*., 13(4):373–385 (2003).
46. Mow VC, Kuei SC, Lai WM, and Armstrong CG. Biphasic creep and stress relaxation of articular cartilage in compression? Theory and experiments. *J. Biomech. Eng*., 102(1):73–84 (1980).
47. Oomens CW, van Campen DH, and Grootenboer HG. A mixture approach to the mechanics of skin. *J. Biomech*., 20(9):877–885 (1987).
48. Winnem AM, Randeberg LL, Larsen ELP, Lilledahl MB, Haaverstad R, Haugen OA, Skallerud B, and Svaasand LO. Biomechanical characterization of soft tissue injuries. In: *Progress in biomedical optics and imaging, Proceedings of SPIE*. San Jose, CA (2007).
49. Zakaria EL, Lofthouse J, and Flessner MF. In vivo hydraulic conductivity of muscle: Effects of hydrostatic pressure. *Am. J. Physiol. Heart. Circ. Physiol*., 273:642–646 (1997).
50. Stute, M. Darcy's Law, cited; Available from: http://www.ldeo.columbia.edu/~martins/hydro/lectures/darcy.html.
51. Jürgens KD, Peters T, and Gros G. Diffusivity of myoglobin in intact skeletal muscle cells. *Proc. Natl. Acad. Sci. U.S.A*., 91(9):3829–3833 (1994).
52. Mosqued, Burnight K, and Liao S. The life cycle of bruises in older adults. *J. Am. Geriatr. Soc*., 53(8):1339–1343 (2005).
53. Randeberg LL, Roll EB, Norvang Nilsen LT, Christensen T, and Svaasand LO. In vivo spectroscopy of jaundiced newborn skin reveals more than a bilirubin index. *Acta Paediatr.*, 94(1):65–71 (2005).
54. Spott T and Svaasand LO, Collimated light sources in the diffusion approximation. *Appl. Opt*., 39(34):6453–6465 (2000).
55. Haskell RC, Svaasand LO, Tsay T-T, Feng T-C, McAdams MS, and Tromberg BJ. Boundary conditions for the diffusion equation in radiative transfer. *J. Opt. Soc. Am. A*, 11(10):2727–2741 (1994).
56. Randeberg LL, Bonesrønning JH, Dalaker M, Nelson JS, and Svaasand LO. Methemoglobin formation during laser induced photothermolysis of vascular skin lesions. *Lasers Surg. Med*., 34(5):414–419 (2004).
57. Saidi IS, Jacques SL, and Tittel FK. Mie and Rayleigh modeling of visible-light scattering in neonatal skin. *Appl. Opt*., 34(31):7410–7418 (1995).
58. Payne G, Langlois N, Lennard C, and Roux C. Applying visible hyperspectral (chemical) imaging to estimate the age of bruises. *Med. Sci. Law*, 47(3):225–232 (2007).
59. Randeberg LL, Baarstad I, Loke T, Kaspersen P, and Svaasand LO. Hyperspectral imaging of bruised skin. In: *Progress in biomedical optics and imaging, Proceedings of SPIE*. San Jose, CA (2006).
60. Gundersen HM and Rasmussen BF. *An application of image processing techniques for enhancement and segmentation of bruises in hyperspectral images*. Doctoral theses at NTNU. Norwegian University of Science and Technology, Department of Computer and Information Science, Trondheim (2007).
61. Bohner M, Schulz K, Belenkaia L, and Liehr AW, Re-oxygenation of haemoglobin in livores after post-mortem exposure to a cold environment. *Int. J. Legal Med.*, 122(2):91–96 (2008).
62. Burke MP, Olumbe AK, and Opeskin K., Postmortem extravasation of blood potentially simulating antemortem bruising. *Am. J. Forensic Med. Pathol.*, 19(1):46–49 (1998).
63. Cabinum-Foeller E and Frasier L. Bruising in children. *Lancet,* 365(9468):16–22 (2005).
64. Maguire S, Mann MK, Sibert J, and Kemp A. Are there patterns of bruising in childhood which are diagnostic or suggestive of abuse? A systematic review. *Arch. Dis. Child.*, 90(2):182–186 (2005).
65. Krug E (ed) Abuse of the elderly. In *World report on violence and health.* World Health Organization, Albany, NY (2002).
66. Webster GF. Common skin disorder in the elderly. *Clin. Cornerstone*, 4(1):39–42 (2001).

Chapter 23
Laser Treatment of Port Wine Stains

Boris Majaron and J. Stuart Nelson

23.1 Introduction

23.1.1 Port Wine Stain (PWS)

Port wine stain (PWS), also called nevus flammeus, is a congenital, cutaneous vascular malformation involving post-capillary venules which produce a light pink to red to dark-red-violet discoloration of human skin [1]. PWS occurs in an estimated 3 children per 1000 live births, affecting males and females and all racial groups equally [2]. There appears to be no hereditary predilection for PWS within families. There are no known risk factors or ways to prevent PWS.

Since most of the malformations occur on the face, PWS is a clinically significant problem in the great majority of patients. PWS should not be considered a cosmetic problem but a disease with potentially devastating psychological and physical complications. Detailed studies have documented lower self-esteem and problems with interpersonal relations [3, 4] in PWS patients.

Histopathological studies of PWS show a normal epidermis overlying an abnormal plexus of dilated blood vessels located in the dermis (Fig. 23.1) [5]. The cause and origin of PWS remains incompletely understood. It is believed that PWS develop within the first 2–8 weeks of gestation. The most likely hypothesis for the development of PWS is the deficiency or absence of surrounding neurons regulating blood flow through the ectatic post-capillary venules [6]. As a result, the blood vessels are unable to constrict normally and remain permanently dilated.

PWS is a progressive vascular malformation of the skin. PWS remain throughout life; there is no involution. PWS are well demarcated and flat and grow proportionately in surface area with the child. In infants and young children, PWS are flat red macules. However, the lesions tend to darken progressively to purple and, by adult

B. Majaron (✉)
Jožef Stefan Institute, Ljubljana, Slovenia
e-mail: boris.majaron@ijs.si

A.J. Welch, M.J.C. van Gemert (eds.), *Optical-Thermal Response of Laser-Irradiated Tissue*, 2nd ed., DOI 10.1007/978-90-481-8831-4_23,

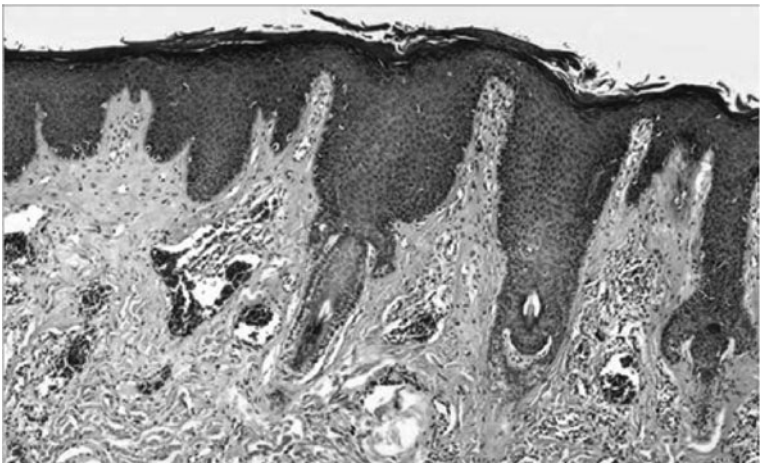


Fig. 23.1 Histopathology of PWS. Note multiple dilated capillaries in the upper dermis

age, often become raised as a result of the development of vascular papules or nodules. These changes in color and contour are attributed to progressive ectasia of the abnormal dermal vascular plexus. If left untreated, PWS often become incompatible with normal life due to the development of vascular nodules on the skin surface which can often bleed spontaneously with incidental trauma. The hypertrophy (increased tissue mass) of the underlying soft tissue that occurs in approximately two-thirds of lesions further disfigures the facial features of many patients. For all of the above reasons, most medical specialists agree that it is essential to begin treatment of PWS as early as possible and to maintain treatment in order to prevent the development of vascular nodules and hypertrophy in later years.

PWS are part of a constellation of vascular malformations and may be a part of a larger disorder or syndrome. Thus, it is important that physicians evaluating PWS patients recognize that associated medical problems may be present, requiring medical intervention separate from treatment of PWS. The most important examples include Sturge-Weber (encephalotrigeminal angiomatosis), Klippel-Trenaunay, Cobb, and Proteus syndromes.

23.1.2 PWS Therapy

PWS have been treated in the past with an array of therapeutic modalities, including skin grafting, ionizing radiation, dermabrasion, cryosurgery, tattooing and electrotherapy [1]. All of these have met with limited success and often left cosmetically unacceptable secondary scarring. These are no longer considered viable treatment options. Once the patient's PWS has developed vascular papules or nodules, masking the lesion with corrective make-up becomes ineffective.

The introduction of the argon laser in the early 1970s represented the first major advance in PWS [7, 8]. The blue-green light (488 and 514 nm) produced by the argon laser is absorbed selectively by hemoglobin within the PWS blood vessels. There, the radiant energy is converted to heat, causing thrombosis and irreversible thermal damage to the targeted PWS blood vessels. Unfortunately, the epidermis also suffers some thermal injury due to undesired absorption by melanin. For many lesions, the threshold for permanent epidermal damage following argon laser therapy was very close to the threshold for blanching of the PWS. Furthermore, in the first generation argon lasers used for medical intervention, the shortest applicable pulse duration was 0.01 s, thereby contributing to non-specific thermal damage. Although the treatment of PWS with the argon laser could produce favorable results, scarring remained a worrisome complication, particularly in infants and young children, due to the propensity for scar formation in younger age groups. Similar results were also reported with other continuous and quasicontinuous lasers including the copper vapor, krypton-ion, argon-pumped dye, Nd:YAG and carbon dioxide lasers. These lasers are therefore no longer recommended for PWS therapy.

If the clinical objective is to cause selective destruction of a specific tissue structure, the laser wavelength should match the highest absorption of the targeted structure relative to surrounding tissue. However, for successful treatment the pulse duration is just as important as optical and tissue factors. During a lengthy laser exposure, most of deposited heat may diffuse away from the target structure, resulting in nonspecific thermal damage to adjacent structures. Conversely, a suitably short laser pulse confines the heating effect to the target structure, resulting in maximum temperature difference between the target and adjacent structures. This paradigm, termed *selective photothermolysis*, was introduced by Anderson and Parrish as a means of achieving selective destruction of subsurface targets by careful selection of wavelength and pulse duration [9]. This was the first of many theoretical constructs around which medical lasers would subsequently be developed.

Selective photothermolysis of PWS blood vessels is currently achieved using pulsed green, yellow or orange lasers. The immediate post-treatment edema and erythematous flare resolve within 48 h, while the temporary purple discoloration (“purpura”) generally resolves within 2 weeks. Thereafter, remnants of coagulated blood and vascular wall components are assimilated within the skin and replaced primarily by dermal collagen. The reduction in the dermal blood volume fraction is observed clinically as PWS lesion blanching.

Pulsed laser treatments are currently administered by moving a laser handpiece, which creates a 3–12 mm diameter spot on the skin surface, across the entire PWS in a methodical fashion. Because PWS blanching is almost never achieved after just one treatment, additional sessions are required, with a 4–8 week interval between successive patient visits. However, if the ultimate standard required is complete blanching of the lesion, the average success rate is below 20% [10, 11].

Skin dyspigmentation (hyper- or hypo-pigmentation) often occurs after pulsed laser therapy in PWS patients with darker skin phototypes. However, dyspigmentation is usually temporary and resolves spontaneously over 6–12 months. Epidermal necrosis and subsequent skin surface textural changes such as scarring, atrophy

or induration were also observed after treatments with high radiant exposures. However, the benefits of laser therapy far outweigh the risks of no treatment and should be considered a medical necessity [12]. The majority of reported scars have occurred in areas accidentally traumatized soon after laser exposure. Moreover, application of active skin cooling, most notably cryogen spray cooling, largely prevents adverse side effects (see Section 23.2.5).

PWS response remains variable and unpredictable, with many patients' PWS fading only minimally. Presently, all patients are treated using laser parameters with selection based on clinical judgment of the physician. However, epidermal thickness (50–150 μm) and melanin concentration in human skin, as well as PWS blood vessel diameter (30–300 μm) and depth distribution (100–1000 μm) vary on an individual patient basis and even from site to site on the same patient [5, 13].

Multiple devices are now available for PWS treatment, each with its own unique wavelength and pulse duration. Changing the wavelength or pulse duration of the laser can result in substantial PWS fading not previously observed with single device therapy. Moreover, several devices are sometimes used during an extended treatment protocol in order to destroy vessels of different sizes.

It has become clear that treatment of a PWS in its macular stage will prevent the development of the hypertrophic component of the lesion. Postoperative biopsies after laser treatment of a PWS reveal that the existing blood vessels are smaller and fewer in number compared with pretreatment biopsies [14]. Thus, the opportunity for progression of these lesions to a more ectatic state is less likely to occur. Although the majority of PWS lesions do not recur, some lesional redarkening many years later has been reported after successful pulsed dye laser (PDL) therapy [15]. One possible explanation might be continuous dilatation of the remaining ectatic vessels which also lack autonomic innervation. Patients who do experience some redarkening will usually only require one or two treatments to return to their former level of PWS blanching.

Studies have recently shown that aggressive treatment of infants and young children at earlier ages improves PWS clearance [16]. There are several "optical" advantages to treating patients at as young an age as possible: (1) less epidermal melanin; which competes for the absorption of laser light; (2) less collagen in the skin, which results in less light being back-scattered out of the skin; and (3) thinner dermis and lower fractional blood volume in younger patients, which allows more light to penetrate into the skin to destroy targeted PWS blood vessels.

It has also been documented that there can be anatomical variation in terms of the PWS response to laser therapy [17]. For example, central face (particularly those lesions involving the V_2 dermatome) PWS respond less effectively to treatment as compared to other areas of the lateral face and neck. PWS involving areas over bony prominences respond well to laser therapy. Upper body PWS respond better to laser therapy than those lesions on the lower body and extremities. PWS lesions associated with Sturge-Weber or Klippel-Trenaunay syndrome respond poorly to laser therapy. Thus, any study evaluating the effectiveness of laser treatment needs to account for PWS anatomical variation of the response to treatment.

23.1.3 Biological and Physical Endpoints of PWS Laser Therapy

A number of different biophysical, biochemical, and biological processes can be induced by irradiation of cutaneous blood vessels with short laser pulses. Depending on radiant exposure, laser wavelength, pulse duration, vessel diameter, etc., the following transient or permanent effects may occur: blood coagulation (with coagulum either embolized, adherent to the vessel wall, or fully obstructing blood flow), vasodilatation, vasoconstriction, vessel occlusion, intravascular cavitation, formation of persistent bubbles, vessel wall rupture, and shrinkage of perivascular collagen.

Clinically, vessel wall rupture in response to explosive vaporization of blood correlates with poor therapeutic outcome in terms of PWS blanching [18, 19]. In such cases, tissue repair mechanisms resolve the dark skin discoloration (purpura) within 7–14 days and revascularize the PWS lesion with minimal blanching. Animal model studies have confirmed that the pathway to irreversible vessel occlusion does not involve blood vaporization [20, 21].

Instead, complete blood vessel wall coagulation is required to bring about their assimilation within the skin and replacement by dermal collagen instead of inducing endothelial cell migration and regeneration of PWS blood vessels [22]. It is not entirely clear, however, whether complete occlusion of blood flow by red blood cell (RBC) coagulum, which is regularly observed in post-treatment histology, is also a prerequisite for a successful therapeutic outcome. In one histological study, clinical outcome correlated well with PWS vessel wall coagulation depth, but not with the (greater) depth of intraluminal blood coagulation [23]. Apparently, the thresholds for denaturation of RBC and circulating proteins are lower as compared to those for tissue proteins that lead to vessel wall coagulation [24, 25]. Intraluminal blood coagulation may thus be an unavoidable intermediate step, because the laser energy is deposited first into the blood vessel due to selective absorption by hemoglobin, before the released heat is transported to the less absorbing vessel wall.

Relations between laser treatment parameters and their influence on PWS therapeutic outcome are quite complex and remain incompletely understood. The sequence of physical processes triggered by laser irradiation is mainly governed by both optical transport and heat transfer dynamics inside the highly heterogeneous tissue structure. Moreover, the photo- and thermally-induced biophysical and biochemical processes (e.g., morphological transitions and membrane fragmentation in RBC, platelet aggregation, formation of met-hemoglobin), with the associated dynamic changes of optical and thermal properties of tissue, further complicate these relations.

The clinical objective of PWS laser treatment is to induce irreversible thermal injury to the ectatic blood vessels, which will stimulate a suitable wound healing response, while avoiding nonselective damage to the healthy epidermis and dermis. For the purposes of mathematical analysis and modeling, the complex laser-tissue interaction can be divided into three main processes with corresponding endpoints:

- absorption of laser light by hemoglobin, which induces selective (and possibly uniform) heating of the vascular lumen relative to surrounding tissue, usually evidenced by complete intraluminal blood coagulation;
- heat transfer from the absorption sites into the entire vessel wall, which enables a sufficient temperature rise in the latter; and
- coagulation of blood vessel wall (or perivascular collagen in capillaries lacking a true vessel wall), which is enabled by maintaining the elevated temperature for a suitable time interval for this dynamic process to occur.

Detailed numerical simulations of optical and thermal transport for increasingly complex PWS model geometries have been successfully developed using the above endpoints as criteria for treatment success. These have been instrumental in advancing our understanding of the interaction between the optical and thermal effects involved in PWS therapy and their relation to the many laser treatment parameters. However, in order to predict reliably the therapeutic effect of a particular laser irradiation protocol, the main limitations are:

- the complexity and large inter- and intra-patient variability of optical and structural characteristics between PWS lesions;
- our limited ability to develop microscopic dynamic models of the clinically most relevant processes, such as aggregation of RBC, vaporization of whole blood, and vessel hemorrhage;
- incomplete knowledge regarding the dynamic changes of optical and thermal properties of individual tissue components during laser irradiation; and
- incomplete understanding of photo-thermally triggered biochemical reactions and biological mechanisms involved in the healing process.

23.2 Optical and Thermal Considerations in PWS Laser Therapy

Laser systems available today for PWS treatment cover a wide range of parameters: wavelengths from 532 to 600 nm, pulse durations from 0.45 to 50 ms, and beam diameters from 3 to 12 mm. Selecting an appropriate PWS treatment protocol has become increasingly more challenging by introduction of devices that allow the operator to vary not only the beam diameter and radiant exposure, but also wavelength and pulse duration. Introduction of active cooling has offered additional flexibility in the treatment protocol. To date, however, no treatment protocol has demonstrated a significant supremacy in terms of optimal therapeutic outcome for all PWS patients.

As mentioned above (Section 23.1.1), there is considerable variation in PWS lesion geometry between individual patients, and even between different areas on the same patient. In this section, we discuss the relationships between individual laser treatment parameters and key PWS lesion characteristics, such as blood vessel

diameter and depth distributions, epidermal thickness, and melanin content. Note that two different aspects of such relationships can be considered: First, which PWS lesions will respond best to a given set of laser treatment parameters. Second, which parameter settings would be optimal for a PWS lesion with known structural and optical characteristics.

Herein (Section 23.2), we limit the discussion to the common treatment approach of today, whereby the local therapeutic effect is achieved using a single monochromatic laser pulse, often applied in conjunction with active cooling of the skin surface. Some recently introduced or proposed extensions of this treatment paradigm and auxiliary techniques to help improve laser therapy of PWS are addressed in Sections 23.3 and 23.4, respectively.

23.2.1 Beam Diameter (Spot Size)

When a confined laser beam irradiates human skin, scattering in radial direction [26, 27] contributes to reduction of beam intensity with subsurface depth, z. Consequently, selection of laser beam diameter (spot size) affects both lateral and axial distribution of fluence rate $\phi\,(r, z)$ [W/cm^2] within the skin (see Chapter 3, Fig. 3.15). In general, using a larger beam diameter – while maintaining the same radiant exposure H [J/cm^2] – enhances delivery of laser energy to deeper absorbing targets.

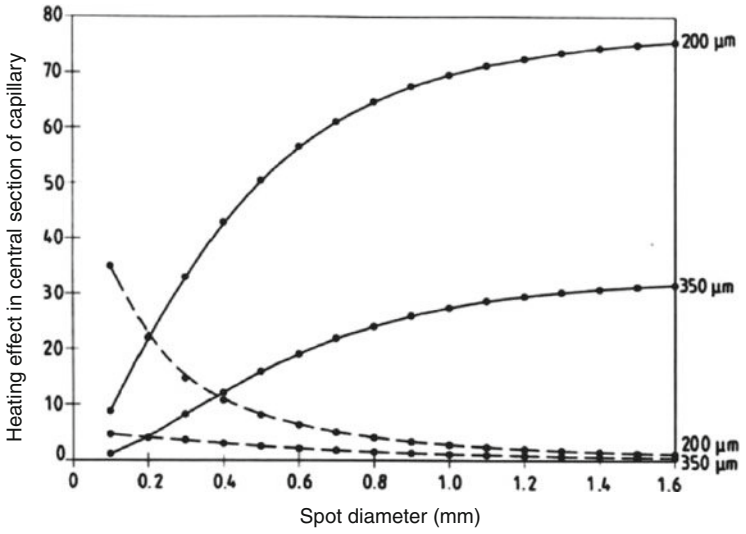


Fig. 23.2 Heating effect on the axis of central capillaries at depths of 200 and 350 µm, as a function of laser beam diameter, D. For constant radiant exposure H [J/cm^2] greater heating is obtained with increasing D (solid lines). For constant energy of the pulse [J] the heating effect always decreases with D (dashed lines). Results from numerical simulation for $\lambda = 577$ nm [30]

Early numerical work indicated that the laser beam diameter for efficient PWS treatment at $\lambda = 577$ nm should be at least $D = 1$ mm or, even better, 3 mm [28]. A small further increase of subsurface fluence rate on the axis of the incident beam, $\phi(0, z)$, was predicted for beam diameters up to 5 mm, in agreement with experimental observations [29].

The effect is illustrated in Fig. 23.2 which presents results of numerical simulations utilizing a 3D Monte Carlo (MC) model of optical transport in skin with a regular array of 50 μm capillaries [30]. When radiant exposure is kept constant, the capillary heating increases sharply with increasing beam diameter until it begins to level off at $D > 1$ mm (solid lines). Note, however, that this involves an increase of the laser pulse energy, proportional to D^2. Increasing D at a constant pulse energy will always reduce the capillary heating (dashed lines).

Note that deeper penetrating wavelengths require a larger beam diameter to prevent an excessive decrease of fluence with depth due to dermal light scattering. In recent experiments on PWS test sites using irradiation at $\lambda = 595$ nm (and $H = 14$ J/cm^2), increasing the beam diameter from 5 to 7 mm enhanced the depth to which blood vessel were successfully coagulated [31]. In addition, larger beam diameters induce more uniform irradiation of blood vessels from all sides, which is particularly important for larger target vessels.

23.2.2 Wavelength

Despite considerable past research, determination of the optimal wavelength for PWS laser treatment remains a controversial issue. Herein, we discuss the basic optical effects that influence laser wavelength selection for PWS therapy, using results from representative mathematical models and experiments in physical models in vitro and in vivo.

23.2.2.1 Absorption Selectivity

The concept of selective photothermolysis, introduced by Anderson and Parrish in 1983, described the induction of thermal injury to specific subsurface targets during irradiation of the skin surface. The first postulate of selective photothermolysis is that the wavelength of light should match high absorption peaks of the targeted structure, relative to the surrounding tissue. For selective targeting of PWS blood vessels, oxyhemoglobin (HbO_2) and deoxyhemoglobin (Hb) serve as suitable chromophore molecules, exhibiting several absorption peaks in the visible and near-IR parts of the spectrum, where the absorption of water – the main constituent of all soft tissues – is lowest (Fig. 23.3).

The absorption spectra of HbO_2 and Hb differ somewhat; therefore, the absorption coefficient of blood, $\mu_{a,\mathrm{bl}}(\lambda)$, depends also on the oxygen saturation level ($SatO_2$). Because PWS vessels are post-capillary venules, we consider optical properties of blood with $SatO_2 = 70\%$ (Table 23.1). By utilizing wavelengths with higher

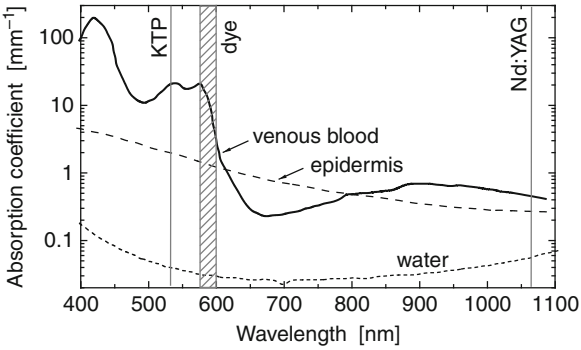


Fig. 23.3 Absorption spectra of venous blood (oxygenation level $SatO_2 = 70\%$), epidermis (light skin) and water. Vertical lines indicate wavelengths of the relevant laser sources

Table 23.1 Absorption coefficient (μ_a) and scattering coefficient (μ_s) of whole blood as a function of wavelength for oxygen saturation levels of 100 and 70% (hematocrit 0.40) [35–39]

	Absorption coefficient (mm^{-1})		Scattering coefficient (mm^{-1})	
Wavelength (nm)	$SatO_2 = 100\%$	$SatO_2 = 70\%$	$SatO_2 = 100\%$	$SatO_2 = 70\%$
532	26.6	23.1	30.6	30.1
577	35.4	31.4	28.8	28.5
585	19.1	18.6	30.4	29.7
590	6.9	9.1	31.3	30.3
595	4.3	6.0	31.4	30.4
600	2.5	4.0	31.3	30.4
1064	0.5	0.22	13.5	12.9

absorption in Hb than in HbO_2 (e.g., 590–600 nm), PWS venules could be targeted selectively as compared to healthy arterioles [32].

Because light passes through the epidermis, the fluence rate at the subsurface target locations will be affected also by absorption by epidermal melanin. Moreover, melanin absorption causes localized heating of the epidermis, which may cause transient or permanent complications such as skin dyspigmentation and scarring. In the visible part of the spectrum, melanin absorption decreases rapidly with wavelength, approximately as $\lambda^{-3.48}$ [33], thereby favoring use of longer wavelengths.

The hemoglobin Soret absorption band (at $\lambda = 418$ nm) offers the highest blood absorption in the visible spectrum and excellent selectivity as compared to melanin. However, the very high $\mu_{a,\mathrm{bl}}$ value induces a pronounced parasitic absorption in dermal capillaries, which prevents effective treatment of subsurface PWS vessels (see Section 23.2.2.3). In addition, penetration of such light into human skin is limited by strong optical scattering. The scattering coefficients of epidermis and dermis decrease with λ throughout the visible and near-IR parts of the spectrum [34], again favoring use of longer wavelengths.

Frequency-doubled Nd:YAG lasers (a.k.a. KTP = potassium titanyl phosphate) emits green light ($\lambda = 532$ nm), which is near the HbO_2 absorption peak at 540 nm. These lasers are widely used for PWS and other vascular treatments.

Pulsed dye lasers (PDL) were developed specifically to explore the HbO_2 absorption peak at 577 nm. PDL technology allows selection of the emitted wavelength for PWS treatment over a range of 577–600 nm (yellow through orange). Because melanin absorption and dermal scattering are lower than at 532 nm, PDLs today represent the most popular and effective approach to PWS laser treatment.

The Nd:YAG laser ($\lambda = 1064$ nm) offers deep penetration into the dermis and is commonly used to treat deep, thick vascular lesions such as cavernous hemangiomas. However, the low spectral selectivity with respect to epidermis and dermis makes the Nd:YAG laser a poor choice for PWS treatment. Nevertheless, perhaps also due to its wide availability, the Nd:YAG laser is sometimes used as a last resort, or as a part of experimental protocols (e.g., Section 23.3.3).

23.2.2.2 Optical Screening Within Individual Blood Vessels

Utilizing laser wavelengths near the HbO_2 absorption peaks at $\lambda = 540$ or 577 nm has one important disadvantage. The associated absorption coefficient $\mu_{a,\mathrm{bl}}$ is so high that larger PWS blood vessels cannot be irradiated uniformly across their entire lumen. A histology study by Tan et al. [40] showed aggregates of RBC at the superficial portion of the vascular lumen after treatment with 577 nm PDL. The authors reported that larger blood vessels recovered from such localized damage. Because increasing the radiant exposure induces unwanted blood vaporization in the superficial vessel lumen, this effect – named optical screening – contributed to the poor efficacy of PWS treatment with early generation PDLs, designed by following literally the concept of selective photothermolysis ($\lambda = 577$ nm).

Because of high scattering anisotropy in blood in this spectral range ($g = 0.995$ for $\lambda = 532$–600 nm) [41], the reduced scattering coefficient $\mu_s' = \mu_s(1-g)$ is much smaller than $\mu_{a,\mathrm{bl}}$ (see Table 23.1). For example, optical penetration depth in venous blood at $\lambda = 577$ nm can thus be estimated as $\delta' = \left(\mu_a + \mu_s'\right)^{-1} \approx 1/\mu_a = 32\ \mu\mathrm{m}$, which indicates that PWS blood vessels with diameters larger than ~30 μm may not be heated uniformly by 577 nm irradiation. This effect was observed directly in a hamster skin flap window model [20]. At energy exposures where vessels with $d = 10{-}30\ \mu\mathrm{m}$ were coagulated uniformly across the entire lumen, larger vessels ($d = 30{-}60\ \mu\mathrm{m}$) exhibited RBC coagula affixed to the superficial vessel wall with blood flowing freely at the contralateral luminal region. A 69% larger H was required to induce complete occlusion of flow in 50% of the vessels in the latter group, compared to the first. In even larger vessels ($d = 60{-}120\ \mu\mathrm{m}$), complete occlusion was not achieved because a further increase of H led to vessel wall rupture.

Figure 23.4 illustrates implications of optical screening as obtained from numerical modeling of optical transport in skin [41]. For a vessel diameter of 60 μm

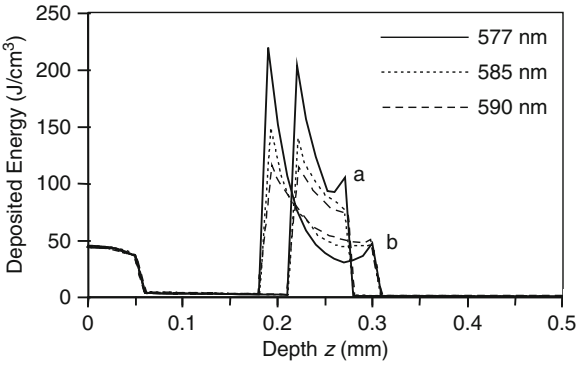


Fig. 23.4 Depth profile of deposited energy density along the laser beam axis ($D = 1$ mm), through the center of a single horizontal vessel with diameter $d = 60$ μm (**a**) and 120 μm (**b**). The laser wavelength is 577 nm (*solid*), 585 nm (*dotted*), and 590 nm (*dashed line*). Results of 3D Monte Carlo simulation with μ_a of blood as given in Table 22.1 for $SatO_2 = 100\%$, except $\mu_{a,bl}(590\text{ nm}) = 15.7\text{ mm}^{-1}$, and $\mu_{s,bl} = 46.8\text{ mm}^{-1}$ for all wavelengths [41]. For the assumed optical properties of epidermis and dermis see Table 23.2

(marked with "a"), irradiation at $\lambda = 577$ nm results in a very non-uniform temperature profile along the axis through the vessel center (solid line). The effect is even more pronounced in the larger blood vessel ("b", solid line), where the ratio between maximal and minimal local heating approaches 10. A significant contribution of diffuse light to energy deposition at the deeper part of the vessel lumen is notable in both cases.

A much more uniform energy deposition across the vessel lumen is obtained in both 60 and 120 μm vessels upon irradiation at $\lambda = 585$ nm, with $\mu_{a,bl}$ 46% smaller than at 577 nm (Fig. 23.4, dotted lines). Moreover, in the larger vessel ("b"), energy deposition near the vessel axis is significantly higher than at 577 nm. Enhanced energy deposition near the vessel center thereby compensates, at least in part, the reduced heating of blood in the superficial portion of the vessel lumen (Table 23.2).

Table 23.2 Absorption coefficient (μ_a), scattering coefficient (μ_s), and anisotropy factor (g) of epidermis and dermis, used in numerical study presented in Figs. 23.4, 23.5, and 23.7 [41]

	Wavelength (nm)	Absorption coefficient (mm^{-1})	Scattering coefficient (mm^{-1})	Anisotropy factor
Epidermis	577	1.85	48.0	0.787
	585, 590	1.80	47.0	0.790
Dermis	577	0.22	13.1	0.787
	585, 590	0.22	12.9	0.790

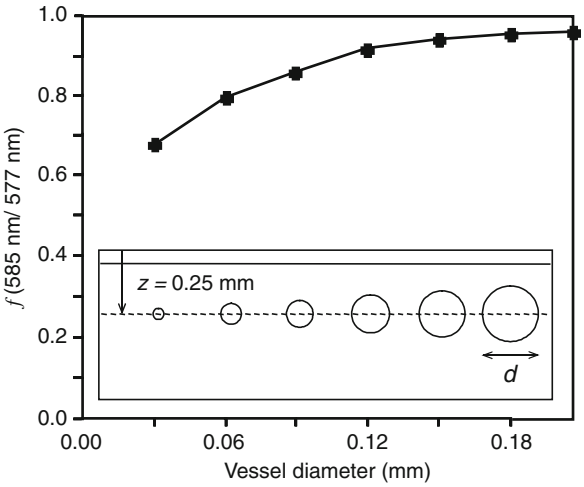


Fig. 23.5 Ratio of total energy deposited in a single cylindrical vessel at $\lambda = 585$ nm as compared to 577 nm. Vessels with different diameters are located at the same subsurface depth (see the sketch). Results from a 3D Monte Carlo simulation [41]. For optical properties see Fig. 23.4

As a result, the ratio between the total energy deposited in the vessel at $\lambda = 585$ nm as compared to 577 nm is not as small as suggested by the ratio of the corresponding values $\mu_{a,\mathrm{bl}}(\lambda)$. In Fig. 23.5, this ratio (f) is plotted as a function of vessel diameter d [41]. For optically thin vessels ($\mu_a d \ll 1$; not included in this simulation), f would approach $\mu_{a,\mathrm{bl}}(585\ \mathrm{nm})/\mu_{a,\mathrm{bl}}(595\ \mathrm{nm})$, which is 0.54 for the presented case. With increasing vessel diameter, f gradually increases to almost 1, which implies that the total deposited energy becomes nearly independent of wavelength.

Note, however, that relation $f(d)$ presented in Fig. 23.5 is not universal. The specific values depend, for example, on the amount and spectral properties of epidermal melanin assumed in the model, as well as depth of simulated blood vessel. A more universal relation, pertaining exclusively to optical screening within a single vessel, was provided by Verkruysse et al. [42]. Under the assumption of isotropic diffuse irradiation, these authors introduced a correction factor, C_{diff}, as the average fluence rate across the vessel lumen, $\langle\phi\rangle$, divided by the value just outside the vessel wall, ϕ_0. The heating power, invested in a certain segment of the vessel, can thus be expressed as

$$P_v = \mu_a \int \phi(r)\, dV = \mu_a \langle\phi\rangle V = C_{\mathrm{diff}}\, \phi_0 V \tag{23.1}$$

Analytically computed C_{diff} and the analogous correction factor for irradiation with a collimated beam, C_{coll}, are presented in Fig. 23.6a as a function of product $R\mu_{a,\mathrm{bl}}$ (solid and dashed lines, respectively). Both correction factors equal 1 for optically thin vessels, and decrease for larger vessel radius and/or absorption coefficient values when the average fluence rate $\langle\phi\rangle$ is progressively reduced by optical screening.

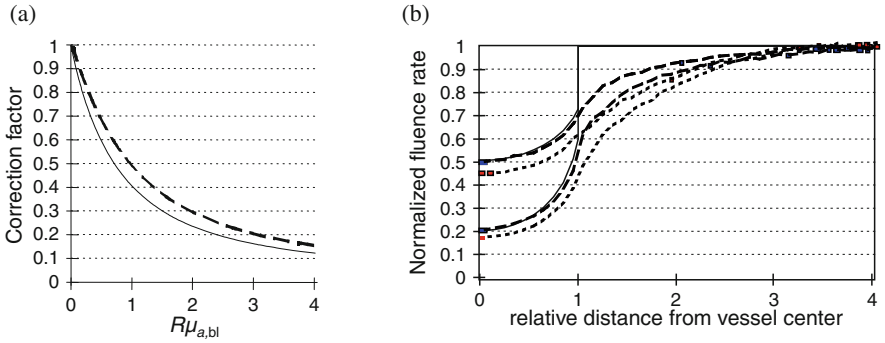


Fig. 23.6 (**a**) Analytically computed correction factor C_{diff} (1) for diffuse (*solid line*) and collimated irradiation of a single vessel (C_{coll}, *dashed line*) as a function of the product $R_{\mu a,\text{bl}}$. (**b**) Fluence rate profile $\phi\,(r)$ within and around the diffusely irradiated vessel as a function of distance from the vessel axis relative to its radius ($R = 50$ μm). Analytical results (*solid lines*) are compared with Monte Carlo results for $\mu_{s,\text{der}} = 20\ \text{mm}^{-1}$(*dashed line*) and 60 mm^{-1} (*dotted*). The upper and lower sets of curves correspond to $R_{\mu a,\text{bl}} = 0.5$ and $R_{\mu a,\text{bl}} = 1.25$, respectively [42]

A closed-form approximation for C_{diff} (1), accurate within ±5%, was provided by van Gemert et al. [43]:

$$C_{\text{diff}} = \begin{cases} \exp\left(-0.803\, R\,\mu_{a,\text{bl}}\right); & 0.6 \le R\,\mu_{a,\text{bl}} \le 1.5 \\ \exp\left(-0.983\,\sqrt{R\,\mu_{a,\text{bl}}}\right); & 1.5 < R\,\mu_{a,\text{bl}} \le 4 \end{cases} \tag{23.2}$$

Analytical analysis of optical screening within a single vessel under collimated irradiation can be found in Kimel et al., Svaasand et al., and Nelson et al. [44–46].

Figure 23.6b presents the analytically obtained fluence rate profiles, $\phi\,(r)$, divided by the value far from the vessel, for two values of product $R\mu_{a,\text{bl}}$ (0.5 and 1.25, respectively; solid lines). Within the vessel, the two profiles are virtually identical to results of a numerical model (MC) for a dermal scattering coefficient of 20 mm^{-1} (dashed lines) but differ somewhat for $\mu_{s,\text{der}} = 60\ \text{mm}^{-1}$(dotted).

Optical screening puts an upper limit on the sizes of blood vessels that can be effectively treated with a selected laser, and an upper limit on the value of $\mu_{\text{a,bl}}(\lambda)$ used for a specific vessel size. Since the optical penetration depths in the spectral band covered by PDLs (577–600 nm) approximately match the range of PWS vessel diameters, it is generally true that smaller and larger PWS blood vessels should be treated using shorter and longer wavelengths, respectively.

Indeed, histology after therapy of PWS at $\lambda = 585$ nm ($t_{\text{p}} = 0.45$ ms, $H = 6$–$8\ \text{J/cm}^2$) indicates complete vessel wall coagulation up to a vessel diameter of $d = 150$ μm; with even larger blood vessels damaged only in the superficial portion of the lumen [23]. However, the enhanced coagulation of larger vessels (as compared to the 577 nm results mentioned above) could in part be attributed also to the larger laser beam diameter (Section 23.2.1) and longer pulse duration (Section 23.2.3).

23.2.2.3 Optical Interaction Between Nearby Vessels – Shading

Due to the scattering-dominated nature of optical transport in dermis at the clinically relevant wavelengths ($\mu_{s,\mathrm{der}} = 13{-}25\ \mathrm{mm}^{-1}, g = 0.77{-}0.8$), highly selective absorption in blood vessels affects the fluence rate profile in their immediate vicinity. As seen in Fig. 23.6b, this effect varies with specific values of $\mu_{a,\mathrm{bl}}$ and $\mu_{s,\mathrm{der}}$, but typically extends more than one vessel diameter (d) into surrounding dermis. When another vessel lies in close proximity, the fluence rate in either vessel will be lowered due to absorption of light in the other vessel.

This effect is illustrated in Fig. 23.7a which presents a numerically obtained cross-sectional map of energy deposition in two parallel vessels ($d = 120\ \mu\mathrm{m}$), which cross the axis of a laser beam at subsurface depths of 0.25 and 0.5 mm [42]. The plotted lines of constant energy deposition clearly indicate the reduction of dermal fluence below and to the side of each vessel. At a sufficient distance from the vessel wall (on the order of d), dermal fluence recovers to the unperturbed level.

Figure 23.7b presents the energy deposition profile through the center of the same vessels (solid line). The dotted line (first vessel) and thin solid line (second vessel) indicate the energy deposition profile in either vessel when the other is absent. Thus, the difference between the two profiles in the lower vessel demonstrates the amount of shading by the superficial vessel. Note, however, that the relative reduction of intraluminal fluence in the superficial vessel, caused by absorption in the deeper vessel, is of similar magnitude. In general, two blood vessels in close proximity will mutually reduce the fluence rates and energy deposition within each other [41].

The discussed shading effect should not be associated with geometrical shadow, which is underlined by results from another Monte Carlo simulation of optical transport in a PWS model involving a regular grid of capillaries (Fig. 23.8) [30]. The indicated column heights are proportional to energy deposition in each volume

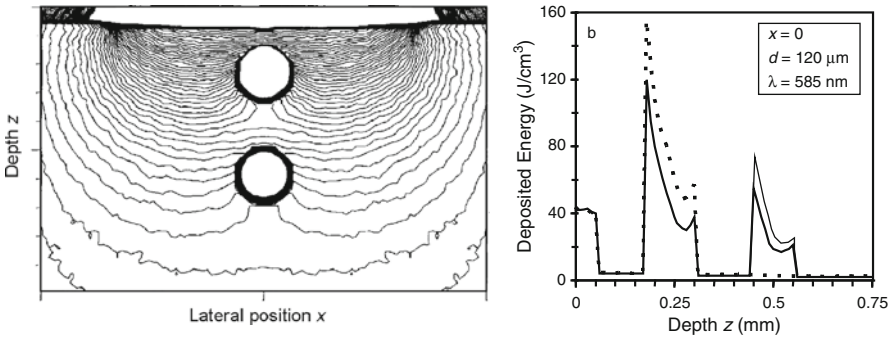


Fig. 23.7 (**a**) Cross-sectional map of deposited energy density for two straight vessels ($d = 120\ \mu\mathrm{m}$), crossing the laser beam axis ($D = 1$ mm) at subsurface depths of 0.25 and 0.5 mm. Lines of constant energy deposition are plotted in steps of 0.1 J/cm^3. (**b**) Depth profiles of energy deposition through both vessel centers (*solid line*). The dotted line (first vessel) and thin solid line (second vessel) indicate the energy deposition profiles in the absence of the other vessel. Results from a 3D Monte Carlo simulation [41] for $\lambda = 585$ nm. (Assumed optical properties as in Fig. 23.4)

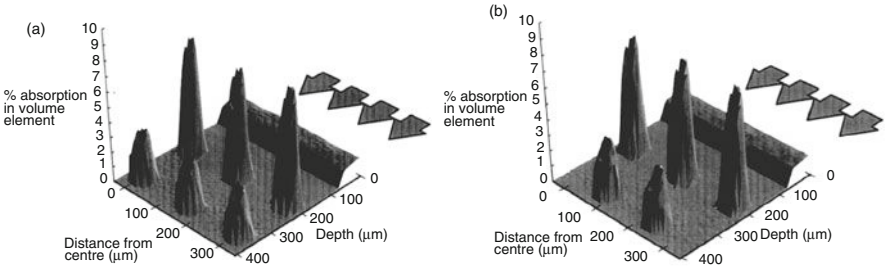


Fig. 23.8 Shading of the deeper vessels due to light absorption in the superficial capillary layer as seen in a numerical model. Note the marginal difference in heating of the lower capillaries when positioned exactly underneath the more superficial vessels (**a**) or between them (**b**). Results from a 2D Monte Carlo model; only a small part of the capillary grid near the axis of the laser beam with diameter $D = 3$ mm is presented [30]. The assumed optical properties at $\lambda = 577$ nm are $\mu_{a,\mathrm{bl}} = 35.4\ \mathrm{mm}^{-1}$ and $\mu_{s,\mathrm{bl}} = 46.8\ \mathrm{mm}^{-1}$, the rest as given in Table 23.2 except $\mu_{a,\mathrm{epi}} = 1.9\ \mathrm{mm}^{-1}$ and $\mu_{s,\mathrm{der}} = 21.0\ \mathrm{mm}^{-1}$

element. Note that the difference between energy deposition in the lower capillaries positioned exactly underneath the more superficial vessels (Fig. 23.8a) or half way between them (Fig. 23.8b) is only marginal. Similarly, energy deposition in the epidermis (seen along the upper right edge) is virtually uniform, despite the fact that strong absorption in the superficial vessels affects the fluence distribution in the vessels' immediate vicinity.

The influence of laser wavelength is now considered by analyzing energy deposition along the laser beam axis through the center of one capillary in each PWS layer (Fig. 23.9) [30]. At $\lambda = 577$ nm, absorption in the superficial capillary layer induces pronounced shading of the deeper capillaries (solid line). At $\lambda = 585$ nm, with a lower absorption coefficient in blood ($\mu_{a,\mathrm{bl}} = 19.1\ \mathrm{mm}^{-1}$, vs. $35.4\ \mathrm{mm}^{-1}$ at 577 nm), the effective decline of energy deposition with depth is significantly reduced (dashed line). As a result, vessels deeper than 0.3 mm absorb progressively more incident radiation at 585 nm as compared to 577 nm. This is in qualitative agreement with histology results, which showed an increase in average depth of vascular injury after PDL treatment of PWS, from 0.72 mm at $\lambda = 577$ nm to 1.16 at 585 nm [40].

Specifically, the deepest vessel in Fig. 23.9 ($z = 0.95$ mm) absorbs at $\lambda = 585$ nm approximately twice as much energy, despite the fact that $\mu_{a,\mathrm{bl}}(585\ \mathrm{nm})$ is approximately half as small as compared to 577 nm. Because the energy deposition rate is equal to the product $\mu_a(\lambda)\ \phi(\lambda)$, the dermal $\phi(585\ \mathrm{nm})$ at this depth is obviously several times higher as compared to 577 nm. As a consequence of the higher fluence levels associated with the less absorbing 585 nm wavelength, even the most superficial capillary absorbs almost as much energy at 585 nm as compared to 577 nm, despite the markedly lower $\mu_{a,\mathrm{bl}}$. Reduced screening within this 50 μm vessel, albeit apparent in Fig. 23.9, cannot fully account for this result. For the same reason, energy deposition in the epidermis is also slightly higher at 585 nm as compared to 577 nm [30, 47].

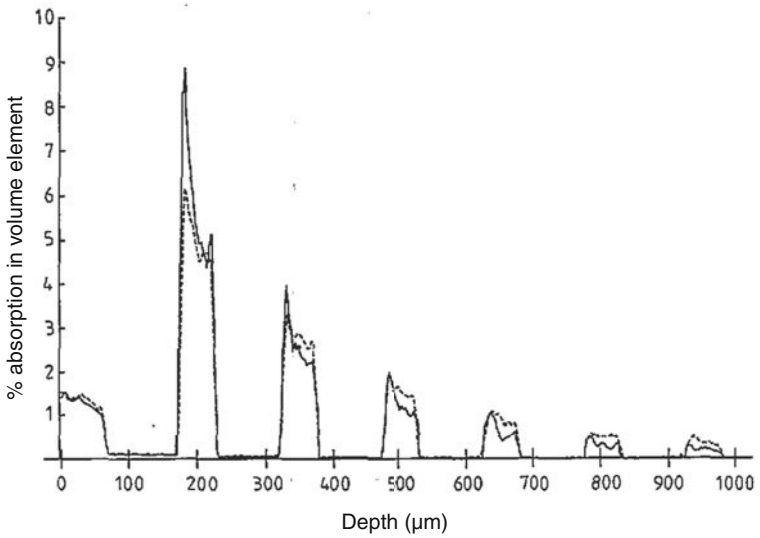


Fig. 23.9 Fractional absorption of incident fluence along the laser beam axis through the center of one capillary in each PWS layer (see Fig. 23.8a) for irradiation at $\lambda = 577$ nm (*solid line*) and 585 nm (*dashed line*). Result from a 2D Monte Carlo model [30]. The optical properties at $\lambda = 585$ nm are $\mu_{a,\mathrm{bl}} = 19.1\ \mathrm{mm}^{-1}$ and $\mu_{s,\mathrm{bl}} = 46.7\ \mathrm{mm}^{-1}$, the rest as in Table 23.2, except $\mu_{a,\mathrm{epi}} = 1.9\ \mathrm{mm}^{-1}$ and $\mu_{s,\mathrm{der}} = 20.5\ \mathrm{mm}^{-1}$; values for $\lambda = 577$ nm are the same as in Fig. 23.8

Numerical modeling involving digitized histology of PWS and spectrally dependent melanin absorption indicated that the same effect may lead to higher epidermal temperature after irradiation at $\lambda = 600$ nm as compared to 585 nm, despite the several times lower $\mu_{a,\mathrm{epi}}$ at 600 nm (Fig. 23.10) [50]. In fact, we have observed this effect in measurements of laser-induced temperature depth profiles in PWS in vivo (see Fig. 23.23a) [49]. At longer PDL wavelengths, with lower absorption in both blood and melanin, the epidermal damage threshold can thus be lower than at $\lambda = 585$ nm.

Due to the shading effect, switching to a slightly less absorbed wavelength can – in addition to reduced optical screening within larger blood vessels – result in increased energy deposition in deeper vessels of any size. However, light that is absorbed too poorly will induce an insufficient temperature rise in the PWS vasculature, at the same or even higher epidermal heating (Fig. 23.10a). What is, therefore, the optimal wavelength for PWS laser therapy?

To answer this conundrum, van Gemert et al. [51] applied a diffusion approximation solution to four geometrical models, representing different types of PWS (adult, pediatric, etc.). In contrast with earlier studies, these authors accounted for the fact that dermal blood, responsible for optical shading, is not distributed homogeneously throughout the superficial dermis, but is contained within vessels of certain radius (R). By multiplying $\mu_{a,\mathrm{bl}}(\lambda)$ with the corresponding correction factor C_{diff} (1), diffusion theory replicates quite well the fluence rate profiles obtained from MC modeling of discrete blood vessels [42].

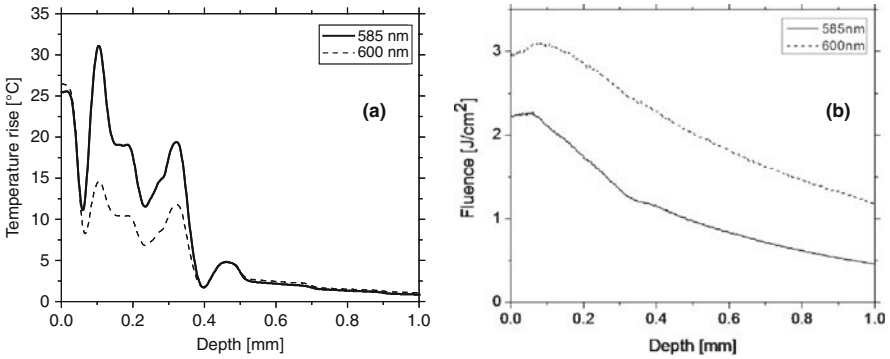


Fig. 23.10 (**a**) Computed temperature depth profiles in model 3D PWS structure after irradiation at $\lambda = 585$ nm (*solid line*) and 600 nm (*dashed*) with radiant exposure $H = 1$ J/cm^2, and (**b**) the corresponding fluence depth profiles. Absorption coefficients: $\mu_{a,\mathrm{bl}}(585\ \mathrm{nm}) = 17.7\ \mathrm{mm}^{-1}$, $\mu_{a,\mathrm{bl}}(600\ \mathrm{nm}) = 3.2\ \mathrm{mm}^{-1}$, $\mu_{a,\mathrm{epi}}(585\mathrm{nm}) = 1.3\ \mathrm{mm}^{-1}$, $\mu_{a,\mathrm{epi}}(600\ \mathrm{nm}) = 1.2\ \mathrm{mm}^{-1}$[50]

The average volumetric heat production in the target blood vessel (Q_3) computed for the geometrical model representing an adult PWS is presented in Fig. 23.11a as a function of λ. If the target vessel with radius $R_3 = 50\ \mu$m lies directly underneath the superficial plexus of similar vessels ($z_2 = 0$; solid line), wavelengths $\lambda = 575$–580 nm are most efficient, and $Q_3(585\ \mathrm{nm})$ is only marginally lower. However, when the two blood vessels are separated by healthy dermis of increasing thickness z_2, the optimal wavelength gradually increases up to ~590 nm at $z_2 = 1$ mm (dashed line). By considering the predicted benefits and drawbacks of specific laser wavelengths (also from results not presented here), the authors suggested that $\lambda = 585$–587 nm would be preferred over 577 nm for treatment of adult PWS patients.

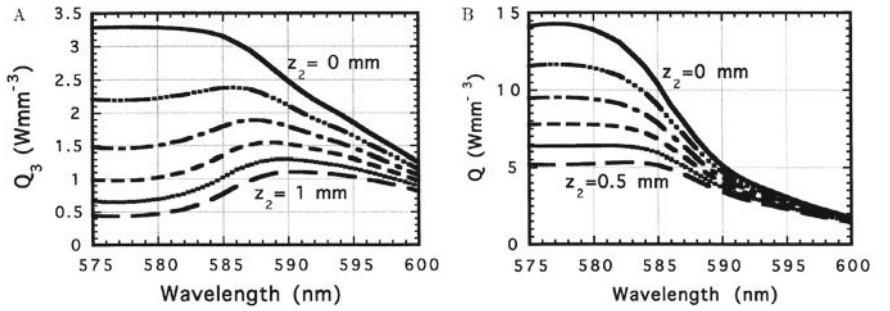


Fig. 23.11 Average volumetric heat production in the target blood vessel (Q_3) computed from diffusion approximation solution with correction factor C_{diff} (1) for two geometrical models of PWS consisting of: (**a**) a target vessel with radius $R_3 = 50\ \mu$m, located under a 0.3 mm thick superficial plexus with vessels of the same size ($R_1 = 50\ \mu$m) and a second dermal layer with normal capillaries ($R_2 = 2\ \mu$m) and variable thickness ($z_2 = 0, 0.2, 0.4, 0.8$, and 1.0 mm; see marks in the graph); (**b**) same elements, but with dimensions $R_1 = R_3 = 12.5\ \mu$m, $z_1 = 0.15$ mm, and thickness z_2 varied in a narrower range (0–0.5 mm; see the graph) [51]

In contrast, when a PWS vessel is smaller and positioned more superficially, as is the case in light-colored pediatric lesions, $\lambda = 577$ nm is either more, or equally efficient as 585–587 nm, again depending on target vessel depth, z_2 (Fig. 23.11b). Nevertheless, because heat deposition in these superficial lesions is sufficiently high also at $\lambda = 585$–587 nm, the authors concluded that the latter wavelength range offers the best treatment efficacy for a broadest range of PWS patients [51].

Note that the computed optimal wavelengths cannot be taken for granted as a clinical guideline due to several limitations of this study. For example, the results rely on diffusion theory and do not account for epidermal scattering or absorption. The spectral dependence of melanin absorption will reduce the values $Q_3(\lambda)$ progressively toward the shorter λ, thus shifting the maxima in the opposite direction. Furthermore, the applied criterion (i.e., maximizing the heating of a selected target vessel) may not be clinically viable because of potentially excessive heating of more superficial vessels. Finally, the influence of pulse duration on redistribution of heat within the vessels was not considered.

Nevertheless, this study demonstrated that even in grossly simplified geometrical models of PWS, variation of one parameter can significantly affect the selection of optimal wavelength, and that the latter can have a critical impact on therapeutic effect. By comparing the results for different plausible, albeit idealized PWS geometries, the study provided an important insight into why some lasers can be very effective in some patients, but fail to induce a significant blanching response in other PWS lesions. This suggests two possible approaches for further improvement of therapeutic success: a systematic variation of treatment parameters throughout an extended treatment protocol; or optimization of treatment parameters on an individual patient basis (Section 23.4.1).

Several recent studies applied the Monte Carlo approach to more complex models of PWS geometry, including the digitized histology of an actual lesion [52–54]. While the predicted effects are increasingly more realistic, the computational complexity of such an approach often prevents systematic variation of structural, optical and/or treatment parameters over the relevant ranges. In addition, heterogeneity and disorder in such model structures makes it difficult to link the observed trends to a specific lesion characteristic. However, this may be the only way to study the response of some PWS components, such as the clusters of small blood vessels. The optical and thermal interaction between the vessels in such clusters can be so strong that they behave almost like a single large vessel.

23.2.2.4 Clinical Experience

A comparison study by Tan et al. [40] showed that 585 nm laser irradiation of PWS can achieve a superior clinical outcome as compared to 577 nm. Histology confirmed that shifting the wavelength to 585 nm increased the dermal depth to which irreversible vascular injury in PWS could be induced. In addition, adult patients who had incomplete clearance of their PWS at 577 nm responded well to 585 nm irradiation.

In a similar study involving previously untreated infants, Raulin et al. [55] found irradiation at 585 nm to be more efficient as compared to 532 nm. Chang et al. [56] and Greve et al. [57] found 585 nm also superior to 595 nm. The difference in blanching after laser treatment with two different wavelengths is presented in Fig. 23.12 [56]. As is commonly observed in clinical practice, the response within each segment ("a" or "b") is quite homogeneous, which suggests that the clinically relevant PWS characteristics vary only marginally between adjacent sites within an optically uniform lesion area.

However, the difference between mean blanching scores achieved with 585 nm as compared to 595 nm in all included patients was much smaller than presented in Fig. 23.12. Similarly, Scherer et al. [59] reported that using wavelengths longer than 585 nm improved the clinical response in some PWS, although 585 nm appeared optimal for all patients. These results were attributed to reduced screening, which enabled larger vessels to be damaged more severely using increased radiant exposures.

Kimel et al. [32] concluded (based on controlled experiments in an animal model and analytical considerations) that 595 nm irradiation should be more effective for deeper vessels while at the same time inducing less purpura, albeit 585 nm was still suggested as the most effective wavelength, overall, particularly for PWS blood vessels with $d = 10$–$60\ \mu m$ [59].

However, Fiskerstrand et al. [14] reported that some (pink) PWS lesions "resistant" to treatment at 585 nm ($t_p = 0.45$ ms) responded very well to irradiation at 577 nm ($t_p = 0.36$ ms). Chowdhury et al. [60] found that a similar group of resistant PWS were treated effectively with the KTP laser at $\lambda = 532$ nm. Furthermore, Laube [61] found that $\lambda = 595$ nm was effective in PWS lesions, resistant to treatment at 585 nm.

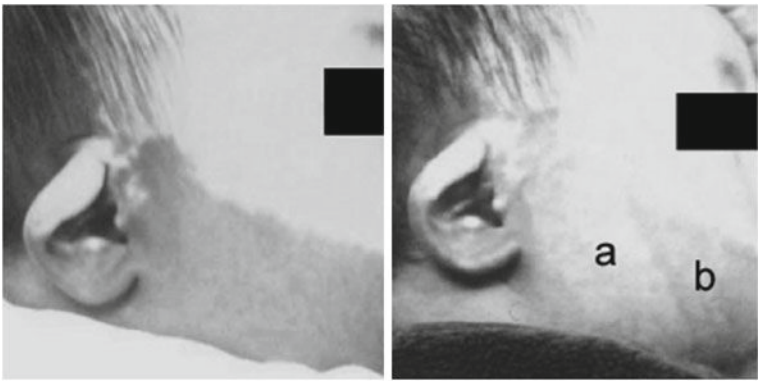


Fig. 23.12 (**a**) PWS lesion on the cheek of an infant before PDL treatment; (**b**) blanching response after two laser treatments at $\lambda = 585$ nm (segment "a") and 595 nm ("b"). Other treatment parameters were fixed [56]

While recalling that the remaining treatment parameters (e.g., pulse duration, beam diameter, radiant exposure) are not always directly comparable due to technical reasons or other practical limitations, we believe that such apparent discrepancies result primarily from large inter- and intra-patient variations in structural and optical characteristics between the lesions. This presents the major difficulty in determining the optimal wavelength for PWS laser therapy, and applies to the other treatment parameters. Numerical results, animal studies and clinical research together strongly suggest that adjustment of treatment parameters on an individual patient basis is a prerequisite for significant improvement in therapeutic outcome in the future (Section 23.4.1).

23.2.3 Pulse Duration

Upon irradiation with a suitably short laser pulse, energy is deposited in the absorbing structure before much heat can be transferred to the surrounding tissue by conduction. The resulting temperature rise in an optically and thermally homogeneous structure is thus directly proportional to the absorbed heat, Q, which is in turn proportional to μ_a and average fluence in the target, $\langle \psi \rangle = \langle \phi \rangle t_p$:

$$\Delta T = \frac{Q}{\rho c V} = \frac{\mu_a \langle \psi \rangle}{\rho c} \tag{23.3}$$

In general, however, a significant fraction of deposited heat may diffuse away from the absorbing structure during laser exposure, which reduces the peak target temperature and impairs the heating spatial selectivity even if the wavelength provides selective absorption of laser light. Therefore, selection of laser pulse duration, which governs spatial confinement of deposited heat in absorbing structures, is just as important as wavelength.

The ensuing biophysical effects will depend directly on temperature field evolution in and around the target vessels. Assuming a predominantly conductive heat transfer within the tissue, the latter can be obtained by solving the heat diffusion equation (see Chapter 10, Eq. (10.6)) with a volumetric heat production term $S(r) = \mu_a(r)\,\phi(r)$

$$\frac{\partial T}{\partial t} - \alpha \nabla^2 T = \frac{\mu_a \phi}{\rho c} \tag{23.4}$$

and appropriate boundary conditions. The perfusion term is not included because its effect is negligible on the time scales relevant for PWS laser therapy.

23.2.3.1 Thermal Relaxation Time of Blood Vessel as a Whole

The second corollary of the selective photothermolysis paradigm [9, 18] calls for a laser pulse duration which ensures significant confinement of deposited heat in

the blood vessels during the laser pulse. This is obtained by observing the thermal relaxation time of an infinite cylinder

$$\tau_v = \frac{d^2}{16\,\alpha} \tag{23.5}$$

In the original derivation, τ_v represents the time interval in which the amplitude of a hypothetical Gaussian temperature profile (dropping to $1/e$ of the maximal value at the vessel wall) decreases by a factor of 2 due to radial diffusion of heat into surrounding tissue. (Because the derivation is somewhat abstract, the exact meaning of τ_v is sometimes misinterpreted by non-specialists.)

Nevertheless, it is clear that a gradual transition from specific to nonspecific heating of a vessel with diameter d occurs as the laser pulse duration (t_p) approaches and then exceeds τ_v. At t_p longer or even comparable to τ_v, a significant amount of heat diffuses from the vessel during laser exposure, thus reducing the peak intravascular temperature and increasing the risk of unwanted thermal injury to adjacent tissue. Only significantly shorter laser pulses ($t_p << \tau_v$) enable a maximal temperature rise in the targeted structure to be achieved.

According to Eq. (23.5), confinement of laser energy within smaller vessels requires progressively shorter pulse durations. For a typical PWS vessel diameter, $d = 50$ μm (and thermal diffusivity of skin, ($\alpha = 0.11$ mm^2/s)), τ_v equals 1.4 ms. Due to the wide range of possible vessel diameters in PWS lesions and quadratic dependence on d, the corresponding values of τ_v are spread over three orders of magnitude (Table 23.3).

Experiments in an animal model (chicken chorio-allantoic membrane – CAM) showed that blood vessels with $d = 50 \pm 10$ μm experience significantly more thermal damage at $t_p = 0.45$ ms than at $t_p = 10$ ms, despite the lower radiant exposure applied during the shorter pulses (3–6 J/cm^2 and 7 J/cm^2, respectively; see Fig. 23.13a) [46]. This correlates well with the fact that 0.45 ms is shorter while 10 ms is several times longer than the corresponding τ_v. Moreover, smaller vessels showed less thermal damage than larger vessels ($\langle d \rangle = 70$ and 110 μm) after irradiation with $t_p = 10$ ms (and a fixed $H = 7$ J/cm^2), a direct consequence of their faster thermal relaxation [44].

Table 23.3 Thermal relaxation time, τ_v (Eq. (23.5)), for the range of PWS blood vessel diameters (d)

vessel diameter (μm)	thermal relaxation time (ms)
10	0.057
20	0.23
30	0.51
50	1.42
100	5.7
150	12.5
200	22.7
300	51

Using thermal diffusivity of skin ($\alpha = 0.11$ mm^2/s)

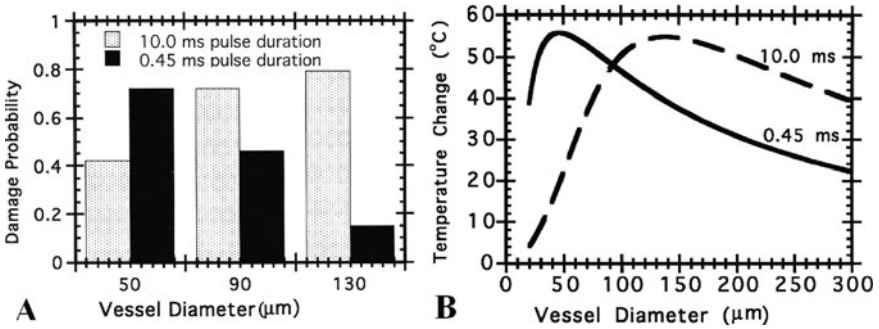


Fig. 23.13 (**a**) Experimentally determined probability of irreversible thermal damage to CAM blood vessels of three diameter classes, after pulsed laser irradiation at $\lambda = 585$ nm and $t_p = 0.45$ ms ($H = 3{-}6$ J/cm^2) or 10 ms ($H = 7$ J/cm^2). (**b**) Analytically computed mean temperature rise as a function of blood vessel diameter for $t_p = 0.45$ ms, $H = 3$ J/cm^2 (*solid curve*), and $t_p = 10$ ms, $H = 7$ J/cm^2 (*dash*), assuming $\mu_{a,\mathrm{bl}} = 17$ mm^{-1} for blood in CAM [46]

Inasmuch as pulses shorter than τ_v provide the highest intravascular temperatures with minimal risk of non-selective injury, can we use sufficiently short pulses ($t_p << 50$ μs) to ensure heat confinement in PWS vessels of all sizes? Unfortunately, no, for a number of reasons:

First, early generation vascular lasers with microsecond pulses ($t_p = 0.3{-}2.0$ μs) produced poor blanching of PWS [19]. Histology showed RBC aggregation and extensive hemorrhage, which was attributed to explosive vaporization of selectively absorbing RBC that could not thermalize with surrounding plasma during irradiation (see Section 23.2.3.2) [62]. Second, it was suggested that such very short pulses may not sustain heat diffusion from the hot lumen through the entire thickness of the vessel wall (Section 23.2.3.4) [45].

And finally, even at somewhat longer t_p, preventing the effects mentioned above, non-uniform deposition of laser energy across the lumen (e.g., due to optical screening – Section 23.2.2.2) may induce localized blood overheating, which can lead to explosive vaporization before the entire vessel wall has been coagulated. Under such circumstances, prolonging t_p helps redistribute the heat within the blood volume during irradiation and allows deposition of more energy without rupturing the vessel wall. Assuming primarily conductive heat transport within the vessel lumen (Eq. (23.4)), τ_v represents the time interval between instantaneous deposition of heat on the vessel axis and occurrence of peak temperature at the vessel wall. Consequently, τ_v provides an estimate of the laser pulse duration required to reduce the temperature inhomogeneity induced by optical screening.

Based on such reasoning, supported by systematic experimentation using PDL devices with different pulse durations, it was concluded that the optimal pulse duration for treatment of blood vessels with known diameter d approximates the corresponding τ_v. Pulse durations of 1–5 ms [44, 46] and $t_p = 1$–10 ms [63] were suggested as optimal for PWS laser therapy.

From a theoretical perspective, the rule of matching t_p to τ_v is neither exact, nor universal. The optimal t_p is determined by a trade-off between optical screening and

heat diffusion, and thus depends on $\mu_{a,\mathrm{bl}}(\lambda)$ (see Section 23.2.3.4). In general, the optimal t_p may thus depend on laser wavelength and PWS specifics such as vessel size distribution, thickness of the vessel wall and basal layer, blood oxygenation level, etc.

However, the related corrections are often minor in view of the wide range of plausible τ_v values (Table 23.3) and our inability to control the above listed factors in clinical practice. Moreover, the transition from specific to nonspecific vessel heating is very gradual and different time constants can be defined, depending on specific criteria used in the derivation. For example, for irradiation of smaller blood vessels and/or poorly absorbed light (i.e., $\mu_a d << 1$), it might be more appropriate to consider the thermal relaxation of a uniformly heated cylinder, which yields

$$\tau_v' = \frac{d^2}{16\,\alpha \ln 2} = 1.44\,\tau_v \tag{23.6}$$

In view of all the ambiguities, the stated simple rule provides a useful order-of-magnitude estimate and, as such, has been widely accepted by the medical laser community.

23.2.3.2 Thermal Relaxation Times of Erythrocytes and Melanosomes

Selectively absorbed lasers with microsecond pulses cause extravasation of RBC without complete necrosis of the vessel wall. A plausible explanation is that such short pulse durations do not allow for sufficient heat transfer from the discrete absorbers (hemoglobin containing RBC) to surrounding plasma and other blood cells. The excessive temperature of the RBC causes explosive vaporization (cavitation) and the associated shock waves rupture the vessel wall.

By restructuring the derivation of τ_v for spherical geometry, we obtain

$$\tau_{\mathrm{eri}} = \left(\sqrt[3]{4} - 1\right) \frac{d_{\mathrm{eri}}^2}{16\,\alpha} \approx \frac{d_{\mathrm{eri}}^2}{27.3\,\alpha} \tag{23.7}$$

The thermal relaxation time of a single RBC, when approximated as a 5 μm diameter sphere, is $\tau_{\mathrm{eri}} \sim 6$ μs (using thermal diffusivity of blood, $\alpha_{\mathrm{bl}} = 0.14\ \mathrm{mm^2/s}$). Laser pulses significantly longer than 10 μs should thus be used to prevent RBC explosion.

The same reasoning applies also to light absorption by epidermal melanin, which is encapsulated in spherical or elliptical melanosomes, typically with $d_{\mathrm{mel}} \leq 1$ μm. When laser pulses shorter than the corresponding $\tau_{\mathrm{mel}} \sim 0.3$ μs are used, melanosomes can experience a temperature rise much higher than the surrounding tissue [64]. This may also cause explosion of the host cell, potentially leading to adverse clinical side effects such as skin dyspigmentation. However, these effects are produced only by nanosecond laser pulses, which are not used for PWS therapy. Because $\tau_{\mathrm{mel}} < \tau_{\mathrm{eri}}$, the risk of melanosome explosion does not present an additional limitation on pulse duration, as long as the latter is longer than τ_{eri}.

Despite the above, the discrete nature of melanosome absorption may induce non-uniform heating with local temperatures far above the average epidermal value. This effect will be strongly reduced when t_p exceeds the characteristic time of heat transfer between individual melanosomes [64]. By replacing d_e in Eq. (23.7) with typical distances between melanosomes (2–10 μm), we can predict that the melanin-containing region of the epidermis should be heated uniformly for laser pulses significantly longer than 25 μs. Note, however, that similar reasoning can be applied also to melanocytes, where the concentration of melanosomes is largest.

Finally, let us consider thermal relaxation of the melanin-rich epidermal layer as a whole. By following the steps of the original derivation for one-dimensional heat transport, we obtain

$$\tau_{\mathrm{epi}} = \frac{3\, d_{\mathrm{epi}}^2}{16\, \alpha} \tag{23.8}$$

In non-tanned, light skin phototypes, melanosomes are concentrated primarily in a ~10 μm thick basal layer (typically located 50–100 μm below the skin surface) while in tanned, darker skin phototypes more melanin is distributed throughout the epidermis. As a result, τ_{epi} can vary from less than 1 ms to over 100 ms. While applying pulses longer than τ_{epi} would help reduce peak epidermal temperatures, this is not a viable approach to enhancing selective heating of PWS blood vessels, because their thermal relaxation time is most often shorter than τ_{epi}.

23.2.3.3 Heat Transport into the Vessel Wall

Because peak blood temperature is limited by the risk of explosive vaporization, laser pulses should be long enough to sustain transfer of heat from the hot lumen through the entire thickness of the vessel wall, d_w [45]. Pulses shorter than the corresponding characteristic time (τ_w) may not induce sufficient vessel wall temperatures adequate for coagulation. Because the vessel wall is much thinner in comparison to vessel diameter, a one dimensional solution of the heat diffusion equation can be applied, whereby

$$\tau_{\mathrm{w}} \sim \frac{d_{\mathrm{w}}^2}{\alpha} \tag{23.9}$$

Given the typical thickness of the PWS blood vessel wall, $d_w = 4{-}6$ μm [5], this implies that t_p should not be shorter than ~0.3 ms (for $\alpha = 0.11$ mm^2/s).

Kimel et al. [32] have extended this concept by approximating the wall thickness of any PWS blood vessel with $d_w = 0.1\ d$, which yields $\tau_w \sim d^2/(100\ \alpha)$. Consequently, the optimal pulse durations for treatment of PWS blood vessels with diameters from 10 to 150 μm should range from 10 μs to 2 ms. For $t_p > \tau_w$, a significant fraction of deposited heat is transferred through the vessel wall into the surrounding tissue, thereby reducing the peak wall temperature and limiting the

applicable radiant exposure due to increased risk of adverse side effects. This theory might help explain the modest overall efficacy of PWS therapy with $t_p > 1.5$ ms.

However, the above reasoning assumes a uniform blood temperature across the vessel lumen, which applies, e.g., to smaller vessels irradiated with deeply penetrating light (i.e., $\mu_{a,\mathrm{bl}} d << 1$) and situations with vigorous forced convection. Since the latter has not been unambiguously demonstrated in millisecond irradiation of blood vessels, the use of longer t_p may be required to overcome optical screening, especially in larger PWS vessels and when using laser wavelengths with higher $\mu_{a,\mathrm{bl}}$.

23.2.3.4 Interaction Between Thermal Relaxation and Optical Screening

The influence of pulse duration (t_p) on photocoagulation of blood vessels with different diameters, as assessed in animal model experiments (CAM), is illustrated in Fig. 23.13a [46]. When irradiated at $t_p = 10$ ms (light columns), the vessel group with smallest diameters ($d = 50 \pm 10$ μm) experienced significantly less thermal injury than the larger vessels, $\langle d \rangle = 90$ and 130 μm, in agreement with the thermal relaxation dynamics of the vessel as a whole (Section 23.2.3.1). Thermal relaxation time τ_v for the 50 μm vessel is 7 times shorter than $t_p = 10$ ms, so heat diffusion to surrounding tissue strongly reduces peak intravascular temperature. For the 130 μm vessel, $\tau_v = 9.4$ ms is comparable to t_p and the heat loss is much smaller.

In contrast, blood vessels with $\langle d \rangle = 50$ μm showed significantly more damage than larger vessels when exposed to laser pulses with $t_p = 0.45$ ms (black columns). In this case, $t_p < \tau_v$ for all three vessel diameters, so the influence of heat diffusion during irradiation is negligible. The observed effect can thus be attributed to optical screening, which progressively reduces the average fluence, $\langle \psi \rangle$, in larger vessels (see Section 23.2.2.2).

Because of the interaction between thermal relaxation and optical screening, irradiation at a given wavelength and pulse duration has the strongest effect on PWS blood vessels within a certain range of diameters. With decreasing vessel size, peak intraluminal temperature is progressively reduced by heat loss to surrounding tissue, as τ_v becomes shorter than t_p. In larger vessels, however, energy deposition is reduced by optical screening (see Figs. 23.4 and 23.6). The combined effect of both mechanisms as estimated using an analytical model of collimated irradiation and heat diffusion is illustrated in Fig. 23.13b [44, 46].

Figure 23.14 presents radiant exposure values required to raise the average temperature in a single subsurface blood vessel of varying diameter to 70°C. The results were obtained by numerically solving the heat diffusion equation (23.4) with $\phi(\mathbf{r})$ determined using a 3D Monte Carlo simulation of irradiation at $\lambda = 585$ nm and six pulse durations from 0.22 to 9.2 ms [65].

For each pulse duration t_p, the computed threshold radiant exposure (H_{th}) has a distinct minimum at a specific vessel diameter, d. At $t_p = 0.45$ ms, e.g., the lowest H is required to induce irreversible injury in blood vessels with $d = 30$–50 μm which supports the results shown in Fig. 23.13, and roughly corresponds to the rule

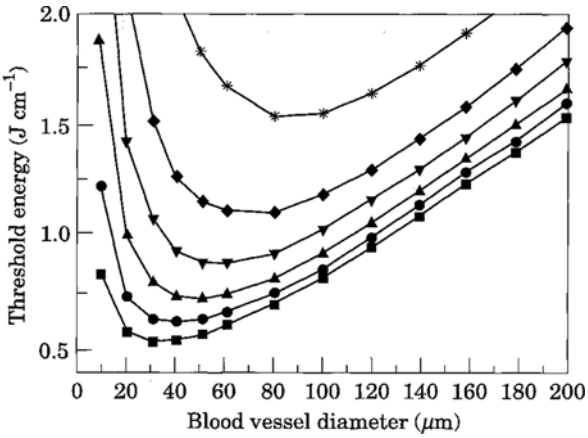


Fig. 23.14 Radiant exposure values required to induce average temperature of 70°C in blood vessels with different diameters when using laser pulse duration of: 0.22 ms (■), 0.45 ms (●), 0.9 ms (▲), 1.8 ms (▼), 3.6 ms (◆), and 9.2 ms (∗). Results from numerical model of optical and thermal transport in skin, including a 50 μm thick epidermal layer and a single blood vessel at dermal depth of 250 μm. The assumed optical properties of skin at $\lambda = 585$ nm are as in Table 23.2, except $\mu_{a,\mathrm{der}} = 0.25\ \mathrm{mm}^{-1}$ and $\mu_{s,\mathrm{der}} = 13.1\ \mathrm{mm}^{-1}$, and $\mu_{a,\mathrm{bl}} = 19.1\ \mathrm{mm}^{-1}$ and $\mu_{s,\mathrm{bl}} = 46.7\ \mathrm{mm}^{-1}$ for blood [65]

of matching t_p and τ_v (see Table 23.3). The steep increase of H_{th} for smaller vessels clearly results from the increased heat loss, as $\tau_v < t_p$ and continues to decrease.

Although not stated by the original authors, the increase of H_{th} with larger vessel diameters can be attributed primarily to optical screening. For the value of $\mu_{a,\mathrm{bl}} = 19.1\ \mathrm{mm}^{-1}$ assumed in the model, the related reduction of $\langle\psi\rangle$ begins at vessel diameters several times smaller than $1/\mu_{a,\mathrm{bl}} = 52$ μm (see Fig. 23.6a). Similar results were obtained also by Svaasand et al. [45], who explicitly accounted for screening in their analytical model of collimated irradiation.

Over a range of laser pulse durations, blood vessels with minimal H_{th} are thus determined by the interaction between the effects of thermal relaxation and optical screening, which primarily affect smaller and larger vessels, respectively. It is therefore not surprising that the minima shift to smaller vessel diameters when the analysis is repeated for $\lambda = 577$ nm with a higher $\mu_{a,\mathrm{bl}}$ [43]. Furthermore, the results vary as a function of target vessel depth [65].

As a result, blood vessel diameters that will be affected most at a given set of laser irradiation parameters (λ, t_p, etc.) in general deviate from the value predicted by simply matching t_p to $\tau_v(d)$. For example, the results in Fig. 23.14 indicate the maximal effect of the 9.2 ms pulses on vessels with $d = 80{-}100$ μm, while the rule suggests $d = 4(\alpha\ t_p)^{1/2} = 127$ μm.

Despite the simplicity of the model geometries used in the above theoretical studies, the results (Figs. 23.13b and 23.14) strongly suggest that blood vessels with certain diameters will be damaged selectively upon irradiation with a specific combination of laser wavelength, pulse duration, and radiant exposure.

However, this does not necessarily imply that the optimal effect in a vessel with diameter d will be provided by a pulse duration with $t_p \approx \tau_v(d)$. As seen in Fig. 23.14, the shortest pulse duration available offers the lowest H_{th} for all vessel sizes. The same effect can be seen also by rescaling the 10 ms result in Fig. 23.13b (dashed line) to the lower radiant exposure used for $t_p = 0.45$ ms (solid line).

This apparent discrepancy may result, in part, from the specific criterion for thermal injury used in this particular study; the average temperature across the vessel lumen reaching $\langle T \rangle = 70°\text{C}$. In examples with pronounced optical screening, achieving this endpoint may not ensure coagulation of the entire vessel wall. Moreover, explosive vaporization of RBC, which is not calculated in the model, may occur locally at even lower average temperatures. Taking these effects into consideration would likely suggest the use of longer pulses for large vessels, but it is impossible to say exactly how much longer. It is safe to conclude, however, that the relatively slow thermal relaxation of larger vessels allows prolonging of t_p (to a value sufficiently smaller than $\tau_v(d)$) with only moderate reduction of $\langle T \rangle$.

23.2.3.5 Clinical Experience

Early clinical studies showed improved efficacy of PDL treatment of PWS when t_p was increased from 0.3 to 360 μs [66] (see Sections 23.2.3.2 and 23.2.3.3).

The modest shifting of the minima in Fig. 23.14 with decreasing t_p and steep increase of H_{th} for smaller d provides insight into one class of PWS lesions resistant to PDL therapy. These PWS often consist of many small vessels, with $\langle d \rangle$ less than 20 μm, resulting clinically in a light pink appearance [14, 45]. As indicated by Fig. 23.14, these vessels can be very difficult to damage within the range of customary treatment parameters ($t_p = 0.5{-}10$ ms at $\lambda = 532, 585,$ or 595 nm).

The response of such lesions might be improved by using even shorter pulse durations, possibly at wavelengths with higher $\mu_{a,\text{bl}}$. In rodent model experiments, 2 μs laser pulses at $\lambda = 577$ nm selectively destroyed vessels as small as $d = 10$ μm [20]. However, it might be preferable to avoid clinical use of such short pulses, not only to avoid hemorrhage due to RBC vaporization in most PWS vessels, but also to prevent skin necrosis due to thermal damage in healthy capillaries lying above the ectatic blood vessels [65, 67]. In fact, the inherent resistance of normal capillaries to thermal injury at the customary treatment parameters may be an important element of safety during this procedure.

In contrast, PDL with a longer pulse duration ($t_p = 1.5$ ms) was found effective in treating a PWS that was resistant to treatment with $t_p = 0.5$ ms [68]. It is important to note that in PWS with larger blood vessels, where treatment efficacy is limited by optical screening, blood cavitation becomes less prominent with increasing pulse duration [21].

Parlette et al. [69] irradiated small leg veins ($d = 0.1{-}1.6$ mm) in vivo with a deeply penetrating Nd:YAG laser ($\lambda = 1064$ nm), thus minimizing optical screening. Each vessel was irradiated at the corresponding threshold for acute response, which turned out to be lower for larger vessels. The authors reported that longer pulses ($t_p = 20{-}60$ ms) resulted in better vessel clearance and less side effects

(purpura and post-inflammatory hyperpigmentation) than shorter pulses (3–10 ms). Histology showed more perivascular collagen injury and vessel contraction with the longer pulses, while intraluminal thrombosis prevailed at $t_p = 3$ ms.

23.2.4 Radiant Exposure

At any given laser pulse duration, a too low radiant exposure will obviously not provide enough heat in the targeted PWS blood vessels to achieve and sustain the elevated vessel wall temperature required for coagulation. At a too high exposure, however, unwanted explosive vaporization of RBC may occur, inducing purpura and stimulating PWS revascularization by tissue repair mechanisms [18, 19]. The therapeutic window between the two corresponding thresholds is a complicated and incompletely understood function of laser wavelength and pulse duration, blood vessel size and depth, average dermal blood content, etc. [23, 70, 71]. In view of the heterogeneity and complexity of each individual PWS lesion, as well as large variability in blood vessel size and depth distribution between different lesions, defining a universally valid safe and therapeutically effective radiant exposure is practically impossible.

Light fluence at the subsurface target location is affected also by epidermal melanin absorption. At the same time, melanin absorption induces localized heating in the epidermis which may lead to acute skin surface disruption (blistering or crusting) resulting in permanent side effects such as dyspigmentation and scarring. Use of excessive radiant exposure thus greatly increases the risk of clinical complications.

In addition to the melanin concentration, which can vary by a factor of more than 20 between different human skin phototypes, the epidermal fluence and subsequent heating depends also on specifics of the PWS lesion, such as volumetric blood content, blood vessel size and depth distribution. Because of the combined influence of all these processes, the threshold radiant exposure for epidermal damage (H_d) will depend on laser wavelength and pulse duration in a complex and – given the unknown structure of the PWS lesion in clinical practice – largely unpredictable manner.

Particularly in moderately dark and heavily pigmented skin phototypes (III–VI), the threshold radiant exposure for thermal damage to the epidermal basal layer is often very close to or even lower than that required for optimal PWS blanching. In many patients, the radiant exposure that can be safely applied is thereby limited by the risk of epidermal injury, which directly affects therapeutic outcome. This limitation can be removed or alleviated by dynamic cooling of the skin surface prior to laser irradiation.

23.2.5 Active Cooling

Although melanin absorption inevitably results in local epidermal heating during laser exposure, the temperature can be prevented from exceeding the threshold for thermal injury by active cooling of the skin surface before irradiation.

This concept was investigated first by Gilchrest et al. [72] and Welch et al. [73], who applied ice and freon spray, respectively, to the skin surface prior to laser irradiation. Despite achieving a reduction of epidermal injury, neither approach led to an improved therapeutic response. Because cooling was applied for a prolonged period of time, temperature of the targeted blood vessels was reduced by a similar amount, thus offsetting the benefit of epidermal cooling.

In contrast to laser heating, cooling can be applied only to the skin surface and the depth of the cooling effect is governed exclusively by laws of heat diffusion. A rigorous solution to the heat diffusion equation matches that for surface heating (see Chapter 10), albeit with a negative heat source. The effective cooling depth scales with the heat diffusion length, l_c, defined as

$$l_c = \sqrt{\alpha\, t_c} \tag{23.10}$$

With a given skin thermal diffusivity ($\alpha = 0.11$ mm^2/s), the extent of cooling can thus be controlled only by varying the cooling time, t_c – the duration the applied coolant is in contact with the skin surface.

According to Eq. (23.10), the cooling effect will extend through the epidermis (50–100 μm thick) for t_c around 20–90 ms. Numerical analysis (Fig. 23.15) shows that, for example, after cooling for $t_c = 60$ ms the temperature drop at the corresponding depth $l_c = 81$ μm is approximately one half of that at the skin surface, and the cooling effect is still significant at depth $2\, l_c$. Because the cooling time is so important, this technique is often referred to as "dynamic cooling".

Aggressive cooling of the skin surface is required to induce a significant epidermal temperature drop over such short cooling times. The technologies implemented in current clinical devices include sprays of cryogenic liquids, actively cooled contact handpieces, and chilled air [74]. Regardless of the approach, the temperature reduction and gradient inside the skin are limited by the rate of heat transfer at the skin surface. The so-called convective boundary condition is applicable ([75], p. 18),

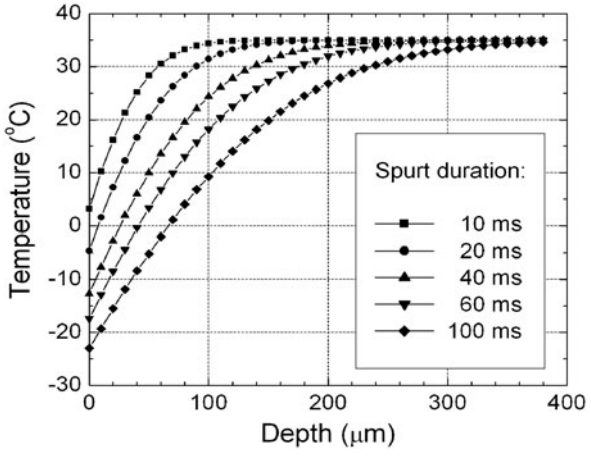


Fig. 23.15 Computed temperature profiles in skin after application of cryogen spray cooling of variable duration t_c (see the legend). Heat transfer coefficient: $h = 6000$ W/m^{2}K; cryogen temperature: $T_c = -50$°C

whereby the heat flux, q_c, is proportional to the temperature difference between the skin surface and cooling medium (T_c):

$$q_c = h\,[T(z=0) - T_c]\,. \tag{23.11}$$

Within this approximation, the influence of all physical mechanisms involved in the process is combined into the heat transfer coefficient, h.

23.2.5.1 Cryogen Spray Cooling

The main principle of cryogen spray cooling (CSC) is rapid evaporation of cryogenic liquid, which extracts latent heat from the skin surface [76, 77]. The cryogenic compound used in current dermatologic laser systems is tetrafluoroethane ($C_2H_2F_4$, also known as R134a). Its low boiling temperature (−26.2°C) and relatively high latent heat of vaporization (117 J/g) enable a higher heat extraction rate as compared to chilled water sprays. During flight to the skin surface, spray droplets evaporatively cool to $T_c = -40$ to -60°C. Heat transfer coefficients up to $h = 10,000{-}60,000$ W/m^{2}K [78–82] and heat extraction rates q_c in excess of 50 W/cm^2 [81, 83] have been reported for various experimental spraying and measurement approaches. For two commercial CSC devices, the average value for a 100 ms cryogen spurt was determined at $h = 6200{-}8400$ W/m^{2}K [82, 84].

Despite the low temperatures experienced transiently near the skin surface (Fig. 23.15), spurt durations up to 100 ms have been used frequently in our clinic with no evidence of cryo-injury. However, if the risk of cryo-injury becomes a serious concern (e.g., with longer spurt durations), less aggressive spray cooling regimes could be obtained by adjustment of the spraying distance and nozzle design [84] or intermittent spraying [83]. The latter approach underlines an important advantage of CSC, the possibility to electronically control timing of the cryogen spurt which offers a predictable cooling effect and reliable safety margin with respect to undesirable thermal injury.

With some CSC devices, a thin layer of liquid cryogen builds up on the skin surface and extends the cooling effect beyond the user-specified spurt duration, t_s (Fig. 23.16b). Since tetrafluoroethane is transparent in the visible and near-IR, such a film does not present a significant optical barrier for laser irradiation [85]. A layer of frost often forms on the cooled skin surface (Fig. 23.16c, d) due to sublimation of atmospheric vapor. Uncontrolled scattering of the incident laser light by the frost could affect the fluence levels inside the skin. In our experience, observable frost forms on human skin only after the cryogen layer has retracted and usually occurs long after the laser pulse is completed and does not interfere with irradiation.

However, condensation of ambient water vapor on the cooled skin surface impairs the efficacy of CSC due to associated deposition of latent heat (Fig. 23.17) [86]. In clinical practice, one can thus safely apply only the radiant exposures as allowed in the worst case scenario (i.e., high humidity). In this regard, a means for controlling humidity around the treatment area (e.g. by blowing dry air or nitrogen

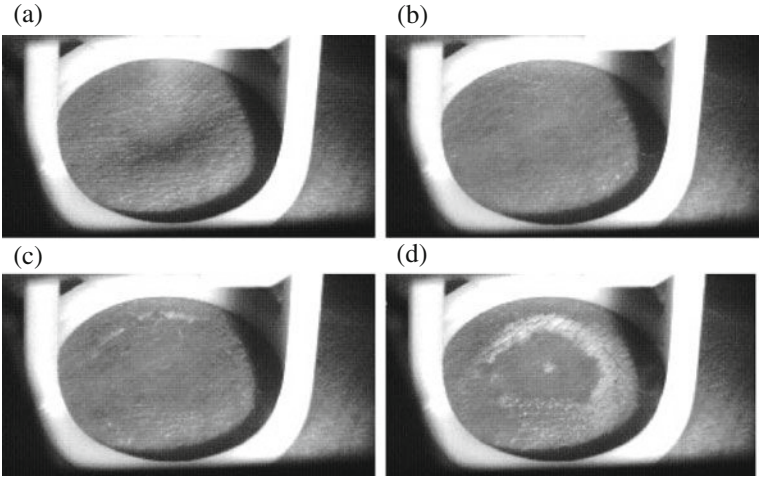


Fig. 23.16 Human skin (forearm) during CSC with an 80 ms cryogen spurt. Photographs were obtained at times (**a**) $t = 30$ ms; (**b**) $t = 130$ ms; (**c**) $t = 180$ ms; (**d**) $t = 230$ ms. Note persistent liquid layer after the spurt (b) and formation of frost around the retracting cryogen pool (c), (d) [86]

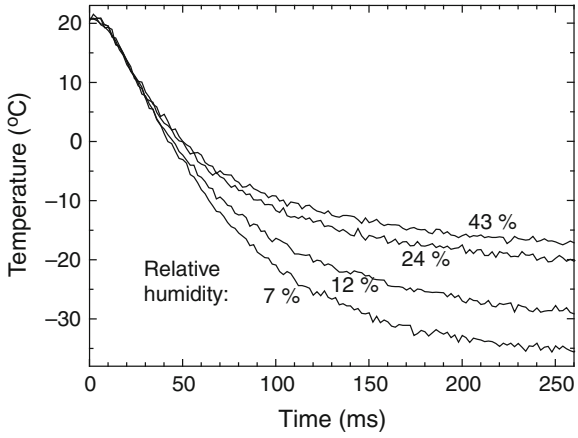


Fig. 23.17 Temperature evolution 90 μm below the surface during CSC in controlled humidity environment (relative humidity values are given next to each curve). Temperature was measured with a miniature thermocouple embedded in epoxy block with thermal properties similar to those of human skin [86]

gas) or a calibration instrument to verify the actual cooling performance of the CSC device might be a useful extension of the existing technology.

23.2.5.2 Contact Cooling and Chilled Air

Contact cooling (CC) relies on heat conduction to a cooled solid body pressed against the patient's skin surface. A transparent window is often used through which laser radiation is delivered. The window can be cooled using chilled water, liquid

cryogen, or thermoelectric elements. The best cooling efficacy is achieved using windows made from highly conductive materials, such as sapphire.

In the laboratory, the heat transfer coefficient between a sapphire plate ($T_c = 6-12°C$) and flat solid body can exceed $h = 2000$ W/m^{2}K [87]. In clinical practice, the heat extraction rate is reduced by imperfect thermal contact between the cooled window and human skin. The quality of thermal contact varies with force applied by the operator, elasticity of treated skin and underlying tissue, density and thickness of hair and, particularly, the type and amount of substances present on the skin surface (sweat, hydrating/lubricating gel, trapped air) [87, 88].

Window temperatures below 0 °C are often employed to achieve sufficiently high heat flux q_c (Eq. (23.11)), which raises concerns with respect to transient freezing of the superficial epidermal layers. In order to increase significantly the therapeutic radiant exposures without risk of epidermal thermal damage or frostbite, it is thus important to have an active control of the induced cooling effect [88]. Ambient water will also condense on the cooled window and build up a layer of frost at temperatures below 0°C, potentially interfering with laser irradiation.

Due to higher T_c and lower h as compared to CSC, the steep temperature gradients required to permit a significant increase in the safe radiant exposure for PWS laser therapy are difficult to achieve with CC. This is illustrated by measurements of basal layer temperature in human skin in vivo during contact cooling with a sapphire plate at $T_c = -4°C$ and a commercial CSC device (Fig. 23.18) [89].

The skin surface can also be pre-cooled with chilled air. Despite using air temperatures as low as −30°C, this method offers the lowest cooling rate, since h for forced convection of gas typically ranges from 25 to 250 W/m^{2}K. Cooling times on the order of several seconds are required to induce a significant

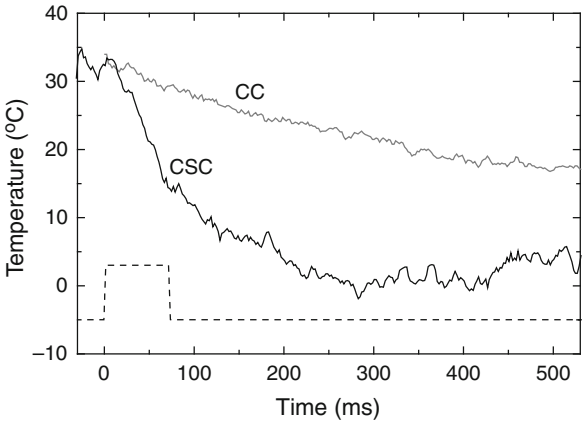


Fig. 23.18 Temperature measured at the epidermal basal layer (depth approx. 80 μm) in response to dynamic cooling with two commercial devices: contact cooling with a sapphire plate at −4°C (CC) and cryogen spray cooling (CSC). *Dashed line* indicates timing of the cryogen spurt ($t_s = 80$ ms). (Data courtesy of Karl Pope; [89])

temperature reduction in the epidermis, inevitably resulting in non-selective ("bulk") cooling of the entire skin. Consequently, this technique offers only minimal protection against laser-induced epidermal damage during PWS laser therapy, but may offer other clinical benefits such as reduction of patient discomfort and pain.

23.2.5.3 Optimal Cooling Duration

Cryogen spray cooling allows independent selection of spurt duration (t_s) and delay before laser irradiation (t_d), which brings additional flexibility to the laser treatment protocol. Based on average epidermal thickness and other practical considerations, the most common settings used for PWS laser therapy are $t_s = 30-50$ ms and $t_d = 30-50$ ms. With commercial CSC devices that deposit a liquid cryogen film on the skin surface (see Fig. 23.16), this provides an effective cooling time of $t_c = 60-100$ ms.

When analyzing the efficacy of CSC for a specific PWS geometry, cooling selectivity can be considered optimal when the temperature difference between the targeted blood vessel and the epidermal basal layer is maximal. As seen in Fig. 23.19a, the temperatures at the basal layer (depth $z_b = 60$ μm) and two hypothetical target depths (150 and 400 μm) continue to decrease after the 100 ms long CSC terminates (solid lines) [90]. Because the basal layer temperature decreases faster than the temperatures at the two target depths, the temperature differences $\Delta T_{t,b}$ also continue to increase for a short time (Fig. 23.19b, solid lines). In the presented example, applying a delay of several tens of milliseconds would therefore improve cooling selectivity by a small margin. Note, however, that the optimal delays and related benefits can be significantly larger, especially for shorter cooling

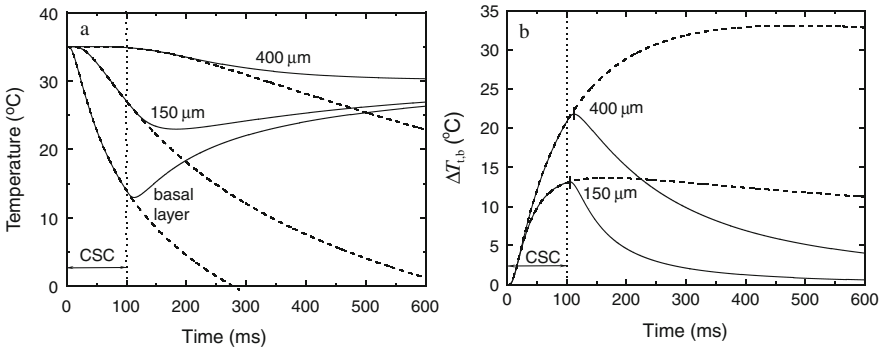


Fig. 23.19 (**a**) Computed temperature evolutions at the basal layer (depth $z_b = 60$ μm) and two target depths in dermis (150 and 400 μm) during and following a 100 ms long CSC. (**b**) Temperature difference between each target and the basal layer. *Vertical lines* (*dotted*) indicate the end of CSC ($h = 5000$ W/m^{2}K, $T_c = -44$°C), *dashed curves* indicate the effects of prolonged cooling [90]

times (t_c) and deeper z_b [90]. Ideally, to permit safe use of the highest radiant exposure, t_c and additional delay should be adjusted according to specific geometry of each PWS lesion.

The same study demonstrated also that prolonged cooling often enhances spatial selectivity beyond that achieved using the additional delay. In the discussed example, a moderate increase of t_c would result in significantly larger $\Delta T_{t,b}$ for the deeper target, without compromising the value for the shallower one (Fig. 23.19b, dashed lines). For the involved PWS geometries, rather long cooling times (t_c) are allowed before $\Delta T_{t,b}$ begins to markedly decrease, suggesting that setting the t_c for a prolonged CSC (with no additional delay) is less critical than setting the optimal delay after a shorter t_c (see Fig. 23.19b, solid lines).

As a practical consequence, prolonged CSC ($t_c > 100$ ms) with no delay before the laser pulse presents a good compromise for a wide range of PWS lesion geometries and values of h. Selecting a delay t_d longer than the residence time of deposited cryogen layer is justified only when the total cooling time t_c is limited to durations shorter than the optimal value, e.g. due to technical constraints or risk of cryo-injury.

23.2.5.4 Maximal Radiant Exposure Applicable with Dynamic Cooling

Application of dynamic cooling allows safe use of significantly higher radiant exposures than was possible previously. Let us consider a hypothetical example, where irreversible thermal damage in the epidermis occurs at a radiant exposure H_d and the corresponding peak temperature rise at the basal layer is $\Delta T_{\max}$. By precooling the latter by ΔT_c, the damage threshold temperature rise will increase to $\Delta T_{\max} + \Delta T_c$). To avoid epidermal damage, the applied radiant exposure at the same irradiation conditions should thus remain below

$$H_c = \frac{(\Delta T_{\max} + \Delta T_c)}{\Delta T_{\max}} H_d \tag{23.12}$$

As a first-order estimate, let us assume that $\Delta T_{\max} = 30$ K (difference between the baseline skin temperature of 35°C and coagulation onset at 65°C). According to Eq. (23.12), precooling of the basal layer by $\Delta T_c = 30$ K (Figs. 23.18 and 23.19) enables a 100% increase in the radiant exposure that can be safely applied. For PWS treatment of Caucasian skin, for example, PDL radiant exposure can be increased from earlier used values of 6–8 to 12–16 J/cm^2. Such an increase can significantly improve the PWS blanching response in cases where the applicable radiant exposure was limited by epidermal damage, and increase the safety margin for other patients.

The predicted benefits of dynamic cooling were confirmed in an early clinical comparison of cryogen-spray cooled (CSC) and non-cooled (NC) laser therapy of 50 patients with PWS [91]. Half of the patients received NC therapy at customary radiant exposures ($H = 5-9$ J/cm^2), the other half received CSC treatment with $H = 10-15$ J/cm^2. The results showed a significantly enhanced blanching response to CSC over NC treatment for all patients combined. Moreover, no permanent scarring or hypopigmentation developed after CSC treatment, vs. two cases

noted in the NC group, despite the significantly higher radiant exposures used with CSC treatment.

In addition, active cooling reduces the pain and distress associated with laser therapy [14, 76, 77, 92] which can be explained by the fact that CSC lowers and shortens the temperature rise in skin, experienced by the pain sensors.

23.3 Beyond the Customary Paradigm

23.3.1 Sequential Application of Multiple Laser Pulses

23.3.1.1 Extended Pulse PDL and KTP Lasers

PDLs with pulse durations longer than 0.5 ms often emit sequences of 3–5 sub-ms spikes. Kimel et al. [32] compared thermal injury caused in CAM (chicken chorioallantoic membrane) blood vessels by two lasers with a nominal pulse duration of 1.5 ms: one with a "smooth" intensity profile, and the other emitting a sequence of three 100 μs spikes. Across the three size groups (d from 40–50 to 110–120 μm) and both vessel types (arterioles and venules) included in the study, no difference in vessel injury was observed between the two lasers when used at the same wavelength (595 nm) and radiant exposure (15 J/cm^2).

It appears that the 0.6 ms delay between the individual laser spikes does not significantly affect vessel wall temperature evolution, probably due to slower transfer of heat from the absorbing lumen. A recent comparison of the same two lasers in PWS patients did not reveal any statistically significant difference in blanching of the test patches ($H = 14$ J/cm^2) [31].

In PDL and KTP lasers with "extended" pulse durations ($t_p = 5-50$ ms), time intervals between the spikes are accordingly longer (~3–20 ms) and interfere with thermal relaxation of most PWS blood vessels (Table 23.3). As illustrated in Fig. 23.20, such irradiation sequences can produce higher peak temperatures in

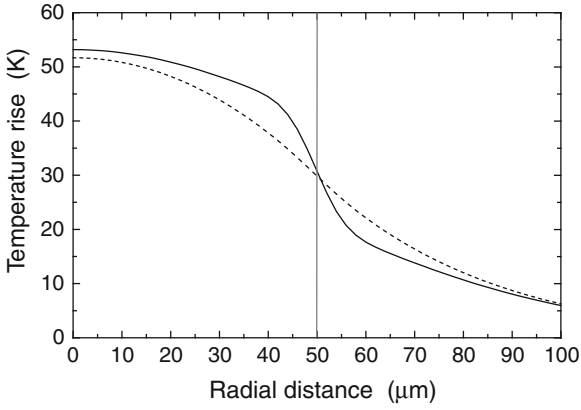


Fig. 23.20 Radial temperature profiles 0.1 ms after the end of an irradiation sequence consisting of 5 instantaneous laser pulses separated by 5 ms (*solid line*) and after continuous deposition of the same energy in 20 ms (*dashed*). Both curves are analytical solutions of the heat diffusion equation for a homogeneously heated cylinder with diameter of 100 μm (indicated by the *vertical line*)

the absorbing vascular lumen as compared with continuous irradiation with the same pulse duration and radiant exposure [93]. Nevertheless, peak temperatures inside the vessel wall differ only marginally between the two irradiation regimes. Of course, individual laser spikes must be long enough to allow thermal relaxation of RBC and melanosomes (Section 23.2.3.2).

23.3.1.2 Multiple Laser Pulses with Multiple Cryogen Spurts

In many PWS patients, particularly those with darker skin phototypes, the main reason for poor efficacy of laser treatment is inadequate heating of larger blood vessels [13, 23, 70, 71]. Due to the combined effect of melanin absorption and optical screening by larger PWS vessels, the latter can be only partially coagulated at radiant exposures below the epidermal damage threshold, despite the use of dynamic cooling.

One possible solution is to divide the radiant exposure into multiple laser pulses (MLP) while controlling the epidermal temperature with multiple cryogen spurts (MCS) applied between successive laser pulses. A suitable pulse repetition rate must be used to ensure accumulation of deposited heat in the targeted blood vessels, raising their temperature with each successive pulse.

We illustrate this approach using numerical solutions of the heat diffusion equation (Eq. (23.4)) with experimentally determined heat flux during CSC (q_c) used in surface boundary condition. For each laser pulse, the fluence rate distribution $\phi(r)$ [W/cm^2] is obtained from a Monte Carlo model of light transport ($\lambda = 532$ nm) in scattering epidermis and dermis with subsurface vessels of different diameters [94]. A sequence of five 1 ms laser pulses ($H = 4$ J/cm^{2}per pulse), separated by 60 ms cryogen spurts, is compared with a single 1 ms laser pulse ($t_p = 1$ ms, $H = 5$ J/cm^2). In both cases, active precooling ($t_c = 80$ ms) is applied before irradiation.

The result in Fig. 23.21a shows that MCS maintain the basal layer temperature below 60°C throughout the irradiation sequence (solid line). Heat accumulation in the 50 μm blood vessel (Fig. 23.21b) is marginal because its thermal relaxation time ($\tau_v = 1.4$ ms) is much shorter than the 60 ms interval between the successive laser pulses. In contrast, temperature at the center of the 150 μm vessel with a longer τ_v increases with each successive laser pulse (Fig. 23.21c, solid line). In this manner, the peak temperature in larger vessels, where light fluence is reduced by optical screening, approaches that achieved in smaller vessels. In comparison, a single laser pulse causes a significantly lower peak temperature in larger vessels, while the epidermal temperature exceeds 70°C.

By using a threshold temperature of 70°C to estimate the extent of irreversible thermal injury, cross-sectional views of the temperature distributions immediately after irradiation suggest that a single laser pulse will not induce photocoagulation of the entire wall in vessels larger than $d = 150-200$ μm (Fig. 23.22a). Meanwhile, MLP with MCS enables complete intraluminal coagulation and uniform heating of the vessel wall up to $d = 300$ μm, while at the same time diminishing the risk of epidermal injury (Fig. 23.22b).

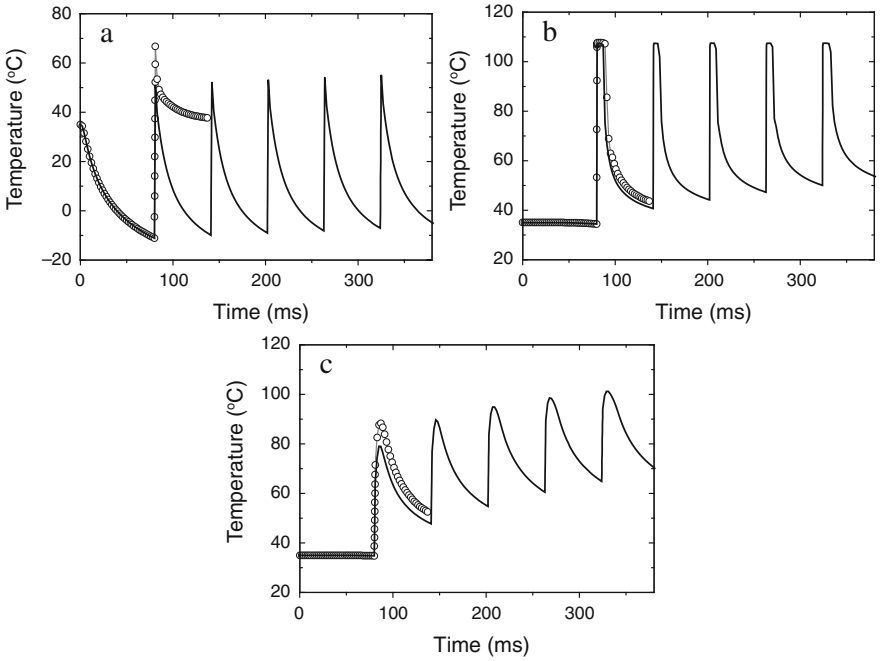


Fig. 23.21 Simulated temperature dynamics during MLP with MCS ($t_p = 1$ ms, $H = 4$ J/cm^2 per pulse; *solid lines*) at the basal layer (**a**) and on the axis of blood vessels with diameters $d = 50$ μm (**b**) and $d = 150$ μm (**c**). Results for a single laser pulse are presented for comparison ($t_p = 1$ ms, $H = 5$ J/cm^2; *circles*)

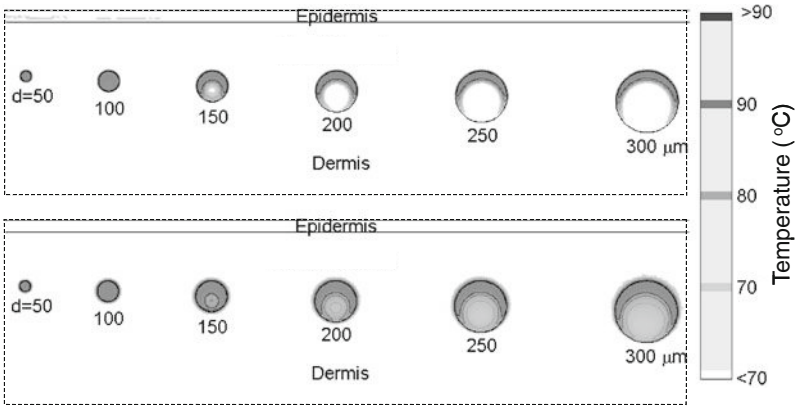


Fig. 23.22 Temperature maps after irradiation of subsurface blood vessels of different diameters with: (**a**) a single laser pulse ($H = 5$ J/cm^2) with CSC; and (**b**) sequential irradiation (at $H = 4$ J/cm^2) with MCS

In preliminary animal experiments, blood vessels ($d \sim 150$ μm) resistant to single-pulse irradiation at 4 J/cm^2 could be irreversibly injured using MLP (5 pulses of 3 J/cm^2 at repetition rate of 27 Hz) without MCS [94]. Although the parameters of MLP-MCS have yet to be optimized and in vivo damage thresholds reevaluated, the potential implications of this approach for PWS treatment are obvious.

23.3.1.3 Multiple Passes of the Laser Beam

In animal model experiments, Kimel et al. [32] observed that cumulative thermal damage to CAM blood vessels 5–10 min after irradiation (at $\lambda = 595$ nm) was significantly greater than immediately (less than 1 min) post irradiation. A follow-up study revealed a steady increase in cumulative damage between 1 and 10 min post irradiation [59]. Application of a second laser pulse 10 min after the first exposure increased the cumulative thermal damage by a lesser amount in all four included vessel groups.

Significantly different diffuse reflectance spectra (DRS; see Section 23.4.1) were measured in PWS lesions 1 and 2 min after irradiation at $\lambda = 585$ nm [95]. Because a 1 min delay allows for complete thermal relaxation of all blood vessels, the difference must be due to photo-thermally initiated biophysical processes or the skin's physiologic response. Numerical analysis of the spectra indicated a larger dermal blood volume at 2 min post irradiation, consistent with the onset of erythema and/or hemorrhage ($t_p = 0.45$ ms, $H \leq 5.3$ J/cm^2).

In recent clinical trials, sequential irradiation of PWS test spots with a PDL separated in time by several minutes (two "passes") lowered the purpuric threshold by 1 J/cm^2 as compared to customary treatment [96]. Further pulse-stacking experiments with dye lasers [97, 98] have confirmed the enduring effect lasting up to 1 min, which could be attributed to irreversible changes in the optical properties of blood (see Section 23.3.2). The effect is likely enhanced (or enabled) by development of a macroscopic coagulum, which slows blood flow and prevents removal of modified blood before application of a subsequent laser pulse [99].

In an earlier pilot study, Bencini et al. [100] found a superior response of PWS lesions to a two-pass laser treatment with different irradiation parameters (first pass: $\lambda = 590-600$ nm, $t_p = 1.5$ ms; second pass: $\lambda = 585$ nm, $t_p = 0.45$ ms) as compared to conventional treatment. However, this result can be attributed to the fact that the applied treatment parameter combinations preferentially affect different blood vessel populations (larger and deeper vs. smaller and shallower vessels, respectively; see Figs. 23.9 and 23.14). It is difficult to speculate whether this result is related to biochemical interaction between the two passes.

The processes mediating the interaction between the first and subsequent laser exposures in vivo are currently incompletely understood and therefore not amenable to quantitative analysis or modeling. The potential advantages and optimal parameters of multi-pass PWS laser therapy remain to be determined.

23.3.2 Dynamic Changes in Tissue Optical Properties During Laser Irradiation

In the theoretical analysis presented so far in this chapter, optical and thermal properties of tissue were assumed to remain constant during and after laser irradiation. In reality, blood photocoagulation is a complicated sequence of events which involve dynamic changes of optical properties, most notably the red shift of the hemoglobin absorption peak, variations of scattering coefficient due to morphological changes of RBCs, and thermally induced formation of methemoglobin (met-Hb).

The red (bathochromic) shift of the hemoglobin absorption spectrum with temperature increases $\mu_{a,\mathrm{bl}}$ on the long-wavelength side of the oxy-Hb absorption peak ($\lambda = 580-620$ nm) [101]. This effect is essentially instantaneous. During a 0.5 ms irradiation, inducing a temperature rise from 20 to 100°C, $\mu_{a,\mathrm{HbO}}$ at 585 nm increases by a factor of 2.5 [102]. On the short-wavelength side of the HbO absorption peak, $\mu_{a,\mathrm{bl}}$ decreases; e.g., by ~10% at $\lambda = 532$ nm [101].

The scattering coefficient of whole blood changes upon heating due to RBC transitioning from a biconcave to spherical shape (at 45–50°C) [103], RBC fragmentation and hemolysis (around 65°C) [104, 105], and subsequent formation of macroscopic coagulum [101].

Conversion of hemoglobin to met-Hb upon laser irradiation of blood was suggested first by Barton et al. [106]. In their in vitro experiments, first signs of transformation began after depositing 3–5 J/cm^2 (delivered over 3–5 ms) at $\lambda = 532$ nm. Met-Hb was formed at 65–72.5°C [106, 107] and occurred almost simultaneously with hemoglobin denaturation during millisecond irradiation [101, 108]. Measurements of diffuse reflectance spectra in vivo indicated the presence of met-Hb in PWS lesions irradiated with radiant exposures greater than 5 J/cm^2 ($\lambda = 585$ nm, $t_p = 0.45$ ms) [95].

The effects listed above clearly play an important role in PWS laser treatment. For example, the red shift, increased blood scattering and formation of met-Hb all combine to reduce the light penetration depth at $\lambda = 595-600$ nm. As a result, threshold radiant exposures for blood vaporization [109] and the extent of vascular damage in human skin [110] obtained with different PDL wavelengths (yellow-orange) differ much less than predicted from static optical properties. Figure 23.23a presents temperature profiles measured in PWS after irradiation at 585 (top) as compared to 600 nm (bottom) at identical radiant exposure [49]. The difference in heating amplitude is much smaller than the ratio between the corresponding absorption coefficients in venous blood (Table 23.1).

Conversely, the red shift and conversion of HbO to met-Hb will both reduce $\mu_{a,\mathrm{bl}}$ at 532 nm [101, 106], which alleviates optical screening and results in more homogeneous energy deposition in larger blood vessels. Moreover, as superficial vessels are heated to 60–70°C during the initial portion of the laser pulse, shading will also be reduced, allowing deeper penetration of the remaining laser exposure.

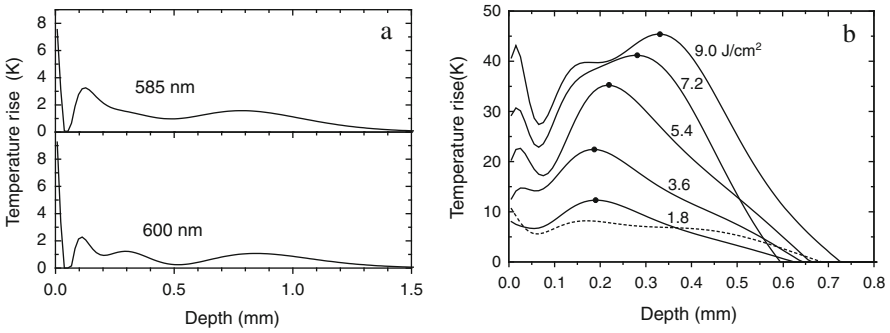


Fig. 23.23 (**a**) Temperature profiles induced in a PWS lesion with a 1.5 ms PDL pulse at $\lambda =$ 585 (*top*) and 600 nm (*bottom*). Both profiles are normalized to radiant exposure $H = 1$ J/cm^2. (**b**) Temperature profiles in PWS after pulsed irradiation ($t_p = 25$ ms) with increasing H (*solid lines*, see values next to each curve). *Dashed line* indicates the temperature profile induced with $H = 1.8$ J/cm^2 after completion of the irradiation sequence. All results were measured in vivo using pulsed photothermal radiometry [49]

Figure 23.23b presents temperature profiles measured in PWS using pulsed photothermal radiometry at $\lambda = 532$ nm (see Section 23.4.1.2) [49]. The temperature rise at $z \sim 0.2$ mm increases sub-linearly with H, indicating a reduced blood absorption at higher radiant exposures. This is supported by the depths of peak PWS temperature (marked by black dots), which are initially 0.19 mm below the skin surface, but move progressively deeper above $H \sim 5$ J/cm^2. The super-linear increase in epidermal temperature above $H \sim 7$ J/cm^2 indicates an increased epidermal fluence, a plausible consequence of reduced absorption in the superficial dermis. Temperature profile induced at the same lesion site with $H = 1.8$ J/cm^2 after completion of the irradiation sequence (dashed line), when compared to that obtained initially at the same H (the lowest solid line), also shows reduced temperature rise at $z \sim 0.2$ mm, higher epidermal temperature, and deeper energy deposition, which confirms that the underlying photo-thermal changes were largely irreversible.

Formation of met-Hb also increases μ_a of venous blood at 1,064 nm by a factor of 3–4 [101, 106], which offers improved selectivity of energy deposition in blood vessels. However, the rapid increase in blood absorption and scattering can lead to uncontrolled heat deposition, which can cause randomly occurring adverse effects [106].

Because met-Hb is only slowly reduced to Hb, the former might be used as a target for photothermolysis for some time after initial light exposure. This may be related to some effects observed in multi-pass PWS therapy (Section 23.3.1.3). Barton et al. [106] have suggested that the reported dynamic changes of blood optical properties could be utilized in development of a treatment approach, whereby two appropriately sequenced laser pulses at different wavelengths are applied to the PWS lesion (Section 23.3.3).

23.3.3 Dual-Wavelength Irradiation

Sturesson and Andersson-Engels [111] have proposed modifying the skin temperature profile prior to delivery of selectively absorbed laser pulses by using a combination of active surface cooling and non-selective ("bulk") laser heating. In their theoretical analysis, pre-irradiating the skin with deeply penetrating light at 1064 nm in conjunction with surface cooling induced the maximum dermal temperature 1 mm below the epidermal-dermal junction (EDJ), with minimal epidermal heating. Because the temperature rise required to reach the coagulation threshold was reduced, a subsequent pulse at 532 nm was predicted to selectively treat blood vessels down to a depth of 0.8 mm below the EDJ, in comparison with 0.5 mm for water cooling alone. This study suggested that dual-wavelength treatment of PWS could compete in effectiveness and safety with the PDL at 577 nm.

Barton et al. [106] proposed a different two-stage laser photocoagulation technique, harnessing the dynamic changes of in vivo blood optical properties (Section 23.3.2). In this approach, selectively absorbed light (e.g., $\lambda = 532$ nm) is applied first to induce heating and initiate blood coagulation. Once the appropriate Arrhenius conditions for met-Hb generation have been met, leading to a significant increase in near-IR absorption inside the blood vessels, the 1064 nm laser pulse is applied.

The attractiveness of these approaches is that the required radiant exposure at both wavelengths can be maintained below their respective therapeutic thresholds. Inasmuch as the two wavelengths have different risk profiles (epidermal damage vs. subcutaneous or full-thickness dermal injury), dual-wavelength treatment should lead to reduced patient discomfort and fewer adverse side effects. The advantage is clearly larger with the second approach, where the near-IR absorption is increased by a factor of 3–4 by met-Hb formation.

Since met-Hb generation during laser exposure can be very rapid ($\sim$1 ms), there is little benefit to waiting between the first and the second pulse. From an engineering standpoint, the simplest solution is to apply a Nd:YAG/KTP laser that emits simultaneously 1064 and 532 nm pulses longer than a few ms. The 1064 nm light should help induce a higher temperature rise in blood vessels relative to the epidermis as compared to irradiation at 532 nm alone, due to combined effect of bulk dermal heating and met-Hb formation.

In a pilot clinical study, all subjects treated with such a dual-wavelength laser system achieved good PWS blanching after a single therapeutic session with a mean 532 nm radiant exposure of 8.2 J/cm^2 and CSC [112]. The response was very similar to that observed after treatment with a conventional KTP laser at $H = 12.4$ J/cm^2 with identical pulse structure ($t_p = 25$ ms). The acute appearance of a brownish skin discoloration, indicative of met-Hb formation, was observed in areas treated with the dual-wavelength laser, but not with the conventional KTP laser.

In numerical and animal studies by our group, we have explored the synergetic effects between the dual-wavelength technique discussed above and the MLP-MCS approach (Section 23.3.1.2) [94]. The numerical results suggest that even in patients with darker skin phototypes this approach could induce photocoagulation of blood

vessels with diameters up to 300 μm, but only when enhanced IR absorption of blood due to met-Hb formation was taken into account [94].

23.3.4 Optical Clearing

The concept of optical clearing involves active stimulation of optical changes beneficial to laser therapy (or optical imaging and diagnostics) in a controlled manner. Because the penetration depth of visible and near-IR light in many soft tissues, including human skin, is limited by optical scattering, several research groups have investigated the possibility to safely and reversibly reduce optical scattering by application of appropriate chemical agents such as glucose, glycerol, propylene glycol, and cosmetic lotions. [113, 114].

Optical scattering in human skin results from its complex morphology combined with the large difference in the index of refraction (n) between tissue water or ground substance on one hand, and proteinaceous skin components on the other. The major contributor to optical scattering in human dermis is collagen ($n = 1.43$–1.53, depending on hydration) organized in fibrils and fibers. In the epidermis, melanin granules ($n = 1.6-1.7$), cell nuclei ($n = 1.38-1.41$) and mitochondria ($n = 1.35-1.36$) have significantly higher refractive indices as compared to extracellular fluid.

The main mechanism of optical clearing is improved index-matching between the various tissue components [115], For example, glycerol ($n = 1.47$) can replace skin water within minutes after injection, dramatically reducing the index mismatch with collagen and keratinocytes and allowing in vivo visualization of cutaneous blood vessels (Fig. 23.24) [116]. Scattering properties are restored when water is reabsorbed into the tissue, which can take up to 30 min. Bulk dehydration due to

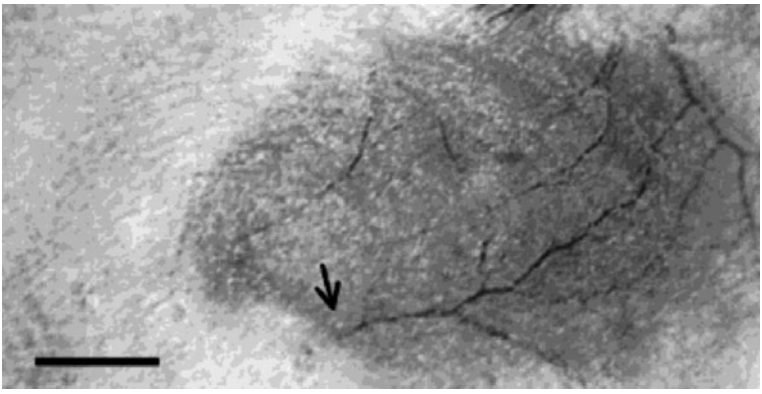


Fig. 23.24 Optical clearing induced in hamster skin in vivo by hypodermal injection of glycerol. Subsurface blood vessels are clearly visible in the affected area (*arrow*), unlike in strongly scattering native skin. Size bar: 1 cm [116]

application of an osmotic agent, which leads to denser and more regular packing of tissue components [117], and structural alteration of collagen fibers [118] have also been identified as additional mechanisms inducing optical clearing of skin.

Reduced optical scattering is expected to benefit PWS laser therapy through the increased penetration depth of visible laser radiation in skin, enabling effective treatment of deeper blood vessels [119]. In addition, reduced backscatter from the dermis should increase the epidermal damage threshold, allowing safe application of higher radiant exposures [120]. Optical clearing might also enable focusing of a laser beam onto subsurface target vessels for improved localization of energy deposition (see Fig. 23.24). Further studies appear warranted to determine how optical clearing of human skin might be most beneficial for PWS therapy. Some of the remaining challenges involve limited penetration of clearing agents through superficial skin layers and the incompletely understood impact on skin cells and extracellular matrix.

23.4 Outlook

23.4.1 Optimization of Laser Treatment Parameters on an Individual Patient Basis

As discussed above (Section 23.2.3.4), irradiation at a given wavelength and pulse duration affects preferentially PWS blood vessels of certain diameter and depth. For optimal clinical efficacy, the treatment parameters should ideally be selected based on knowledge of these and other structural and optical characteristics of each particular PWS lesion.

Moreover, the maximal radiant exposure that can be safely applied during PWS laser therapy is often limited by threshold for epidermal injury. Even the rather rudimentary relation (Eq. (23.12)) indicates that the latter varies with several factors. First, H_d obviously depends on the local melanin concentration and its absorption coefficient at the laser wavelength, but varies also with the fluence level at the basal layer. The last, in turn, depends on blood volumetric fraction, vessel size and depth distribution, epidermal thickness, as well as the scattering properties of both the epidermis and dermis. Secondly, ΔT_c depends on specifics of the cooling technology (reflected primarily in parameters h and T_c), the selected cooling time (t_c), and epidermal thickness at the treatment site (Figs. 23.15 and 23.19). Finally, due to kinetics of protein denaturation (see Chapter 13), the threshold peak temperature for the onset of epidermal injury varies with the laser pulse duration (t_p).

The main obstacles to optimizing PWS laser therapy are the complexity and heterogeneity of the vasculature, compounded by the large variation of relevant anatomical and optical characteristics between different patients and even between different anatomic locations on the same patient. Moreover, the sizes and depths of PWS capillaries appear to change as treatment progresses, due to selective vessel clearance [14, 23, 31, 121]. Therefore, one must exercise great caution when transferring therapeutically efficient as well as maximal safe radiant exposure values

between lasers systems with different wavelengths, pulse durations, active cooling devices, etc., as well as between patients or between different locations in the same patient.

This is corroborated by apparently conflicting results of several clinical studies. For example, Raulin et al. [55] reported PDL ($\lambda = 585$ nm) to be more efficient in infants than the KTP laser, but the latter produced blanching of PWS resistant to PDL treatment [60]. Similarly, although 585 nm was found to be the most effective PDL wavelength overall [56–58], irradiation at 595 nm ($t_p = 1.5$ ms) induced further improvement in a majority of PWS lesions that had stopped responding to treatment with 585 nm and $t_p = 0.5$ ms [31, 61]. In addition, PDLs with a longer pulse duration ($t_p = 1.5$ ms) were found effective for treating some PWS resistant to treatment at $t_p = 0.5$ ms) [68], although even shorter pulses appeared beneficial for a special group of PWS lesions resistant to similar treatment [14].

The combined evidence from theoretical, experimental and clinical research strongly suggests that a significant improvement in PWS therapy efficacy could be achieved by optimizing laser treatment parameters on an individual patient basis. Based on the above discussion, this would entail determination of blood vessel size and depth distribution, epidermal thickness and melanin concentration, and possibly also skin scattering properties and blood oxygenation level.

Unfortunately, existing non-invasive imaging techniques have various disadvantages for the purpose of PWS characterization, such as insufficient spatial resolution to resolve small blood vessels, long image acquisition times and prohibitive cost (all typical for magnetic resonance imaging), or limited contrast (ultrasound, optical coherence tomography – OCT). As a result, physicians select treatment parameters on the basis of very limited information (e.g., lesion appearance, skin type) or by testing several parameter combinations on small test patches. Extensive research efforts are under way to develop different techniques for noninvasive PWS characterization, as well as quantitative criteria for selection of optimal treatment parameters on an individual patient basis.

23.4.1.1 Diffuse Reflectance Spectroscopy and Colorimetry

Diffuse reflectance spectroscopy (DRS) applies white light irradiation to analyze diffusely reflected light from the skin surface (see Chapters 8 and 17). The reflectance spectrum contains information about the anatomic and optical properties of the measurement site, such as epidermal thickness, concentration of epidermal melanin and other chromophores, size and depth distribution of PWS blood vessels, and even epidermal and dermal scattering properties. Differences between the DRS measured on normal skin and untreated and treated PWS skin on the same patient are presented in Fig. 23.25a [43].

Determination of skin anatomic and optical properties requires solution of the DRS inverse problem which can be approached using the diffusion theory solutions for light propagation in a multi-layered skin model (see Fig. 23.25b). An alternative approach involves Monte Carlo (MC) computations of fluence rate distribution in model PWS skin which may contain discrete absorbing structures. Despite the

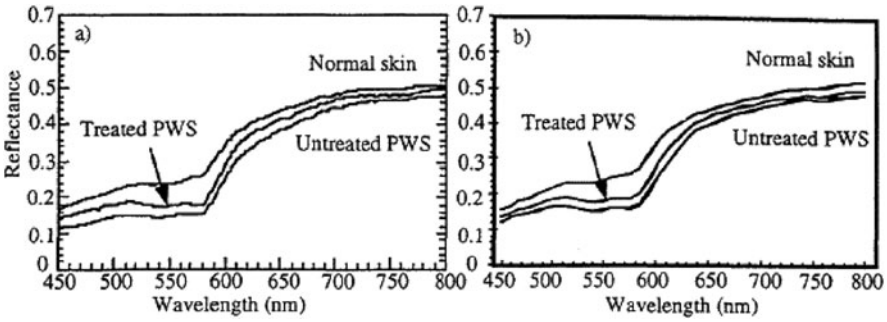


Fig. 23.25 (**a**) Measured and (**b**) matched simulated diffuse reflectance spectra in normal skin, reddish-purple PWS lesion before and 6 weeks after treatment [43]

recent progress in this area [95, 122], a general, robust solution for extraction of all clinically relevant PWS characteristics from DRS spectra has not been developed.

The DRS spectrum is very closely related to perceived PWS color. Using the multi-layered skin model approach, Svaasand et al. [123] concluded that PWS color varies systematically with vessel depth, from bright ($z < 0.4$ mm) to deep red ($z = 0.4{-}0.5$ mm) to bluish-purple ($z > 0.5$ mm). Histology showed that lesion color also correlated with average blood vessel diameter, ranging from 16 to 51 μm in pink and purple PWS lesions, respectively. Numerical predictions of lesion DRS and color based on MC simulations in a multi-layer model of PWS were provided by Verkruysse et al. [124].

In comparison with DRS, the advantages of colorimetry include less costly instrumentation and simpler data analysis. Thereby, it provides a practical tool for objective monitoring of PWS treatment efficacy, using either dedicated instrumentation [11, 125] or digital photography [126, 127]. Because object color is measured and expressed in 3D color space, colorimetry is practically limited to determination of one or two lesion properties, e.g., melanin and hemoglobin index [128]. Nevertheless, if clinical relevance of such information is confirmed, colorimetry might provide a practical approach to individualized optimization of PWS therapy.

23.4.1.2 Novel Optical Techniques for PWS Characterization

The earliest noninvasive investigations of PWS blood vessel size and location used transcutaneous microscopy [129, 130]. Recently, Sivarajan et al. [131] developed a dedicated depth-measurement videomicroscope (DMV) for assessment of PWS blood vessel diameter and depth in vivo.

To extend the depth to which highly scattering human skin can be imaged, optical coherence tomography (OCT) utilizes special light sources with longer wavelengths (typically around 800 or 1300 nm) and interferometric analysis of backscattered light. Recently developed optical Doppler tomography (ODT) additionally explores

characteristics of light scattering from moving RBC to provide enhanced contrast and enable measurement of blood flow velocity [132, 133]. Recent advances in ODT technology have allowed imaging of PWS birthmarks in vivo with spatial resolution of 10 μm and blood flow velocity sensitivity of ∼10 μm/s [134–136].

Pulsed photothermal radiometry (PPTR) measures transient changes in mid-IR radiation emitted from a sample surface following pulsed laser exposure. The laser-induced temperature depth profile in strongly scattering biological tissue can be reconstructed by solving the inverse problem of heat diffusion and radiative emission [137]. Of particular interest is the possibility of 3D imaging of PWS vascular structure by recording IR emission from the skin surface with high spatial resolution [138]. Sequential irradiation with two PDL wavelengths allows separation of the epidermal and vascular contributions to the temperature rise, enabling location and imaging of PWS blood vessels in close proximity to epidermis [50, 139–141].

While PPTR axial resolution appears limited (spreading of a thin absorbing layer amounts to ∼ 15% of its depth) [142], an important advantage is that it can provide – in addition to relevant structural information – absolute values of laser-induced temperature rises (see Fig. 23.23). This enables objective assessment of the required therapeutic radiant exposure and epidermal damage threshold, which are key to efficacy and safety of PWS laser therapy.

The same information can also be obtained by photoacoustic depth profiling and imaging, which combines high optical contrast (e.g., absorption in blood vs. dermis) with linear propagation, low scattering and moderate attenuation of ultrasonic waves in tissue [143, 144]. Using this approach, in vivo depth profiling of PWS [145] and cross-sectional imaging of subsurface blood vessels [146] were recently demonstrated.

Laser speckle imaging is analogous to laser Doppler imaging, but features substantially faster acquisition speeds. It can provide intraoperative evaluation of skin perfusion dynamics, which enables re-treatment of persistent areas in the same therapeutic session [48].

23.4.2 Real-Time Guidance of Therapy Using Feedback Signals

As an extension of therapy planning on an individual patient basis, Sebern et al. [147] developed a semi-autonomous laser device (called the “Smart scalpel”), which employs DRS as part of a line scan imaging system to identify and selectively target PWS blood vessels for treatment with a focused laser beam.

An alternative approach to (or augmentation thereof) such individualized therapy planning could be real-time monitoring of the effects of laser irradiation. Detection of either the desired endpoints (Section 23.1.3) or onset of some adverse effect could be used in a feedback loop to terminate the exposure or adjust treatment parameters for subsequent laser pulses. For example, increased backscattering or various spectroscopic signatures of met-Hb formation could be used to indicate that the endpoint of blood coagulation has been reached [21, 95, 106].

Alternatively, Suthamjariya et al. [21] proposed detection of the cavitation onset in treated vessels, which would automatically stop laser irradiation. A sudden increase of optical reflection or acoustic emission could be used as the feedback signal. Furthermore, light source intensity and pulse duration could be adjusted such that the required exposure is delivered in approximately one thermal relaxation time of the target vessels, τ_v (Eq. (23.5)).

23.4.3 Photodynamic Therapy

Photodynamic therapy (PDT) combines application of an exogenous chromophore, which allows tailoring of absorption properties for specific light sources, with optical activation of the drug for further localization of its biochemical effect [148]. The photosensitizing drug is administered first to the patient. After a suitable time interval, the tissue-localized drug is irradiated at a wavelength selectively absorbed by the photosensitizer. Excited molecules subsequently react with a substrate to generate highly reactive short-lived species, such as singlet oxygen, which cause irreversible oxidative damage to intracellular structures.

A laser or noncoherent light source provides low-power irradiation at a wavelength suitable to drive the photochemical reaction with insignificant heat generation. Milliwatt power levels avoid epidermal thermal injury while long exposure times (up to several minutes) ensure accumulation of sufficient dose in the targeted structure [149]. This property contrasts with photothermal therapy, where the required temperature rise must be achieved with millisecond laser exposure.

Animal model studies at our institution have demonstrated that PDT with benzoporphyrin derivative monoacid (BPD) and yellow light can achieve selective vascular destruction [150, 151]. BPD is a second-generation PDT drug that remains confined to the vascular compartment and has a strong absorption at 576 and 690 nm, making it very suitable for treatment of hypervascular skin lesions, including PWS. Moreover, rapid metabolic clearance of BDP reduces the duration of skin photosensitivity, the main disadvantage of PDT drugs that require systemic application.

Unlike photocoagulation, which tends to spare smaller blood vessels (diameter $d < 20\ \mu m$; see Fig. 23.14), PDT destroys all vessels containing photosensitizer. This can be an advantage, but also calls for careful treatment planning, because complete destruction of a dermal vascular network could potentially lead to skin ulceration, necrosis and permanent scarring/dyspigmentation.

A treatment protocol which combines the benefits of photothermal therapy and PDT was recently designed by our group [150]. Treatment was initiated with subtherapeutic PDT exposure at 576 nm, causing initial vascular damage. PDL irradiation was then applied to heat selectively the vessels compromised by PDT. An enhanced treatment effect compared with either PDT or PDL alone was demonstrated in two animal models [151, 152]. Further studies are required to fully determine the role of PDT, alone or in conjunction with other regimens, in the clinical management of PWS patients.

References

1. Mulliken JB and Young AR. *Vascular birthmarks – hemangiomas and malformations*. WB Saunders, Philadelphia, PA (1988).
2. Jacobs AH and Walton RG. The incidence of birthmarks in the neonate. *Pediatrics*, 58:218–222 (1976).
3. Heller A, Rafman S, Zvagulis I, and Pless IB. Birth defects and psychosocial adjustment. *Am. J. Dis. Child*., 139:257–263 (1985).
4. Malm M and Carlberg M. Port-wine stain – a surgical and psychological problem. *Ann. Plastic Surg*., 20:512–516 (1988).
5. Barsky SH, Rosen S, Geer DE, and Noe JM. The nature and evolution of port wine stains: A computer assisted study, *J. Invest. Dermatol*., 74:154–157 (1980).
6. Smoller BR and Rosen S. Port wine stains: A disease of altered neural modulation of blood vessels? *Arch. Dermatol*., 122:177–179 (1986).
7. Cosman B. Experience in the argon laser therapy of port wine stains. *Plast. Reconstr. Surg*., 65:119–129 (1980).
8. Noe JM, Barsky SH, Geer DE, and Rosen S. Port wine stains and the response to argon laser therapy: Successful treatment and the predictive role of color, age, and biopsy, *Plast. Reconstr. Surg*., 65:130–136 (1980).
9. Anderson RR and Parrish JA. Selective photothermolysis – precise microsurgery by selective absorption of pulsed radiation. *Science*, 220:524–527 (1983).
10. Katugampola GA and Lanigan SW. Five years′ experience of treating port wine stains with the flashlamp-pumped pulsed dye laser. *Br. J. Dermatol*., 137:750–754 (1997).
11. van der Horst CMAM, Koster PHL, de Borgie CAJM, Bossuyt PMM, and van Gemert MJC. Effect of the timing of treatment of port-wine stains with the flash-lamp-pumped pulsed dye-laser. *New Eng. J. Med*., 338:1028–1033 (1998).
12. Geronemus RG and Ashinoff R. The medical necessity of evaluation and treatment of port-wine stains. *J. Dermatol. Surg. Oncol*., 17:76–79 (1991).
13. Selim MM, Kelly KM, Nelson JS, Wendelschafer-Crabb G, Kennedy WR, and Zelickson BD. Confocal microscopy study of nerves and blood vessels in untreated and treated port wine stains: Preliminary observations. *Dermatol. Surg*., 30:892–897 (2004).
14. Fiskerstrand EJ, Svaasand LO, Kopstad G, Dalakar M, Norvang LT, and Volden G. Laser treatment of port-wine stains: Therapeutic outcome in relation to morphological parameters, *Br. J. Dermatol*., 134:1039–1043 (1996).
15. Huikeshoven M, Koster PHL, de Borgie CAJM, Beek JF, van Gemert MJC, and van der Horst CMAM. Redarkening of port-wine stains 10 years after pulsed-dye-laser treatment. *New Eng. J. Med*., 356:1235–1240 (2007).
16. Chapas AM, Eickhorst K, and Geronemus RG. Efficacy of early treatment of facial port wine stains in newborns: A review of 49 cases, *Lasers Surg. Med*., 39:563–568 (2007).
17. Renfro L and Geronemus RG. Anatomical differences in the treatment of port wine stains with the pulsed dye laser. *Arch Dermatol*., 29:182–188 (1993).
18. Anderson RR and Parrish JA. Microvasculature can be selectively damaged using dye lasers: A basic theory and experimental evidence in human skin. *Laser Surg Med*., 1:263–276 (1981).
19. Hulsbergen Henning JP, van Gemert MJC, and Lahaye CTW. Clinical and histological evaluation of portwine stain treatment with a microsecond-pulsed dye-laser at 577 nm. *Laser Surg Med*., 4:375–380 (1984).
20. Barton JK, Hammer DX, Pfefer TJ, Lund DJ, Stuck BE, and Welch AJ. Simultaneous irradiation and imaging of blood vessels during pulsed laser delivery. *Laser Surg Med*., 24:236–243 (1999).
21. Suthamjariya K, Farinelli WA, Koh W, and Anderson RR. Mechanisms of microvascular response to laser pulses. *J. Invest. Derm*., 122:518–525 (2004).
22. Arndt KA. Treatment techniques in argon-laser therapy – comparison of pulsed and continuous exposures. *J. Am. Acad. Dermatol*., 11:90–97 (1984).

23. Hohenleutner U, Hilbert M, Wlotzke U, and Landthaler M. Epidermal damage and limited coagulation depth with the flashlamp-pumped pulsed dye laser: A histochemical study. *J. Invest. Dermatol*., 104:798–802 (1995).
24. Birngruber R, Hillenkamp F, and Gabel VP. Theoretical investigations of laser thermal retinal injury. *Health Phys*., 48:781–796 (1985).
25. McKenzie AL. Physics of thermal processes in laser-tissue interaction. *Phys. Med. Biol*., 35:1175–1209 (1990).
26. Anderson RR and Parrish JA. The optics of human skin. *J. Invest. Dermatol*., 77:13–19 (1981).
27. van Gemert MJC, Jacques SL, Sterenborg HJCM, and Star WM. Skin optics, *IEEE Trans. Biomed. Eng*., 36:1146–1154 (1989).
28. Keijzer M, Pickering JW, and vanGemert MJC, Laser-beam diameter for port wine stain treatment. *Laser Surg. Med*., 11:601–605 (1991).
29. Tan OT, Motomedi M, Welch AJ, and Kurban AK. Spotsize effects on guinea pig skin following pulsed irradiation. *J. Invest. Dermatol*., 90:877–81 (1988).
30. Smithies DJ and Butler PH. Modelling the distribution of laser light in port-wine stains with the Monte Carlo method. *Phys. Med. Biol*., 40:701–731 (1995).
31. Sivarajan V, Maclaren WM, and Mackay IR. The effect of varying pulse duration, wavelength, spot size, and fluence on the response of previously treated capillary vascular malformations to pulsed-dye laser treatment. *Ann. Plast. Surg*., 57:25–32 (2006).
32. Kimel S, Svaasand LO, Cao D, Hammer-Wilson MJ, and Nelson JS. Vascular response to laser photothermolysis as a function of pulse duration, vessel type, and diameter: Implications for port wine stain laser therapy. *Laser Surg. Med*., 30:160 –169 (2002).
33. Jacques SL. Melanosome absorption coefficient. On: http://omlc.ogi.edu/spectra/melanin/mua.html, 2003.
34. Saidi IS, Jacques SL, and Tittel FK. Mie and Rayleigh modeling of visible-light scattering in neonatal skin. *Appl. Opt*., 34:7410–7418 (1995).
35. Kampen van EJ, and Zijlstra WG. Determination of hemoglobin and its derivatives. In: H Sobotka, CP Stewart (eds) *Advances in clinical chemistry*. Academic, New York, pp. 158–187 (1965).
36. van Gemert MJC, Welch AJ, Miller ID, and Tan OT. Can physical modeling lead to an optimal laser treatment strategy for port wine stains? In: ML Wolbrasht (ed) *Laser applications in medicine and biology*, Vol. 5. Plenum, New York, pp. 199–275 (1991).
37. Kuenster JT and Norris KH. Spectrophotometry of human hemoglobin in the near-infrared region from 1000 to 2500 nm. *J. Near Infrared Spectrosc*., 2:59–65 (1994).
38. Roggan A, Friebel M, Dorschel K, Hahn A, and Muller G. Optical properties of circulating human blood in the wavelength range 400–2500 nm. *J. Biomed. Opt*., 4:36–46 (1999).
39. Faber DJ, Aalders MCG, Mik EG, Hooper BA, van Gemert MJC, and van Leeuwen TG. Oxygen saturation-dependent absorption and scattering of blood. *Phys. Rev. Lett*., 93:028102 (2004).
40. Tan OT, Morrison P, and Kurban AK., 585 nm for the treatment of port-wine stains, *Plast. Reconstr. Surg*., 86:1112–1117 (1990).
41. Lucassen GW, Verkruysse W, Keijzer M, and van Gemert MJC. Light distributions in a port wine stain model containing multiple cylindrical and curved blood vessels. *Laser Surg. Med*., 18:345–357 (1996).
42. Verkruysse W, Lucassen GW, de Boer JF, Smithies DJ, Nelson JS, and van Gemert MJC. Modelling light distributions of homogeneous versus discrete absorbers in light irradiated turbid media. *Phys. Med. Biol*., 42:51–65 (1997).
43. van Gemert MJC, Nelson JS, Milner TE, Smithies DJ, Verkruysse W, de Boer JF, Lucassen GW, Goodman DM, Tanenbaum BS, Norvang LT, and Svaasand LO. Non-invasive determination of port wine stain anatomy and physiology for optimal laser treatment strategies. *Phys. Med. Biol*., 42:937–950 (1997).

44. Kimel S, Svaasand LO, Hammer-Wilson M, Schell MJ, Milner TE, Nelson JS, and Berns MW. Differential vascular response to laser photothermolysis. *J. Invest. Dermatol*., 103:693–708 (1994).
45. Svaasand LO, Fiskerstrand EJ, Kopstad G, Norvang LT, Svaasand EK, Nelson JS, and Berns MW. Therapeutic response during pulsed laser treatment of port-wine stains: Dependence on vessel diameter and depth in dermis. *Laser Med. Sci*., 10:235–243 (1995).
46. Nelson JS, Milner TE, Svaasand LO, and Kimel S. Laser pulse duration must match the estimated thermal relaxation time for successful photothermolysis of blood vessels. *Laser Med. Sci*., 10:9–12 (1995).
47. Kienle A and Hibst R. A new optimal wavelength for treatment of port wine stains? *Phys. Med. Biol*., 40:1559–1576 (1995).
48. Choi B, Kang NM, and Nelson JS. Laser speckle imaging for monitoring blood flow dynamics in the in vivo rodent dorsal skin fold model. *Microvasc. Res*., 68:143–146 (2004).
49. Majaron B, Milanič M, and Nelson JS. Interaction of a dual-wavelength laser system with cutaneous blood vessels. In: A Vogel (ed) *Therapeutic Laser Applications and Laser-Tissue Interactions III, Proc. SPIE*, Vol. 6632, 66320C. Bellingham, WA (2007).
50. Choi B, Majaron B, and Nelson JS. Computational model to evaluate port wine stain depth profiling using pulsed photothermal radiometry. *J. Biomed. Opt*., 9:299–307 (2004).
51. van Gemert MJC, Smithies DJ, Verkruysse W, Milner TE, and Nelson JS. Wavelengths for port wine stain laser treatment: Influence of vessel radius and skin anatomy. *Phys. Med. Biol*., 42:41–50 (1997).
52. Smithies DJ, van Gemert MJC, Hansen MK, Milner TE, and Nelson JS. Three-dimensional reconstruction of port wine stain vascular anatomy from serial histological sections. *Phys Med Biol*., 42:1843–1847 (1997).
53. Barton JK, Pfefer TJ, Welch AJ, Smithies DJ, Nelson JS, and van Gemert MJC. Optical Monte Carlo modeling of a true port wine stain anatomy. *Opt. Express*, 2:391–396 (1998).
54. Pfefer TJ, Barton JK, Smithies DJ, Milner TE, Nelson JS, van Gemert MJC, and Welch AJ. Modeling laser treatment of port wine stains with a computer-reconstructed biopsy. *Laser Surg. Med*., 24:151–166 (1999).
55. Raulin C and Greve B. Retrospective clinical comparison of hemangioma treatment by flashlamp-pumped (585 nm) and frequency-doubled Nd:YAG (532 nm) lasers. *Laser Surg. Med*., 28:40–43 (2001).
56. Chang CJ, Kelly KM, van Gemert MJC, and Nelson JS. Comparing the effectiveness of 585-nm vs., 595-nm wavelength pulsed dye laser treatment of port wine stains in conjunction with cryogen spray cooling. *Lasers Surg. Med*., 31:352–358 (2002).
57. Greve B and Raulin C. Prospective study of port wine stain treatment with dye laser: Comparison of two wavelengths (585 nm vs 595 nm) and two pulse durations (0.5 milliseconds vs. 20 milliseconds). *Laser Surg. Med*., 34:168–173 (2004).
58. Scherer K, Lorenz S, Wimmershoff M, Landthaler M, and Hohenleutneret U. Both the flashlamp-pumped dye laser and the long-pulsed tunable dye laser can improve results in port-wine stain therapy. *Br. J. Dermatol*., 145:79–84 (2001).
59. Kimel S, Svaasand LO, Hammer-Wilson MJ, and Nelson JS. Influence of wavelength on response to laser photothermolysis of blood vessels: Implications for port wine stain laser therapy. *Laser Surg. Med*., 33:288–295 (2003).
60. Chowdhury MM, Harris S, and Lanigan SW. Potassium titanyl phosphate laser treatment of resistant portwine stains. *Br. J. Dermatol*., 144:814–817 (2001).
61. Laube S, Taibjee S, and Lanigan SW. Treatment of resistant port wine stains with the V Beam pulsed dye laser. *Laser Surg. Med*., 33:282–287 (2003).
62. Nakagawa H, Tan OT, and Parrish JA. Ultrastructure changes in human skin after exposure to a pulsed laser. *J. Invest. Dermatol*., 84:396–400 (1985).
63. Dierickx CC, Casparian JM, Venugopalan V, Farinelli WA, and Anderson RR. Thermal relaxation of port-wine stain vessels probed in vivo: The need for 1–10 millisecond laser pulse treatment. *J. Invest. Dermatol*., 105:709–714 (1995).

64. Svaasand LO, Milner TE, Anvari B, Norvang LT, Tanenbaum BS, Kimel S, Berns MW, and Nelson JS. Epidermal heating during laser induced photothermolysis of port wine stains: Modeling melanosomal heating after dynamic cooling the skin surface. In: *Laser interaction with hard and soft tissue II*, Proc. SPIE, Vol. 2323, pp.366–377 (1995).
65. deBoer JF, Lucassen GW, Verkruysse W, and van Gemert MJC. Thermolysis of port-wine-stain blood vessels: Diameter of a damaged blood vessel depends on the laser pulse length, *Lasers. Med. Sci.*, 11:177–180 (1996).
66. Pickering JW and van Gemert MJC. 585 nm for the laser treatment of port wine stains: A possible mechanism. *Laser Surg. Med*., 11:616–618 (1991).
67. Svaasand LO, van Gemert MJC, Verkruysse W, Fiskerstrand EJ, and Norvang LT. Dosimetry for laser treatment of port-wine stains. In: SL Jacques, GJ Mueller, A Roggan, and DH Sliney (eds) *Laser-tissue interactionX: Photochemical, photothermal, and photomechanical*, Proc. SPIE, Vol. 601, pp. 463–471 (1999).
68. Bernstein EF. Treatment of a resistant port-wine stain with the 1.5-msec pulse duration, tunable, pulsed dye laser. *Dermatol. Surg*., 26:1007–1009 (2000).
69. Parlette EC, Groff WF, Kinshella MJ, Domankevitz Y, O'Neill J, and Ross EV. Optimal pulse durations for the treatment of leg telangiectasias with a neodymium YAG laser. *Laser Surg. Med*., 38:98–105 (2006).
70. Goldman MP and Fitzpatrick RE. Treatment of cutaneous vascular lesions. In: MP Goldman and RE Fitzpatrick (eds) *Cutaneous laser surgery*. Mosby, St. Louis, pp. 19–105 (1994).
71. Lanigan SW. Port-wine stains unresponsive to pulsed dye laser: Explanations and solutions. *Br. J. Dermatol*., 139:173–177 (1998).
72. Gilchrest BA, Rosen S, and Noe J. Chilling port wine stains improves the response to argon laser therapy. *Plast. Reconstr. Surg*., 69:278–283 (1982).
73. Welch AJ, Motamedi M, and Gonzalez A. Evaluation of cooling techniques for the protection of the epidermis during Nd:YAG laser irradiation of the skin. In: SN Joffe (ed) *Neodymium:YAG laser in medicine and surgery*. Elsevier, New York, pp. 196–204 (1983).
74. Nelson JS, Majaron B, and Kelly KM. Active skin cooling in conjunction with laser dermatologic surgery: Methodology and clinical results. *Semin. Cutan. Med. Surg*., 19:253–266 (2000).
75. Carslaw HS and Jaeger JC. *Conduction of heat in solids* Clarendon, Oxford, 2nd Edition (1959).
76. Nelson JS, Milner TE, Anvari B, Tanenbaum BS, Kimel S, Svaasand LO, and Jacques SL. Dynamic epidermal cooling during pulsed laser treatment of port-wine stain. *Arch. Dermatol*., 131:695–700 (1995).
77. Nelson JS, Milner TE, Anvari B, Tanenbaum BS, Svaasand LO, and Kimel S. Dynamic epidermal cooling in conjunction with laser-induced photothermolysis of port wine stain blood vessels. *Laser Surg. Med*., 19:224–229 (1996).
78. Anvari B, Milner TE, Tanenbaum BS, and Nelson JS. A comparative study of human skin thermal response to sapphire contact and cryogen spray cooling. *IEEE Trans. Biomed. Eng*., 45:934–941 (1998).
79. Verkruysse W, Majaron B, Aguilar G, Svaasand LO, and Nelson JS. Dynamics of cryogen deposition relative to heat extraction rate during cryogen spray cooling. In: *Lasers in surgery: Advanced characterization, therapeutics, and systems X*, Proc. SPIE, Vol. 3907, pp. 37–58, Bellingham (2000).
80. Aguilar G, Verkruysse W, Majaron B, Svaasand LO, Lavernia EJ, and Nelson JS. Measurement of heat transfer coefficient during continuous cryogen spray cooling. *IEEE J. Sel. Top. Quant. Electr*., 7:1013–1021 (2001).
81. Aguilar G, Wang GX, and Nelson JS. Effect of spurt duration on the heat transfer dynamics during cryogen spray cooling. *Phys. Med. Biol*., 48:2169–2181 (2003).
82. Svaasand LO, Randeberg LL, Aguilar G, Majaron B, Kimel S, Lavernia EJ, and Nelson JS. Cooling efficiency of cryogen spray during laser therapy of skin. *Laser Surg. Med*., 32:137–142 (2003).

83. Majaron B, Svaasand LO, Aguilar G, and Nelson JS. Intermittent cryogen spray cooling for optimal heat extraction during dermatologic laser treatment, *Phys. Med. Biol*., 47:3275–3288 (2002).
84. Aguilar G, Majaron B, Pope K, Svaasand LO, Nelson JS, and Lavernia EJ. Influence of nozzle-to-skin distance in cryogen spray cooling for dermatologic laser surgery. *Laser Surg. Med*., 28:113–120 (2001).
85. Edris A, Choi B, Aguilar G, and Nelson JS. Measurements of laser light attenuation following cryogen spray cooling spurt termination. *Laser Surg. Med*., 32:143–147 (2003).
86. Majaron B, Kimel S, Verkruysse W, Aguilar G, Pope K, Svaasand LO, Lavernia EJ, and Nelson JS. Cryogen spray cooling in laser dermatology: Effects of ambient humidity and frost formation. *Laser Surg. Med*., 28:469–476 (2001).
87. Anvari B, Milner TE, Tanenbaum BS, Kimel S, and Nelson JS. Selective cooling of biological tissues: Application for thermally mediated therapeutic procedures. *Phys. Med. Biol*., 40:241–252 (1995).
88. Altshuler GB, Zenzie HH, Erofeev AV, Smirnov MZ, Anderson RR, and Dierickx C. Contact cooling of the skin. *Phys. Med. Biol*., 44:1003–1023 (1999).
89. Pope K and Lask G. *Epidermal temperature evaluation during dynamic spray cooling, contact cooling, and ice*. Presented at the 20th Annual meeting of the ASLMS, Reno, NV, April (2000).
90. Verkruysse W, Majaron B, Tanenbaum BS, and Nelson JS. Optimal cryogen spray cooling parameters for pulsed laser treatment of port wine stains. *Laser Surg. Med*., 27:165–170 (2000).
91. Chang CJ and Nelson JS. Cryogen spray cooling and higher fluence pulsed dye laser treatment improve port wine stain clearance while minimizing epidermal damage. *Dermatol. Surg*., 25:767–772 (1999).
92. Waldorf HA, Alster TS, McMillan K, Kauvar ANB, Geronemus RG, and Nelson JS. Effect of dynamic cooling on 585 nm pulsed dye laser treatment of port wine stain birthmarks. *Dermatol. Surg*., 23:657–662 (1997).
93. Lukač M, Majaron B, and Rupnik T. Ablative and thermal effects of Er:YAG laser on human tissue. In: R Waidelich et al. (eds) *Laser in der medizin/in medicine*, Proc., 13th Internat. Congress LASER 97. Springer, Berlin, pp. 566–572 (1998).
94. Jia W, Choi B, Franco W, Lotfi J, Majaron B, Aguilar G, and Nelson JS. Treatment of cutaneous vascular lesions using multiple-intermittent cryogen spurts and two-wavelength laser pulses: Numerical and animal studies. *Laser Surg. Med*., 39:494–503 (2007).
95. Randeberg LL, Bonesronning JH, Dalaker M, Nelson JS, and Svaasand LO, Methemoglobin formation during laser induced photothermolysis of vascular skin lesions. *Laser Surg. Med*., 34:414–419 (2004).
96. Tanghetti E, Sierra RA, Sherr EA, and Mirkov M. Evaluation of pulse-duration on purpuric threshold using extended pulse pulsed dye laser (Cynosure V-Star). *Laser Surg. Med*., 31:363–366 (2002).
97. Tanghetti EA, Sherr EA, and Alvarado SL. Multipass treatment of photodamage using the pulse dye laser. *Dermatol. Surg*., 29:686–691 (2003).
98. Rohrer TE, Chatrath V, and Iyengar V. Does pulse stacking improve the results of treatment with variable-pulse pulsed dye lasers? *Dermatol. Surg*., 30:163–167 (2004).
99. Ross EV and Domankevitz Y. Laser leg vein treatments: A brief overview. *J. Cosmetic Laser Ther*., 5:192–197 (2003).
100. Bencini PL. The multilayer technique: A new and fast approach for flashlamp-pumped pulsed (FLPP) dye laser treatment of port-wine stains. *Dermatol. Surg*., 25:786–789 (1999).
101. Black JF and Barton JK. Chemical and structural changes in blood undergoing laser photocoagulation. *Photochem. Photobiol*., 80:89–97 (2004).
102. Verkruysse W, Nilsson AMK, Milner TE, Beek JF, Lucassen GW, and van Gemert MJC. Optical absorption of blood depends on temperature during a 0.5 ms laser pulse at 586 nm. *Photochem. Photobiol*., 67:276–281 (1998).

103. Nilsson AMK, Lucassen GW, Verkryusse W, Andersson-Engels S, and van Gemert MJC. Changes in optical properties of human whole blood in vitro due to slow heating. *Photochem. Photobiol*., 65:366–373 (1997).
104. Baranov VY, Chekhov DI, Leonov AG, Leonov PG, Ryaboshapka OM, Semenov SY, Splinter R, Svenson RH, and Tatsis GP. Heatpinduced changes in optical properties of human whole blood in vitro. In: *Optical diagnostics of biological fluids IV,* Proc. SPIE, Vol. 3599, pp.180–187 (1999).
105. Mordon S, Rochon P, Dhelin G, and Lesage JC. Dynamics of temperature dependent modifications of blood in the near-infrared. *Laser Surg. Med*., 37:301–307 (2005).
106. Barton JK, Frangineas G, Pummer H, and Black JF. Cooperative phenomena in two-pulse, two-color laser photocoagulation of cutaneous blood vessels. *Photochem. Photobiol*., 73:642–650 (2001).
107. Randeberg LL, Daae Hagen AJ, and Svaasand LO. Optical properties of human blood as a function of temperature. *Proc. SPIE*, 4609:20–29 (2002).
108. Black JF, Wade N, and Barton JK. Mechanistic comparison of blood undergoing laser photocoagulation at 532 and 1,064 nm. *Laser Surg. Med*., 36:155–165 (2005).
109. Kimel S, Choi B, Svaasand LO, Lotfi J, Viator JA, and Nelson JS. Influence of laser wavelength and pulse duration on gas bubble formation in blood filled glass capillaries. *Laser Surg. Med*., 36:281–288 (2005).
110. Pikkula BM, Chang DW, Nelson JS, and Anvari B. Comparison of 585 and 595 nm laser-induced vascular response of normal in vivo human skin. *Laser Surg. Med*., 36:117–123 (2005).
111. Sturesson C and Andersson-Engels S. Mathematical modelling of dynamic cooling and pre-heating, used to increase the depth of selective damage to blood vessels in laser treatment of port wine stains. *Phys. Med. Biol*., 41:413–428 (1996).
112. Ahčan U, Zorman P, Recek D, Ralca S, and Majaron B. Port wine stain treatment with a dual-wavelength Nd:YAG laser and cryogen spray cooling: A pilot study. *Laser Surg. Med*., 34:164–167 (2004).
113. Tuchin VV. Optical clearing of tissues and blood using the immersion method. *J. Phys. D: Appl. Phys*., 38:2497–2518 (2005).
114. Tuchin VV. *Optical clearing of tissues and blood.* SPIE, Bellingham, WA (2005).
115. Tuchin VV, Maksimova IL, Zimnyakov DA, Kon IL, Mavlutov AH, and Mishin AA. Light propagation in tissues with controlled optical properties. *J. Biomed. Opt*., 2:401–417 (1997).
116. Vargas G, Chan EK, Barton JK, Grady HG III, and Welch AJ. Use of an agent to reduce scattering in the skin. *Laser Surg. Med*., 24:133–141 (1999).
117. Rylander CG, Stumpp OF, Milner TE, Kemp NJ, Mendenhall JM, Diller KR, and Welch AJ. Dehydration mechanism of optical clearing in tissue. *J. Biomed. Opt*., 11:041117 (2006).
118. Yeh AT, Choi B, Nelson JS, and Tromberg BJ. Reversible dissociation of collagen in tissues. *J. Invest. Dermatol*., 121:1332–1335 (2003).
119. Khan MH, Choi B, Chess S, Kelly KM, McCullough JL, and Nelson JS. Optical clearing of in vivo human skin: Implications for light-based diagnostic imaging and therapeutics. *Laser Surg. Med*., 34:83–85 (2004).
120. Khan MH, Chess S, Choi B, Kelly KM, and Nelson JS. Can topically applied optical clearing agents increase the epidermal damage threshold and enhance therapeutic efficacy? *Laser Surg. Med*., 35:93–95 (2004).
121. Sivarajan V and Mackay IR. Noninvasive in vivo assessment of vessel characteristics in capillary vascular malformations exposed to five pulsed dye laser treatments, *Plast. Reconstr. Surg*., 115:1245–1252 (2005).
122. Zonios G, Bykowski J, and Kollias N. Skin melanin, hemoglobin, and light scattering properties can be quantitatively assessed in vivo using diffuse reflectance spectroscopy. *J. Invest. Dermatol*., 117:1452–1457 (2001).
123. Svaasand LO, Norvang LT, Fiskerstrand EJ, Stopps EKS, Berns MW, and Nelson JS. Tissue parameters determining the visual appearance of normal skin and port-wine stains. *Laser Med. Sci*., 10:55–65 (1995).

124. Verkruysse W, Lucassen GW, and van Gemert MJC. Simulation of color of port wine stain skin and its dependence on skin variables. *Laser Surg. Med*., 25:131–139 (1999).
125. Rah DK, Kim SC, Lee KH, Park BY, and Kim DW. Objective evaluation of treatment effects on port-wine stains using L∗a∗b∗ color coordinates. *Plast. Reconstr. Surg*., 108:842–847 (2001).
126. Setaro M and Sparavigna A. Quantification of erythema using digital camera and computer-based colour image analysis: A multicentre study. *Skin Res. Technol*., 8:84–88 (2002).
127. Jung B, Kim CS, Choi B, Kelly KM, and Nelson JS. Use of erythema index imaging for systematic analysis of port wine stain skin response to laser therapy. *Laser Surg. Med*., 37:186–191 (2005).
128. Takiwaki H, Miyaoka Y, Kohno H, and Arase S. Graphic analysis of the relationship between skin colour change and variations in the amounts of melanin and haemoglobin. *Skin Res. Technol*., 8:78–83 (2002).
129. Jones SG, Shakespeare PG, and Carruth JAS. Transcutaneous microscopy and argon laser treatment of port wine stains. *Lasers Med. Sci*., 4:73–78 (1989).
130. Motley R and Lanigan SW. Videomicroscopy predicts outcome in treatment of port-wine stains. *Arch. Dermatol*., 133:921–922 (1997).
131. Sivarajan V and Mackay IR. The depth measuring videomicroscope (DMV): A non-invasive tool for the assessment of capillary vascular malformations. *Laser Surg. Med*., 34:193–197 (2004).
132. Chen Z, Milner TE, Dave D, and Nelson JS. Optical Doppler tomographic imaging of fluid flow velocity in highly scattering media. *Opt. Lett*., 22:64 (1997).
133. Izatt JA, Kulkarni MD, Yazdanfar S, Barton JK, and Welch AJ. In vivo bidirectional color Doppler flow imaging of picoliter blood volumes using optical coherence tomography. *Opt. Lett*., 22:1439–1441 (1997).
134. Zhao YH, Chen ZP, Saxer C, Shen QM, Xiang SH, de Boer JF, and Nelson JS. Doppler standard deviation imaging for clinical monitoring of in vivo human skin blood flow. *Opt. Lett*., 25:1358–1360 (2000).
135. Nelson JS, Kelly KM, Zhao Y, and Chen Z. Imaging blood flow in human port-wine stain in situ and in real time using optical Doppler tomography. *Arch. Dermatol*., 137:741–744 (2001).
136. Milner TE, Yazdanfar S, Rollins AM, Izatt JA, Lindmo T, Chen ZP, Nelson JS, and Wang XJ. Doppler optical coherence tomography. In: BE Bouma and GJ Tearney (eds) *Handbook of optical coherence tomography*. Marcel Decker, New York (2002).
137. Milner TE, Goodman DM, Tanenbaum BS, and Nelson JS. Depth profiling of laser-heated chromophores in biological tissues by pulsed photothermal radiometry. *J. Opt. Soc. Am. A*, 12:1479–1488 (1995).
138. Telenkov SA, Tanenbaum BS, Goodman DM, Nelson JS, and Milner TE. In vivo infrared tomographic imaging of laser-heated blood vessels. *IEEE J. Sel. Top. Quant. Elect*., 5:1193–1199 (1999).
139. Majaron B, Verkruysse W, Tanenbaum BS, Milner TE, Telenkov SA, Goodman DM, and Nelson JS. Combining two excitation wavelengths for pulsed photothermal profiling of hypervascular lesions in human skin. *Phys. Med. Biol*., 45:1913–1922 (2000).
140. Choi B, Majaron B, Vargas G, Jung B, Stumpp OF, Kang NM, Kelly KM, Welch AJ, and Nelson JS. In vivo results using photothermal tomography for imaging cutaneous blood vessels. In: T Kundu (ed) *Smart nondestructive evaluation and health monitoring of structural and biological systems II*, Proc. SPIE, Vol. 5047, pp. 350–361 (2003).
141. Li B, Majaron B, Viator JA, Milner TE, Chen Z, Zhao Y, Ren H, and Nelson JS. Accurate measurement of blood vessel depth in port wine stained human skin in vivo using pulsed photothermal radiometry. *J. Biomed. Opt*., 9:961–966 (2004).
142. Milanič M, Serša I, and Majaron B, Spectrally composite reconstruction approach for improved resolution of pulsed photothermal temperature profiling in water-based samples. *Phys. Med. Biol*., 54:2829–2844 (2009).

143. Gusev VE and Karabutov AA. *Laser optoacoustics*, American Institute of Physics, New York (1993).
144. Wang LHV. Ultrasound-mediated biophotonic imaging: A review of acousto-optical tomography and photo-acoustic tomography. *Dis. Markers* 19:123–138 (2004).
145. Viator JA, Au G, Paltauf G, Jacques SL, Prahl SA, Ren H, Chen Z, and Nelson JS. Clinical testing of a photoacoustic probe for port wine stain depth determination. *Laser Surg. Med*., 30:141–148 (2002).
146. Kolkman RGM, Hondebrink E, Steenbergen W, and de Mul FFM. In vivo photoacoustic imaging of blood vessels using an extreme-narrow aperture sensor. *IEEE Sel. Top. Quant. Electr*., 9:343–346 (2003).
147. Sebern EL, Brenan CJ, and Hunter IW. Design and characterization of a laser-based instrument with spectroscopic feedback control for treatment of vascular lesions: The "Smart Scalpel". *J. Biomed. Opt*., 5:375–382 (2000).
148. Nelson JS, McCullough JL, and Berns MW. Principles and applications of photodynamic therapy in dermatology. In: KA Arndt, JE Dover, SA Olbricht (eds) *Lasers in cutaneous and aesthetic surgery*. Lippincott-Raven, Philadelphia, pp. 349–382 (1997).
149. Orenstein A, Nelson JS, Liaw LH, Kaplan R, Kimel S, and Berns MW. Photochemotherapy of hypervascular dermal lesions: A possible alternative to photothermal therapy? *Laser Surg. Med*., 10:334–343 (1990).
150. Kimel S, Svaasand LO, Kelly KM, and Nelson JS. Synergistic photodynamic and photothermal treatment of port-wine stain? *Laser Surg. Med*., 34:80–82 (2004).
151. Kelly KM, Kimel S, Smith T, Stacy A, Hammer-Wilson MJ, Svaasand LO, and Nelson JS. Combined photodynamic and photothermal induced injury enhances damage to in vivo model blood vessels. *Laser Surg. Med*., 34:407–413 (2004).
152. Smith TK, Choi B, Ramirez-San-Juan J, Nelson JS, Osann K, and Kelly KM. Microvascular blood flow dynamics associated with photodynamic therapy, pulsed dye laser irradiation and combined regimens. *Laser Surg. Med*., 38:532–539 (2006).

Chapter 24
Infrared Nerve Stimulation: A Novel Therapeutic Laser Modality

Jonathon D. Wells, Jonathan M. Cayce, Anita Mahadevan-Jansen, Peter E. Konrad, and E. Duco Jansen

24.1 Introduction

Neural stimulation is the process of activating neurons using an external source to evoke action potential propagation down an axon. Electrical, chemical, thermal, optical, and mechanical methods have all been reported to stimulate neurons in both the central nervous system (CNS) and the peripheral nervous system (PNS) [1]. For nearly 2 centuries electrical stimulation has been the gold standard for the stimulation of neurons and other excitable tissues. It functions by increasing the transmembrane potential to activate voltage-gated ion channels which induce action potential propagation down the axon of a neuron [2–5]. However, electrical stimulation lacks spatial precision due to the inherent electrical field propagation which results in the recruiting of multiple (unwanted) neuronal fibers. Additionally, electrical stimulation induces a stimulation artifact which can mask neuronal signals resulting from the simulation [6, 7].

Optical stimulation of neural tissue is a fundamentally new approach that has distinct advantages when compared to other neural excitation modalities, including standard electrical nerve stimulation. From a laser-tissue interaction and clinical laser application point of view, this application of laser technology falls in the therapeutic category although it may also be used as a means to provide guidance during surgical procedures, in which case it is more of a diagnostic or mapping tool. However, it is fundamentally different from most therapeutic laser applications in that the aim is not to permanently damage, coagulate or ablate the tissue. Instead, the objective is to induce a physiological response using laser light (and – as we will show later in this chapter – laser-induced temperature gradients). This novel

E.D. Jansen (✉)
Departments of Biomedical Engineering and Neurosurgery, Vanderbilt University, Nashville, TN 37235, USA
e-mail: duco.jansen@vanderbilt.edu

A.J. Welch, M.J.C. van Gemert (eds.), *Optical-Thermal Response of Laser-Irradiated Tissue*, 2nd ed., DOI 10.1007/978-90-481-8831-4_24,

methodology for nerve excitation relies on irradiation of the nerve surface, in a non-contact fashion, with a pulsed (infrared) laser operating at an optimized radiant exposure and wavelength for the generation of compound action potentials (CAP) and associated physiological effect (i.e. muscle contraction or sensory response). The response is spatially precise, providing the possibility for selective targeting of individual nerve fascicles, and does not result in tissue damage at laser radiant exposures at least two times that required for stimulation. Thus, optical stimulation presents an innovative approach to neural activation that may benefit clinical nerve stimulation as well as fundamental neurophysiology and neuroscience. Here we will define and characterize optical nerve stimulation as well as detail the methodology of this technique. The fundamental principles and experiments that help define the optimal parameters needed for safe and effective stimulation of the peripheral nerve will be described. The biophysical mechanism will be briefly discussed. Finally, the potential impact of optical stimulation is noted in terms of applications and clinical utility for this technology.

24.1.1 Limitations of Electrical Nerve Stimulation

The standard method for excitation of neural tissue relies on the injection of electrical current in excitable tissues. This method is used both clinically (in diagnostics and therapeutics), and for basic scientific research in electrophysiology and the neurosciences. Although this method is widely used, it has several fundamental limitations. Electrical stimulation of a nerve requires physical contact between a (metal) electrode and the tissue. Often the electrode is physically pierced into the tissue, which can give rise to tissue damage. Spatial precision of stimulation is limited due to the size of electrodes and, more importantly, the inherent induction of an electric field spanning a spatial area much greater than the size of the electrode. This spread of electrical current in a graded fashion beyond the electrode causes poor spatial specificity with this technique [8]. Many applications of neural stimulation require precision of the stimulus in a small target tissue. Electrodes designed to deliver precise stimulation have inherently high impedance characteristics which, in turn, impose higher voltage requirements to deliver the same charge as dictated by Ohm's law. In addition, in applications where electrophysiological recordings are performed to measure the response to stimulation, an unavoidable "stimulation artifact" is recorded owing to the fact that the stimulation technique occurs in the same domain as the recording technique. The result is a recorded signal that contains electrical artifact from the stimulus. Usually the amplitude of this artifact is much larger than the recorded action potential, and precise interpretation is often not possible, or requires significant processing in order to obtain meaningful results [6, 9]. These limitations in the current technique have driven researchers to pursue other means for neural stimulation. The literature thoroughly documents the notion that action potentials can be triggered in neurons using many different stimuli; these include electrical, magnetic, mechanical, thermal, chemical, and optical means. Here we will focus our attention on optical methods for the stimulation of neural tissue.

24.1.2 Definition of Optical Stimulation

We define *optical stimulation* as the direct induction of an evoked potential (EP) or action potential (AP) in response to a transient targeted deposition of optical energy. This implies that only a pulsed source can be used for stimulation of neural tissue, and that continuous wave (CW) irradiation will not lead to compound action potential generation. As will be discussed later in this chapter, we have preliminary evidence that the induction of a temperature gradient (dT/dz or dT/dt) is in fact required to induce an action potential.

This strict definition for optical stimulation clearly defines what is and is not considered optical stimulation. It also distinguishes optical stimulation from other related applications such as the modulation of the excitability of nerves using light [10–12] or from the use of low-level light therapy (LLLT) or biostimulation where low irradiance levels at laser wavelengths that are weakly absorbed in tissue are applied continuously for several minutes [13, 14].

24.1.3 Previous Work in Optical Stimulation

The first report of low-level, direct laser stimulation of neural tissue was published in 2005 [15]. Historically this was certainly not the first time light was used to stimulate neural tissue. Numerous reports can be found in the literature that use a plethora of ways to use light to interact with the neural system. Optical stimulation was first reported [16] as action potentials generated in *Aplysia* neurons through a reversible mechanism. This was the first indication that optical irradiance of nerve cells could indeed induce action potentials in neural tissue. Other studies reported stimulation in the spinal cord as a side effect to ablation using a short pulse, ultra-violet excimer laser [17]; and the induction of firing of pyramidal neurons during exposure to high intensity mode-locked infrared femtosecond laser [18]. Prior to our work described in the subsequent sections of this chapter, there have been no systematic studies published on the application of optical energy for neural activation.

24.2 Optical Stimulation in the Peripheral Nervous System (PNS)

The basis of this work is that delivery of pulsed laser light can be used for contact-free, damage-free, artifact-free stimulation of discrete populations of excitable cells. We have recently shown that low-energy laser pulses elicit compound nerve action potentials (CNAPs) and compound muscle action potentials (CMAPs), with resultant muscle contraction [19]. The stimulation threshold (0.3–0.4 J/cm^2) at optimal wavelengths in the infrared (1.87, 2.1, 4.0 μm) is at least 2 times less than the threshold at which any histological tissue damage occurs (0.8–1.0 J/cm^2). For this the stimulation threshold in the rat sciatic nerve was defined as the minimum radiation

exposure required to induce a visible muscle twitch. This stimulation threshold, at least in the rat sciatic nerve, is largely independent of the pulse duration which was varied between 5 μs and 5 ms. Given that such a motor response requires recruitment of numerous axons, the actual stimulation threshold for individual axons is likely to be significantly lower. In fact, stimulation thresholds reported for stimulation of sensory neurons in the gerbil cochlea were as low as 0.005 J/cm^2 [20]. Recently, we demonstrated the ability to stimulate neural pathways in a thalamocortical brain slice preparation producing artifact free electrical recordings of neural activation. Other investigators have started to adopt this approach and have shown feasibility in stimulation of the spiral ganglion cells in the cochlea [20–22], the facial nerve [23], the cavernous nerve [24], vagus nerve [25] and the vestibular system. Optical nerve stimulation has three fundamental advantages over electrical stimulation [15] that make it a potentially powerful tool for a number of procedures that currently employ electrical means as the standard of care: (1) the precision of optically delivered energy is superior to electrical stimulation techniques and can easily be confined to individual nerve fascicles or neural pathways in the brain without requiring separation between the area of stimulation and neighboring areas; (2) optical stimulation does not produce the previously mentioned stimulation artifact since the stimulation and recording occurs in different domains; and (3) optical stimulation is achieved in a non-contact fashion, a technical advantage that can minimize the risk of nerve trauma or metal-tissue interface concerns.

The following section illustrates the methodology and fundamental considerations of optical nerve stimulation. This will be discussed in terms of peripheral nerve stimulation in the sciatic nerve followed by a description of recent work in the central nervous system and its potential benefits in the clinic.

24.2.1 Feasibility, Methodology, and Physiological Validity in the PNS

Optical stimulation was first demonstrated and characterized in the peripheral nervous system (PNS) using the rat sciatic nerve. The typical experimental setup to perform optical stimulation with electrical recording of the nerve and muscle potentials is depicted in Fig. 24.1.

In general, an infrared pulsed laser source is coupled into a fiber optic or focused through a lens to provide a specific spot size for stimulating a nerve. A Holmium:YAG laser ($\lambda = 2.12$ μm, $\tau_p = 350$ μsec (FWHM)) was used to characterize optical stimulation in the PNS. This wavelength was shown to be optimal for peripheral nerve stimulation through a wavelength optimization study performed using the Free Electron Laser (FEL) at Vanderbilt University. The importance of this parameter will be discussed in detail in Section 24.2.2. Stimulation experiments in the rat sciatic nerve reveal that a 400–600 μm fiber diameter most efficiently results in excitation while maintaining precision in stimulation. Naturally, the diffraction limit dictates that the theoretical limits of spatial precision (i.e. the laser spotsize)

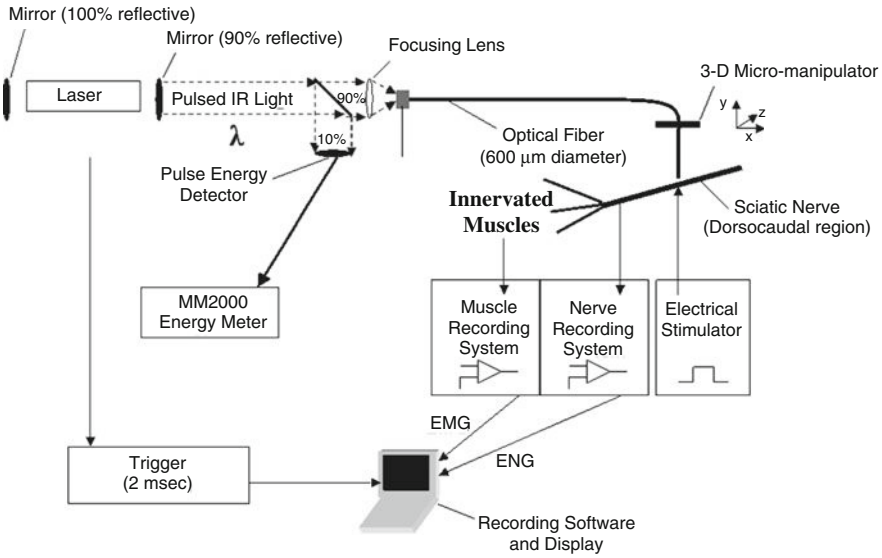


Fig. 24.1 Typical experimental setup for optical stimulation and recording in the rat sciatic nerve. (Adapted from [26])

for both delivery methods are on the order of a few micrometers. Radiant exposures required to stimulate vary depending on the wavelength of the laser source used. Electrical stimulation and recording of the resulting compound nerve action potential (CNAP) and compound muscle action potential (CMAP) can be employed to verify the validity of the evoked response from laser stimulation and compare this to the standard electrical stimulation methods. Application of a depolarizing neuromuscular blocker (succinylcholine) resulted in a measurable CNAP but loss of CMAP, validating the physiological nature of the laser induced APs and confirming the involvement of normal propagation of impulses from nerve to muscle upon optical stimulation.

In a proof of principle study, CNAPs and CMAPs were consistently observed and recorded using conventional electrical recordings (Fig. 24.2) from both electrical and optical peripheral nerve excitation methods. The similarity in the shape and timing of the signals from optical and electrical stimulus in Fig. 24.2 show that conduction velocities, represented by the time between the CNAP and CMAP, are equal. These traces imply the motor fiber types recruited and seen in the recorded compound action potentials are identical regardless of excitation mechanism. Stated differently, based solely on observation of the physiologic portions of recorded signals (nerve and muscle), one cannot discern between the two stimulation techniques. However, one can discern two important signal characteristics in Fig. 24.2 that allow one to differentiate between optically and electrically evoked potentials. First, the inherent electrical stimulation artifact is only seen in the electrically stimulated peripheral nerve recordings. Second, the spatial selectivity of the optical stimulation is superior. This is represented by the number of axons recruited with

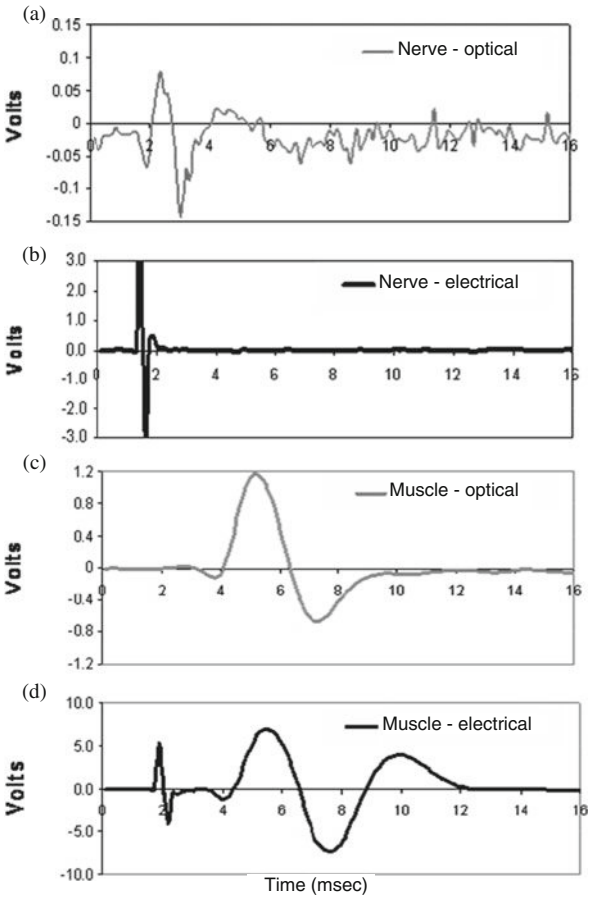


Fig. 24.2 Compound nerve and muscle action potentials recorded from sciatic nerve in rat. (**a**) CNAP recorded using optical stimulation at 2.12 μm; (**b**) CNAP from electrical stimulation; (**c**) Biceps Femoris CMAP recorded using optical stimulation at 2.12 μm; (**d**) Biceps Femoris CMAP using electrical stimulation. The stimulation time for all recordings occurred at $t = 1.8$ msec. CNAP responses were amplified 5000x and filtered using a high pass filter (>20 Hz) and low pass filter (<3 KHz). CMAP responses were amplified 1000x and filtered using a high pass filter (>0.05 Hz) and low pass filter (<5 KHz) (from [15])

optical stimulation when compared to electrical stimulation (evident by the order of magnitude difference in amplitude of the compound potentials – which is proportional to the number of axons recruited – between electrically and optically induced compound potentials).

The artifact associated with electrical stimulation prevents scientists from recording neural potentials near the site of stimulation. This is demonstrated in Fig. 24.2(b) which shows the CNAP response recorded from the rat sciatic nerve following electrical stimulus. Recording occurs 22 mm away from the site of stimulation. A large electrical artifact completely conceals the nerve response for over 1 msec following stimulation. Thus, the onset time and, in some cases, peak amplitude of the response is very difficult to distinguish from background, and therefore no relevant response characteristics can be inferred.

In contrast, Fig. 24.2(a) depicts the nerve response to optical stimulation (same stimulation and recording site as electrical) using laser radiant exposures above

stimulation threshold intensities, which does not contain a noise artifact. Thus, the nerve conduction velocities from the fast and slower conducting motor fibers within the sciatic nerve can be quantified in terms of timing and amplitude. The distance from stimulation to recording in the nerve was 22 mm and two peaks are seen at 0.6 and 2.5 msec following the laser stimulus (t =1.8 msec), yielding conduction velocities measured to be 36.7 m/sec with fast conducting axons and slower conduction fiber velocity of 8.8 m/sec. It is worth noting that the velocity of conduction within the nerve subsequent to laser stimulation falls within the normal range for the rat sciatic nerve fast conducting A_α motor neurons and slower conducting A_δ motor neurons.

With regards to the spatial precision, it is well known in electrophysiology that electrical stimulation has an unconfined spread of charge radiating far from the electrode. In the case of peripheral nerve stimulation, as the injected current required for stimulation increases, the volume of tissue affected by the electric field increases

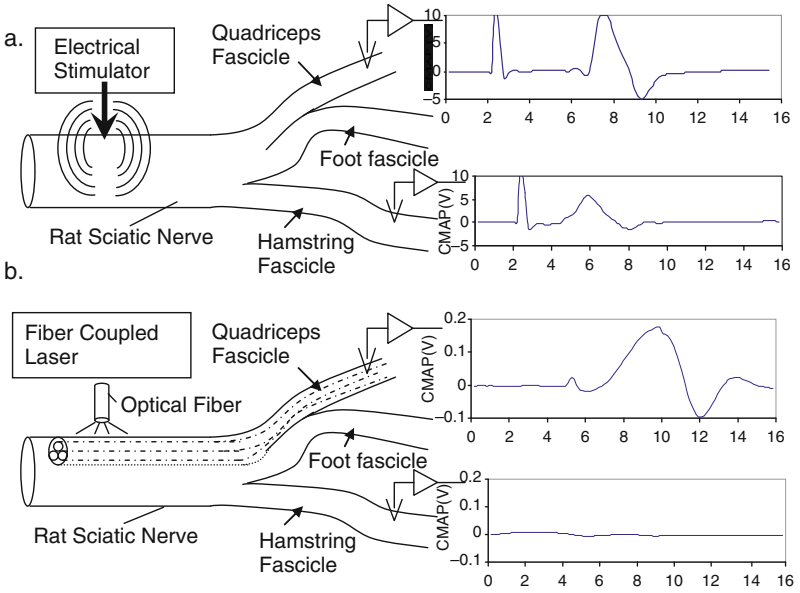


Fig. 24.3 Selective recruitment of isolated nerve fascicles within a large peripheral nerve (sciatic nerve of the rat) using electrical versus optical stimulation techniques. CMAP recording electrodes were placed within the gastrocenemius and biceps femoris of a rat approximately 40 and 55 mm from the site of stimulation, respectively. (**a**) Electrical stimulation with threshold energy (1.02 A/cm^2), was delivered proximal to the first nerve branch point on the fascicle leading to the gastrocenemius. Muscular responses within gastrocenemius and biceps femoris were simultaneously recorded. Note that using the minimum energy required to stimulate contraction of the gastrocnemius still results in stimulation of the neighboring biceps femoris fascicle. (**b**) Laser stimulation at threshold (0.4 J/cm^2) recorded in the gastrocnemius and no response observed in the biceps femoris. Grossly, the electrical stimulation results in excitation of the entire nerve and a subsequent twitch response from all innervated muscles. In contrast, the optical stimulation results in a muscle twitch of the muscle innervated by the targeted nerve fascicle. By moving the laser spot across the nerve, different individual muscle groups can indeed be stimulated (from [15])

proportionally (for review see: [8]). Thus, the CNAP and CMAP represent a population response to stimulation, made up of individual all-or-none responses from constituent axons, where a linear relationship exists between stimulation intensity and strength of the CNAP response [27]. Thus, a minimum value for spatial selectivity in activation exists, and appreciably limits the precision of electrical stimulation [28]. In contrast, lasers, owing to their wavelength-dependent optical penetration depth, variable spot size, and laser radiant exposure, excel in applications necessitating a precisely controlled and quantifiable volume of action in biological tissue [29, 30]. As a result of this spatial confinement of energy deposition, optical stimulation allows for more selective excitation of fascicles (small groups of axons within a nerve), resulting in isolated, specific muscle contraction (see Fig. 24.3).

As in electrical stimulation, increasing optical energy results in a linear increase in recruitment of axons (and thus in the recorded CNAP). A limit to laser excitation does exist at about 2 J/cm^2 stimulation radiant exposures, where a decrease in the physiologic response occurs. This is attributed to axon damage within the nerve as stimulation energies approach the laser thermal damage and ablation threshold, affecting the tissue's ability to generate and propagate action potentials.

24.2.2 Threshold for Stimulation Dependence on Wavelength

In the IR, tissue absorption is dominated by water which is the dominant chromophore in the peripheral nerve. As a first order approximation, light penetration in peripheral nerve tissue can be described by the wavelength dependent property of tissue absorption using Beer's Law. Because of this, we can also assume that the light propagation into the tissue will be confined to regions directly under the irradiated spot on the nerve surface.

Theoretically, the most appropriate wavelengths for stimulation will depend on the tissue geometry. A typical rat sciatic nerve section stimulated in this study was approximately 1.5 mm in diameter, with a 100–200 μm epineural and perineural sheath between the actual axons and the nerve surface. Despite the fact that the number of fascicles per nerve varies greatly across all mammalian species, the typical fascicle thickness is constant and tends to be between 200 and 400 μm [31]. Thus, to theoretically achieve selective stimulation of individual fascicles within the main nerve, the penetration depth of the laser must be greater than the thickness of the outer protective tissue (200 μm) and in between the thickness of the underlying fascicle (penetration depth of 300–500 μm). The near-infrared part of the spectrum (700–1300 nm) represents an area where light is relatively poorly absorbed while in the mid- to far-infrared ($\lambda > 1400$ nm) absorption by tissue water dominates and results in shallow penetration [30]. By irradiating the nerve surface overlying the target fascicle for stimulation within the main branch, infrared laser light may provide profound selectivity (in terms of spot size and optical penetration depth) in excitation of individual fascicles resulting in isolated muscle contraction without thermal damage to tissue if the appropriate wavelength and spot size is utilized.

Using a tunable Free Electron Laser (FEL) [32], we determined the stimulation threshold (defined as the minimum radiant exposure required for a visible muscle contraction) as well as the ablation/damage threshold (defined as the minimum radiant exposure required for visible cavitation or ejection of material from the nerve) for a number of wavelengths ($\lambda = 2.1,\ 3.0,\ 4.0,\ 4.5,\ 5.0,$ and $6.1\ \mu m$) that cover a range of optical penetration depths [33]. While the FEL is an excellent source for gathering experimental data and exploring the wavelength dependence of the interaction owing to its tunability, it is neither easy to use nor clinically viable. Nevertheless, experimental data gathered with this tunable light source can provide guidance towards the design of an appropriate and optimized turn-key bench top laser system for optical nerve stimulation.

The stimulation threshold exhibits a wavelength dependence which mirrors the inverse of soft tissue absorption curve. This trend is illustrated in Fig. 24.4a, which shows the stimulation and ablation threshold radiant exposures for five trials with each of the six wavelengths used in this study. The water absorption spectrum is included to discern general trends. Suitable wavelengths for optimal stimulation, those with maximum efficacy and minimum damage, can be inferred and generally lay away from the water absorption peaks (2.1 and 4.0 μm). As one would expect, based on a photothermal mechanism, the ablation threshold for neural tissue is inversely proportional to the (water) absorption coefficient. The stimulation threshold is lower at wavelengths with high absorption, but it is also easier to ablate tissue (less radiant exposure) at these wavelengths. Thus, a more useful indicator of optimal wavelengths is the safety ratio (defined as the ratio of threshold radiant exposure for ablation to that for stimulation).

This ratio (Fig. 24.4b) identifies spectral regions with a large margin between radiant exposures required for excitation and damage and thus of safety. Results indicate that the highest safety ratios (>6) are obtained at 2.1 and 4.0 μm, which correspond to valleys in tissue absorption and have nearly equivalent absorption coefficients. We can conclude that clinically relevant wavelengths for optimal stimulation, at least in the peripheral nerves and their anatomy/geometry, will not occur at peaks in tissue absorption because the energy required to produce action potentials within the nerve is roughly equal to the energy at which tissue damage occurs. We can also predict that absolute valleys in the absorption curve (i.e. visible and NIR region, 400–1400 nm) will not yield optimal wavelengths, because the low absorption, owing to lack of endogenous chromophores for these wavelengths in neural tissue, will distribute the light over a large volume leading to insufficient energy delivered to the nerve fibers for an elicited response. Instead, the most appropriate wavelengths for stimulation of the sciatic nerve occurs at relative valleys in IR soft tissue absorption, which produce an optical penetration depth of 300–500 μm (which corresponds to the optical penetration depth at $\lambda = 2.12\ \mu m$). In this scenario, the optical penetration depth matches up with the target geometry in order to stimulate one fascicle within the nerve. It should be clear that this finding also implies that other parts of the nervous system (with different morphology) may require a different wavelength for optimal stimulation such that the optical penetration depth is matched to the morphology of the targeted excitable tissue.

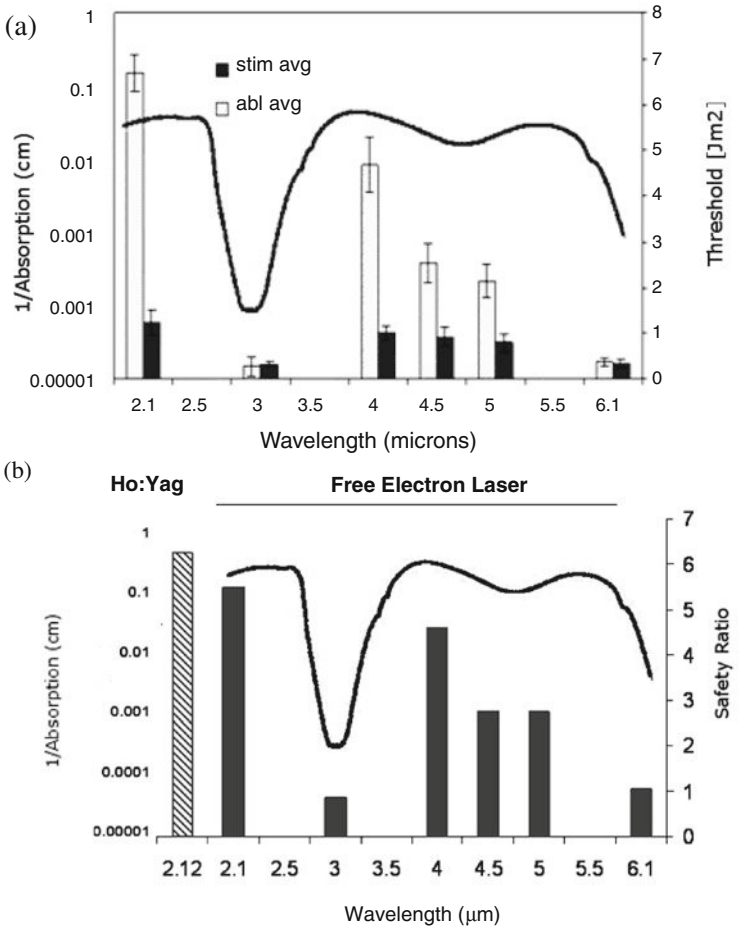


Fig. 24.4 Wavelength dependence of the (**a**) stimulation versus the ablation thresholds (**b**) the safety ratio = ablation threshold/stimulation threshold. The *solid line* in both figures indicates the optical penetration depth (left *y*-axis) (from [15]). In figure (**b**) safety ratio obtained for the Ho:YAG laser is shown in stripes (from [19])

There are few lasers that emit light at 4.0 μm in wavelength, and fiberoptic delivery at this wavelength is problematic as regular glass fibers do not transmit beyond 2.5 μm. However, the Holmium:YAG laser at 2.12 μm is commercially available, can be delivered via fiberoptics, and is currently used for a variety of clinical applications [34–38]. The Ho:YAG laser was successfully used for neural stimulation with an average stimulation threshold radiant exposure of 0.32 J/cm^2 and an associated ablation threshold of 2.0 J/cm^2($n = 10$), yielding a safety ratio of greater than 6.

24.2.3 Nerve Histological Analysis

While tissue ablation served as a good indicator for safe wavelengths by allowing calculation of a safety ratio for stimulation, this phenomenon is not a synonym for thermal damage resulting in altered tissue morphology and function. It is essential to define an exact range of "safe" laser radiant exposures, or the values between threshold and the upper end of radiant exposures which do not result in permanent tissue damage in order to strictly define what is appropriate for clinical use. Histological analysis was performed on excised rat sciatic nerves, extracted acutely (<1 h after stimulation) or 3–5 days following stimulation. Indications of damage include, but are not limited to, collagen hyalinization, collagen swelling, coagulated collagen, decrease or loss of birefringence image intensity, spindling of cells in perineurium and in nerves (thermal coagulation of cytoskeleton), disruption and vacuolization of myelin sheaths of nerves, disruption of axons, and ablation crater formation. For the survival study, histological sections were examined for evidence of any delayed neuronal damage and Wallerian degeneration.

Figure 24.5 shows the summary of the histological assessment of damage as function of radiant exposure for the survival study. Nerves from the survival study do not reveal damage to the nerve or surrounding perineurium in 8 of the 10 specimens, with damage occurring at radiant exposures above 2 times the stimulation threshold. These histological findings suggest that nerves can be consistently stimulated using optical means at or near threshold without causing any neural tissue damage. These findings are further corroborated by a functional analysis of toe spreading in the

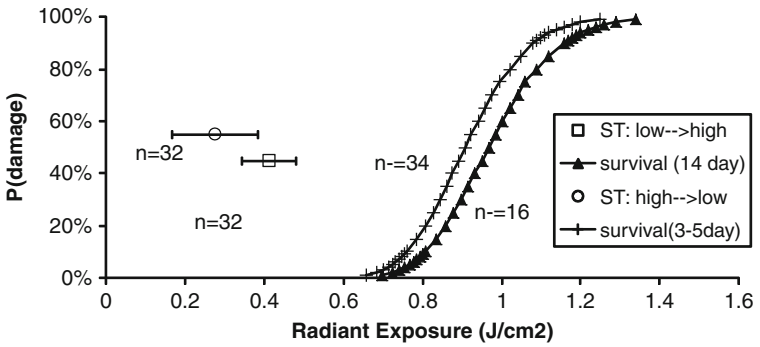


Fig. 24.5 Probability of damage as a function of laser radiant exposure compared to the stimulation threshold. Results from statistical analysis show the probability of histological nerve damage versus the laser radiant exposure used for transient optical nerve stimulation from data collected from 3- to 5-day survival studies (crosses, n=34) and 14 day survival studies (triangles, n=16). The results from studies to determine the range of threshold radiant exposures needed for stimulation are shown for comparison as well as the associated 95% confidence intervals. This graph illustrates that a safe margin exists between the maximum laser radiant exposures required to stimulate and the minimum radiant exposures necessary for damage [P(damage)=0%]. These results quantify the upper limits for brief, low repetition rate optical stimulation of the rat sciatic nerve (from [39])

survival animals. No functional neurological deficits were seen in any of the animals treated with at <2*x* the stimulation threshold [15].

24.2.4 Biophysical Mechanism

While optical stimulation has been shown to be an effective and advantageous method for stimulation of neural tissue, the obvious and intriguing question of the underlying mechanisms of this phenomenon are still poorly understood, although some progress has been made recently [40]. Exactly what biophysical stimulus is induced in the tissue by the absorbed laser light that ultimately results in an action potential? And, given this biophysical stimulus, what is the biological mechanism responsible for the transduction into action potentials (i.e. the opening of transmembrane sodium channels)? In general, many studies have been conducted investigating potential interaction effects by using all types of laser systems and tissue targets. Three main interaction mechanisms are classified today: photochemical, photothermal, and photomechanical. It is worth noting here that chemical, thermal, and mechanical means have all been previously shown to produce action potentials in neurons. Niemz makes the observation that all these seemingly different interaction types share a common property: the characteristic radiant exposure [J/cm^2] ranges from approximately 1 to 1000 J/cm^2. This is surprising since the irradiance itself [W/cm^2] varies over more than 15 orders of magnitude. Thus a single parameter distinguishes and primarily controls these processes: the duration of the laser exposure, which is largely similar to the interaction time itself [41]. According to a graph of the laser radiant exposure versus the duration of pulse width, the time scale can roughly be divided in three major sections: continuous wave or exposure times > 1 s for photochemical interactions; 100 s down to 1 μs for photothermal interactions; and 1 μs and shorter for photomechanical interactions. It should be clear, however, that these boundaries are not strict and adjacent interaction types cannot always be separated.

In brief, the group of photochemical interactions is based on the fact that light can induce chemical effects and reactions within macromolecules or tissues. Recent and exciting developments that rely on photochemical interactions include experimental applications in the field of photostimulation of neurons where light may be used to activate “caged” compounds [42–44]. In this scenario, stimulatory neurotransmitters are linked to an inactivating group (a “caged” compound). Upon UV light exposure cleavage of the neurotransmitter from its “cage” is achieved, rendering the active form of the stimulatory neurotransmitter only there where activated by light exposure. This technique takes advantage of the high spatial resolution of uncaging molecules with light.

We examined whether the mechanism for optical nerve stimulation is a result of photochemical effects from laser tissue interaction. The stimulation thresholds in the infrared part of the spectrum in essence follow the water absorption curve [15] suggesting that no “magical wavelength” has been identified, effectively excluding a single tissue chromophore responsible for any direct photochemical effects. This

also provides some evidence that the effect is directly thermally mediated or a secondary effect to photothermal interactions (i.e. photomechanical effects), as tissue absorption from laser irradiation can be directly related to the heat load experienced by the tissue. Theoretically, one can predict that a photochemical phenomenon is not responsible since infrared photon energy (<0.1 eV) is too low for a direct photochemical effect of laser tissue interactions and the laser radiant exposures used are insufficient for any multiphoton effects [45].

Maxwell's EM theory suggests an inherent electric field exists within laser light, which is associated with the propagation of light itself and driven by a time and space varying electric and magnetic field [46]. We questioned whether the electric field within the light beam used to irradiate and stimulate the peripheral nerve is large enough to directly initiate action potentials, considering the standard method of stimulation is through electrical means. To test this proposition, we used an alexandrite laser operating at 750 nm (near infrared light) to attempt stimulation of the peripheral nerve. This wavelength, unlike the Holmium:YAG wavelength, has minimal absorption in soft tissue; however, the electric field of intensity is similar regardless of wavelength. Thus, any stimulation reported with a low absorption wavelength would indicate that the electric field of the laser light mediates stimulation. Results explicitly prove a direct electrical field effect due to laser radiation traversing the tissue is highly unlikely as a means for optical stimulation because light from the alexandrite laser did not stimulate even at radiant exposures 50 times higher than those used for the Ho:YAG laser. This has been further substantiated by calculations that show that the maximum instantaneous intensity of the electric field owing to the laser radiant exposure is at least two orders of magnitude lower than that needed to for electric stimulation of the peripheral nerve [40]. Moreover, it is important to realize that the electric field, owing to light, oscillates at 10^{14}–10^{15} Hz, which is orders of magnitude higher than the typical electrical stimulation field oscillator frequency.

Photomechanical effects are secondary to rapid heating with short laser pulses (<1 μs) that produce forces, such as explosive events and laser-induced pressure waves, able to disrupt cells and tissue. Since we are operating well below the ablation threshold, ablative recoil can be excluded as a source of mechanical effects. In contrast, tissue heating will always result in thermo-elastic expansion. Nerve stimulation using pressure waves (rapid mechanical displacement, ultrasound) is well documented in the literature [47, 48]. We sought to prove or disprove photomechanical effects (thermoelastic expansion or pressure wave generation) leading to optical stimulation.

Contributions from pressure waves to optically stimulate the peripheral nerve were studied by examining the effect of pulse duration on stimulation threshold. It is clear from our results that the stimulation threshold radiant exposure does not change with pulse width through almost 3 orders of magnitude (5 μs–5 ms). Moreover, all pulse durations lie well outside of the stress confinement zone. Thus, there is strong evidence that laser-induced pressure waves are not implicated in the optical stimulation mechanism. Since pressure wave generation has been discarded as a plausible means, next tissue displacement during the laser pulse was

measured using a phase sensitive OCT setup [49] to test the actual magnitude of thermoelastic expansion of the tissue resulting from optical stimulation. The change in surface displacement of the rat sciatic nerve (ex vivo) upon irradiation with Ho:YAG radiant exposures slightly above threshold (0.4 J/cm^2) were measured to be 300 nm. Displacement of 300 nm in a 350 μsec pulse width is small, but not negligible. Nevertheless, while at this point we cannot exclude contributions of the thermoelastic expansion, this effect is thermal in origin.

Through this process of elimination we have systematically shown that the electric field, photochemical and photomechanical effects from laser tissue interactions do not result in excitation of neural tissue. Thus, we have arrived at the hypothesis that laser stimulation of neural tissue is mediated by a photothermal process resulting from transient irradiation of peripheral nerves using infrared light.

Photothermal interactions include a large group of interaction types resulting from the transformation of absorbed light energy to heat, leading to a local temperature increase and thus a temperature gradient both in time and in space. Photothermal effects generally tend to be non-specific and are mediated primarily by absorption of optical energy and secondly governed by fundamental principles of heat transport [50]. With regards to tissue damage, thermal interactions in tissue are typically governed by rate processes, i.e. not just the temperature plays a role but the duration for which the tissue is exposed to a particular temperature is also a parameter of major importance.

Infrared camera measurements during laser stimulation of rat sciatic nerves showed a peak temperature rise at the center of the spot of 8.95°C, yielding an average temperature rise of 3.66°C across the Gaussian laser spot. The thermal relaxation time of the rat peripheral nerve, was measured to be about 90 msec, which corresponds well to the theoretical value of about 100 msec (given the penetration depth of 0.25 mm). We can infer that the pulse width of light delivered to the tissue must be significantly less than 90 msec in duration to result in the desired stimulation effect. We can also infer that temperature superposition will begin to occur at higher repetition rates (>5 Hz) as the tissue requires slightly greater than 200 msec to return to baseline temperature. At repetition rates greater than 5 Hz tissue temperatures will become additive with each ensuing laser pulse, and resulting tissue damage may begin to occur with long term stimulation.

It was further shown that the nerve temperature increases linearly with laser radiant exposure. Recent literature suggests that thermal changes to mitochondria begin to occur as low as 43°C (protein denaturation begins at tissue temperature close to 57°C). This temperature corresponds to the radiant exposure associated with the onset of thermal damage found in histological analysis of short term laser nerve stimulation (0.8–1.0 J/cm^2). These results imply that optical stimulation of peripheral nerves is mediated through a thermal gradient as a result of laser tissue interaction and that this phenomenon is safe at radiant exposures of at least two times the threshold required for action potential generation. In the case of non-hydrated tissue, the temperature as a function of radiant exposure shifts upward 6°C. Here the mitochondrial damage will theoretically begin to occur between 0.5 and 0.6 J/cm^2, thus illustrating the importance of tissue hydration for safe and efficient nerve excitation. It should also be emphasized that with more sensitive endpoints

the stimulation threshold will be significantly lower than what is required for a visible muscle twitch. In that case, the temperature rise will be proportionally less, allowing higher pulse repetition rates. In fact, Izzo et al. [21] have shown that in the cochlea, optical stimulation of the spiral ganglion cells can be achieved with as little as 0.01 J/cm^2. At those radiant exposures they have continuously stimulated for several hours at repetition rates as high as 400 Hz without any loss of function.

Although the exact cellular mechanism is unknown, two separate theories have been developed on how the thermal energy from optical stimulation is tranduced into an action potential. The first theory concerns temperature sensitive TRPV1 and TRPV2 sodium ion channels which are found in numerous areas of both the central and peripheral nervous systems [51–55]. The second theory for the cellular mechanism of optical stimulation assumes all protein channels in a neuron's membrane may be affected by the heat which is deposited by the laser light. In this scenario the temperature influences the dynamic equilibrium between the "open," "closed" and "inactive" states of ion channels. Nevertheless, at this stage neither theory can fully explain the experimental findings and further research is needed to unravel these mechanistic underpinnings.

24.3 Other Applications

24.3.1 Sensory Nerve Stimulation – Towards an Optically-Based Cochlear Implant

Contemporary cochlear implants use electric current to stimulate neurons, but current spread limits the number of independent stimulation contacts to 10, decreasing the spatial resolution of the auditory frequencies which can be encoded by the implant. This fundamentally limits the fidelity of sound reproduction in state-of-the-art cochlear implants. However, optical stimulation has the potential to overcome this limitation due to high spatial resolution making it possible to independently stimulate more sites along the cochlea [21]. The increased number of independent stimulation sites would allow for encoding more sound frequencies along the basilar membrane of the cochlea, thereby increasing sound discrimination and speech recognition of the patient.

Izzo et al. have demonstrated the increased spatial selectivity of INS over electrical stimulation in several animal models using a variety of methods, including electrophysiology and histology [20–22, 56, 57]. Recordings from the inferior colliculus and tone masking experiments demonstrate that an optically stimulated population of neurons can be confined to an area similar to that stimulated by a moderate-level acoustic tone (Fig. 24.6) [56]. Non-damaging high frequency stimulation, 200–400 Hz, would be needed to adequately restore a sense of hearing and to correctly encode sound intensity. However, high rates of optical stimulation raise the possibility that an increase in temperature from each laser pulse would accumulate and cause tissue damage. Initial chronic stimulation experiments, up to 6 h and 400 Hz, showed that optical stimulation evoked a stable compound action

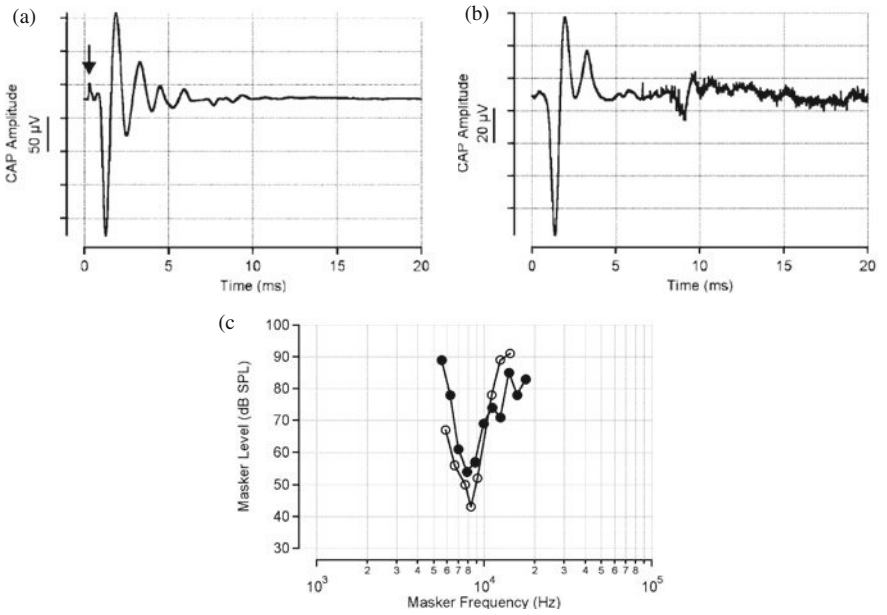


Fig. 24.6 Acoustically and optically evoked action potentials. Time waveforms of optically evoked compound action potentials (CAPs) are shown. (**a**) CAPs evoked from an acoustic click stimulus (81 dB SPL peak pressure), are shown. (**b**) This is a representative optical stimulation CAP with 0.06 J/cm^2, obtained from the same animal as (**a**). There is no CM present in the optical CAP. Noise seen from 6 to 20 ms in the optical CAP trace is electrical noise from the laser power supply (from Izzo et al. [21]). (**c**) Shows the spatial selectivity of optical stimulation (*closed circles*) which is nearly identical to that of acoustic stimulation (represented by the *open circles*) (from [56])

potential, a sensitive marker for the physiological state of the cochlea. Damage to the cochlea would have shown a drastic decrease in amplitude [57].

Optical stimulation has been used to successfully stimulate the cochlea of acute and chronic deafened animals in which acoustic thresholds are significantly elevated. The evoked potentials from optical stimulation in acutely deafened animals and chronic deafened animals were not significantly different from optical stimulation in normal animals, which demonstrates the feasibility of optical stimulation in neural systems which have undergone trauma resulting in neural degeneration [57]. Further studies will continue to investigate the effects of chronic optical stimulation with the aim of developing an implantable cochlear stimulator.

24.3.2 Optical Stimulation in the Central Nervous System (CNS)

Until recently, optical stimulation research had been confined to the peripheral nervous system, imploring the question of whether it is possible to stimulate the brain

using this technique. Thus the next step in the development of infrared neural stimulation was to achieve and optimize stimulation in the CNS, specifically the brain. The vastly different geometry and physiology of the brain implies that a different wavelength of infrared light may be needed to achieve optimal stimulation in the CNS. Moreover, whereas in the PNS detecting induced action potentials up or downstream in a nerve is relatively straightforward, stimulation in the brain with its intricate networks of interconnected neurons, especially when using a spatially precise method of stimulation poses a daunting challenge. Due to the complexity of the brain, a well characterized thalamocortical brain slice model was chosen to simplify the complexity of the neuronal networks of the brain. The brain slice model that was used preserved a three neuron network between cortical and thalamic neurons that reproduce action potential activity similar to that seen with in vivo animal studies [58, 59]. The three neuronal network of the thalamocortical brain slice allows for some organization and repeatability between individual preparations when optimizing infrared neural stimulation techniques. The purpose of this study was to prove feasibility and optimize the parameters of optical stimulation in vitro in the thalamocortical brain slice model.

Like the initial research performed in the PNS, initial feasibility studies were performed using the FEL to identify wavelengths (i.e. penetration depths) that could be used to stimulate the brain slices. Wavelengths in the infrared with penetration depths varying between 200 and 450 μm were used. Each slice had a thickness of 500 μm with a 180 μm section of dead or damaged tissue on each face of the brain slice. Figure 24.7 represents one of the first recordings made in the central nervous system (brain slice). It can be see that the signal exhibits a frequency lock with the repetition rate of the laser and exhibits only the biological signal with no stimulation artifact being recorded. It was found that light at each wavelength could stimulate thalamocortical brain slices with the resulting signal being blocked by tetratodotoxin (TTX). However, greater repetition rates were needed to efficiently stimulate the brain slice. In this initial study 3.65 μm was identified as the optimal wavelength (Fig. 24.8a). Light with a wavelength of 3.65 μm has a penetration depth which optimally delivers photons to 300 μm of living tissue in the brain slice model when compared to the other wavelengths tested. Nevertheless, this wavelength is not easily obtainable by conventional sources, and delivery of the laser light is challenging due to transmission limits of standard fiberoptics. Other wavelength ranges (with similar penetration depths) are being explored.

The central nervous system processes information at a higher rate when compared with the peripheral nervous system, requiring higher repetition rates when performing electrical stimulation [60–62]. Izzo et al. showed the need for higher repetition rates when using optical stimulation in the cochlea and the auditory-vestibular nerve to transfer stimulated sensory information to the brain [20, 22]. Therefore, optical stimulation should function at a higher repetition rate in the central nervous system than what was found in the sciatic nerve during the original development of optical stimulation. As expected, higher repetition rates reduced the threshold radiant energy needed to evoke action potentials in the brain slice (Fig. 24.8b). Here we were limited by the FEL laser to a 30 Hz repetition rate;

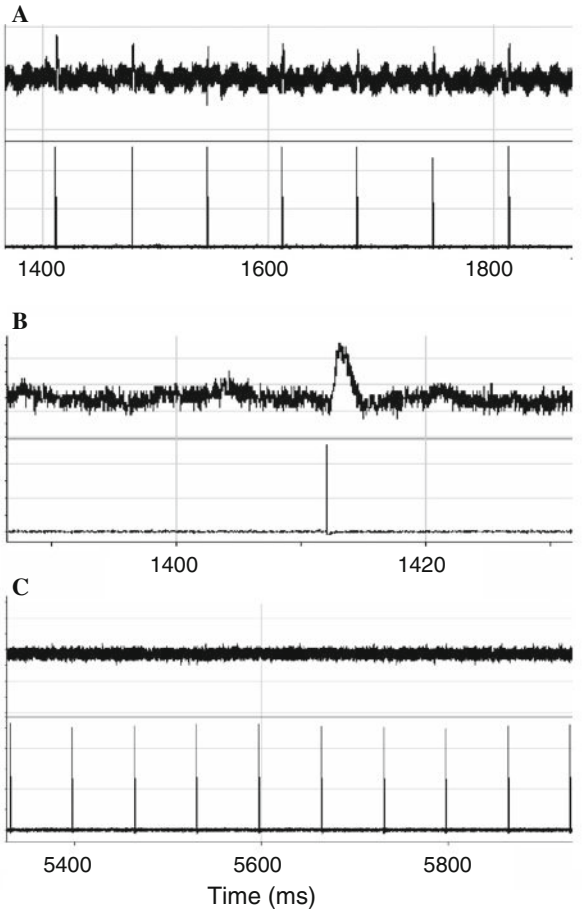


Fig. 24.7 Infrared stimulation evokes neural spike activity in the thalamocortical slice as seen in extracellular recordings. (**a**–**c**). *Top trace* shows recorded extracellular potentials; *bottom trace* represents timing of each pulse. (**a**) Infrared light evoked action potentials. (**b**) Zoomed trace on one spike. (**c**) Lack of infrared evoked spike activity after TTX application. Laser parameters: 1.46 J/cm^2, 15 Hz, 220 μm spot size, 4.0 μm light.

however, initial results from in vivo optical stimulation in the cortex has shown we can stimulate at 250 Hz with energy at the μJ level.

As in the PNS, the spot size was crucial in determining the quantity of neuronal tissue recruited during optical stimulation. As spot size increased, the threshold energy needed to stimulate the brain decreased (Fig. 24.8c). In the brain slice, a larger spot size increased the probability of detecting an optically induced signal due to the numerous projections of neurons in the brain slice by our glass pipette electrode located in one region of the slice. When spot size was approximately 500 μm

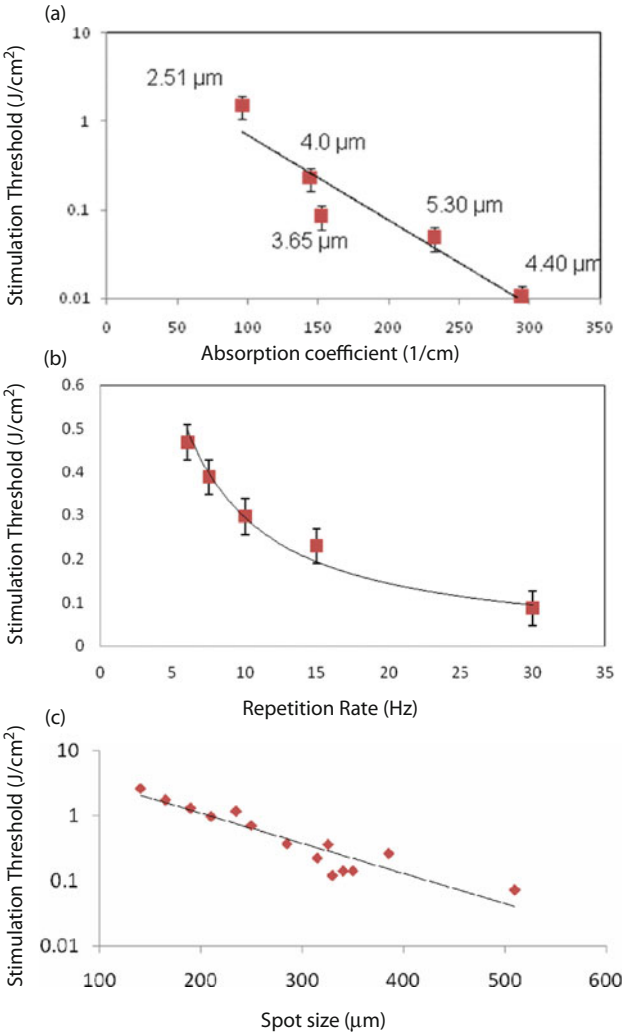


Fig. 24.8 Effect of wavelength (absorption coefficient), pulse repetition rate and spotsize on stimulation thresholds. (**a**) Infrared radiant energy levels at neural activation thresholds for different absorption coefficients. Laser parameters: Spotsize = 350±50 µm, Pulse duration = 1 ms; repetition rate = 30 Hz; (**b**) Increasing pulse frequency reduces radiant energy levels at neuronal activation thresholds. Laser parameters: λ=3.65 µm and spot size = 320±20 µm; (**c**) Increasing spot size reduces radiant energy levels at neuronal activation thresholds. Laser parameters: λ=3.65 um, repetition rate = 30 Hz, pulse duration = 1 ms

in diameter, stimulation threshold could be decreased to approximately 0.07 J/cm^2 using 3.65 µm light at a repetition rate of 30 Hz. Using the results from the brain slice experiments, we plan to continue the development of optical stimulation for use in applications in vivo.

24.4 Impact

24.4.1 Applications

The advantages of optical stimulation over conventional electrical means of stimulating neural tissue drive the applications where this novel technology may have a place. The exquisite spatial precision of optical stimulation can overcome limitations in electrical stimulation techniques used when precision stimulation of neural structures is required for peripheral nerve surgery during which small clusters of nerve fibers are stimulated to determine their viability in peripheral nerve repair [63]. In peripheral nerve surgery, electrical stimulation is utilized to identify the connectivity and functionality of specific nerve roots to be selectively avoided or resected. This usually requires dissecting the nerve bundles apart to determine which ones conduct through a damaged area and which bundles do not. The optical method could confine the stimulation easily to segments of a nerve without requiring separation between the intended area to be stimulated and other areas. Similarly, surgeries involving cranial nerves would benefit from precise functional testing, such as differentiating nerve tissue from tumor in small areas such as the central pontine angle through which the vestibular and facial nerve traverse. Auditory nerve stimulation could be significantly enhanced with a larger number of distinct stimulation sites along the cochlea than is currently possible using electrical means. This suggests that it may be possible to develop a better cochlear implant. In the brain, the field of neurosurgery gains the ability to better discriminate critical areas during resection surgery by using optical stimulation techniques instead of tradition electrical stimulation which could lead to a better quality of life for the patient after surgery. Deep brain stimulation (DBS) represents another area of neurosurgery where optical stimulation stands to have significant impact on clinical techniques used to treat patients with CNS ailments. At present, DBS functions by delivering a small amount of current to a specific area of the brain to treat diseases such as Parkinson's or epilepsy or to control chronic pain. The current essentially creates an "active" lesion which removes a tremor, or quiets neurons which are causing pain to be sensed or seizures to occur. Current electrical techniques suffer from current spread when electrical stimulation is applied to the brain, which sometimes causes debilitating effects and even death if the electrode is not placed correctly [64–67]. As we have shown, the potential of confining the optical energy to a small volume will allow for a smaller active lesion to be created in the brain, potentially removing any debilitating effects seen with electrical techniques.

Second, the lack of a stimulation artifact on recording electrodes enables a whole new area of investigation that was previously not possible [68–70]. Simply stated, this approach will allow recording of neural potentials near the source of stimulation. Additionally, fewer stimuli need to be applied due to reduction in the need for signal averaging (which usually requires hundreds of stimuli), which in turn facilitates higher throughput of mapping.

The third group of applications originates from the fact that optical stimulation, unlike its electrical equivalent, does not require direct contact between the stimulating probe (fiber optic) and the target tissue. This circumvents issues associated with properties of impedance, current shunting, and field distortion around the area of contact between the electrode and the tissue in the acute setting [8]. In the chronic setting, issues of half-cell potential differences, metal toxicity and tissue reaction to various implanted electrodes significantly limit the materials and sizes of materials used for chronic electrode implants. Furthermore, plating and deplating of metal in an ionic medium (interstitial fluid) is not an issue with optical stimulation. In principle, a fiberoptic-based chronic implant may be advantageous due to longer tissue stability (no unstable impedance characteristics) and safer (i.e. inert) interface materials (glass/fiber optic cable versus metal) [71]. Nevertheless, before such claims can be substantiated, the effects of chronically stimulating neural tissues optically must be studied carefully and thoroughly.

24.4.2 Future Directions

Optical stimulation presents a paradigm shift in neural activation that potentially has major implications for clinical neural stimulation as well as fundamental neurophysiology and neuroscience. To date, this concept has been demonstrated using large, cumbersome and expensive laboratory laser sources (FEL, Ho:YAG). For optical stimulation to find its way to practical utility and clinical use, a simple, user friendly, portable, reliable and low cost device must be developed. An initial, research grade optical stimulator based on solid-state laser technology has been developed and is now commercially available (Lockheed-Martin-Aculight, Capella). Built on a versatile solid state (diode) laser platform, this device can be used for routine use in laboratory settings, with the ultimate goal of delivering a clinically usable product. The prototype of the optical stimulator is a pulsed 1.85–1.87-μm (resulting in similar absorption as Ho:YAG laser in soft tissue) diode laser with a fiber-coupled output, representing a >95% reduction in size compared to the Ho:YAG laser. The laser-based stimulator, which appears similar to telecom products offered today, has the advantages of compactness & portability, high reliability, and low life cycle cost. Furthermore, the unit plugs into a regular power outlet (110 V) and has no special power or cooling requirements Experiments described above have been repeated with this laser system, and, as expected, show similar results.

The diode lasers will be used to advance the work of optical stimulation in the CNS by taking advantage of the high frequency repetition rates produced by these laser sources. As alluded to previously, we plan to develop optical stimulation in vivo by stimulating the somatosensory cortex of the rat brain. Here we use high frequency stimulation to simulate conditions seen in DBS to stimulate the surface of the somatosensory cortex. We plan to use optical imaging techniques to visualize activation of cortical neurons since optical imaging has been shown to have spatial

resolution on the order of 10 μm [72]. By combining two optical techniques, we are showing the usefulness of optics in neuroscience and how these tools can be used to further advance our knowledge of the nervous system.

In summary, we have shown a novel alternative to electrical stimulation to interface with the neural system using light. This method provides several unique advantages over traditional methods. However, transient optical stimulation of neural tissue is in its infancy and many questions remain open with regards to the underlying mechanism, the limitations of its utility, and applications that have not even been thought of at this time. Moving this field forward will require multidisciplinary approaches and intense research efforts on all fronts.

References

1. Schoonhoven R and Stegeman DF. Models and analysis of compound nerve action potentials. *Crit. Rev. Biomed. Eng*., 19:47–111 (1991).
2. Cole K. *Membranes, ions, and impulses: A chapter of classical biophysics*. University of California, Berkely, p. 569 (1968).
3. Plonsey R and Barr R. *Bioelectricity: A quantitative approach.* Kluwer Academic, New York, 2nd edition (2002).
4. Hille B. *Ion channels of excitable membranes.* Sinaure Associates, Sunderland, MA, 3rd edition (2001).
5. Huang C, Harootunian A, Maher M, Quan C, Raj C, McCormack K, Numann R, Negulescu P, Gonzalez J. Characterization of voltage-gated sodium channel blockers by electrical stimulation and fluorescence detection of membrane potential. *Nat. Biotechnol*., 24:439–446 (2006).
6. McGill K, Cummins K, Dorfman L, Berlizot BB, Leutkemeyer K, Nishimura D, Widrow B. On the nature and elimination of stimulus artifact in nerve signals evoked and recorded using surface electrodes. *IEEE Trans. Biomed. Eng*., 29:129–137 (1982).
7. Civillico E and Contreras D. Comparison of responses to electrical stimulation and whisker deflection using two different voltage-sensitive dyes in mouse barrel cortex *in Vivo*. *J. Membr. Biol*., 208:171–182 (2005).
8. Palanker D, Vankov A, Huie P, and Baccus S. Design of a high-resolution optoelectronic retinal prosthesis. *J. Neural. Eng*., 2:S105–S120 (2005).
9. Miller CA, Abbas PJ, and Brown CJ. An improved method of reducing stimulus artifact in the electrically evoked whole-nerve potential. *Ear Hear*, 21:280–290 (2000).
10. Wu WH, Ponnudurai R, Katz J, Pott CB, Chilcoat R, Uncini A, Rapoport S, Wade P, Mauro A. A failure to confirm report of light-evoked response of peripheral nerve to low power helium-neon laser light stimulus. *Brain Res*., 401:407–408 (1987).
11. Balaban P, Esenaliev R, Karu T, Kutomkina E, Letokhov V, Oraevsky A, and Ovcharenko N. He-Ne laser irradiation of single identified neurons. *Lasers Surg. Med*., 12:329–337 (1992).
12. Bragard D, Chen AC, and Plaghki L. Direct isolation of ultra-late (C-fibre) evoked brain potentials by CO_2 laser stimulation of tiny cutaneous surface areas in man. *Neurosci. Lett*., 209:81–84 (1996).
13. Gigo-Benato D, Geuna S, de Castro Rodrigues A, Tos P, Fornaro M, Boux E, Battiston B, and Giacobini-Robecchi MG. Low-power laser biostimulation enhances nerve repair after end-to-side neurorrhaphy: A double-blind randomized study in the rat median nerve model. *Lasers Med. Sci*., 19:57–65 (2004).
14. Ilic S, Leichliter S, Streeter J, Oron A, DeTaboada L, and Oron U. Effects of power densities, continuous and pulse frequencies, and number of sessions of low-level laser therapy on intact rat brain. *Photomed. Laser Surg*., 24:458–466 (2006).

15. Wells J, Kao C, Jansen ED, Konrad P, and Mahadevan-Jansen A. Application of infrared light for in vivo neural stimulation. *J. Biomed. Opt*., 10:064003 (2005).
16. Fork RL. Laser stimulation of nerve cells in Aplysia. *Science* 171:907–908 (1971).
17. Allegre G, Avrillier S, and Albe-Fessard D. Stimulation in the rat of a nerve fiber bundle by a short UV pulse from an excimer laser. *Neurosci. Lett*., 180:261–264 (1994).
18. Hirase H, Nikolenko V, Goldberg JH, and Yuste R. Multiphoton stimulation of neurons. *J. Neurobiol*., 51:237–247 (2002).
19. Wells J, Kao C, Mariappan K, Albea J, Jansen ED, Konrad P, and Mahadevan-Jansen A. Optical stimulation of neural tissue *in vivo*. *Opt. Lett*., 30:504–506 (2005).
20. Izzo AD, Walsh JT Jr, Jansen ED, Bendett M, Webb J, Ralph H, and Richter CP. Optical parameter variability in laser nerve stimulation: A study of pulse duration, repetition rate, and wavelength. *IEEE Trans. Biomed. Eng*., 54:1108–1114 (2007).
21. Izzo AD, Richter C-P, Jansen ED, and Walsh J. Laser stimulation of the auditory nerve. *Lasers Surg. Med*., 38:745–753 (2006).
22. Izzo AD, Walsh JT, Ralph H, Webb J, Bendett M, Wells J, and Richter C-P. Laser stimulation of auditory neurons: Effect of shorter pulse duration and penetration depth. *Biophys. J. Biophysj*.107.117150 (2008).
23. Teudt IU, Nevel AE, Izzo AD, Walsh JT Jr, and Richter CP. Optical stimulation of the facial nerve: A new monitoring technique? *Laryngoscope* 117:1641–1647 (2007).
24. Fried NM, Lagoda GA, Scott NJ, Su L-M, and Burnett AL. Optical stimulation of the cavernous nerves in the rat prostate. In: *Photonic therapeutics and diagnostics IV*. SPIE, San Jose, CA, USA, 1st edition, pp. 684213–684216 (2008).
25. Duke AR, Cayce JM, Gessel IV, and Jansen ED. Optical stimulation for the efferent vagus nerve *in vivo*. In: *Neural Interfaces Conference*. Cleveland, OH (2008).
26. Wells J, Konrad P, Kao C, Jansen ED, and Mahadevan-Jansen A. Pulsed laser versus electrical energy for peripheral nerve stimulation. *J. Neurosci. Methods*, 163:326–337 (2007).
27. Geddes LA and Bourland JD. The strength-duration curve. *IEEE Trans. Biomed. Eng*., 32:458–459 (1985).
28. Geddes LA and Bourland JD. Tissue stimulation: Theoretical considerations and practical applications. *Med. Biol. Eng. Comput*., 23:131–137 (1985).
29. van Hillegersberg R. Fundamentals of laser surgery. *Eur. J. Surg*., 163:3–12 (1997).
30. Vogel A and Venugopalan V. Mechanisms of pulsed laser ablation of biological tissues. *Chem. Rev*., 103:577–644 (2003).
31. Paxinos G. *The rat nervous system*. Elsevier, Sydney, Australia, 3rd edition (2004).
32. Edwards GS and Hutson MS. Advantage of the Mark-III FEL for biophysical research and biomedical applications. *J. Synchrotron. Radiat*., 10:354–357 (2003).
33. Hale GM and Querry MR. Optical-constants of water in 200-nm to 0.2-mm wavelength region. *Appl. Opt*., 12:555–563 (1973).
34. Razvi HA, Chun SS, Denstedt JD, and Sales JL. Soft-tissue applications of the holmium:YAG laser in urology. *J. Endourol*., 9:387–390 (1995).
35. Topaz O, Rozenbaum EA, Luxenberg MG, and Schumacher A. Laser-assisted coronary angioplasty in patients with severely depressed left ventricular function: Quantitative coronary angiography and clinical results. *J. Interv. Cardiol*., 8:661–669 (1995).
36. Kabalin JN, Gilling PJ, and Fraundorfer MR. Application of the holmium:YAG laser for prostatectomy. *J. Clin. Laser Med. Surg*., 16:21–27 (1998).
37. Fong M, Clarke K, and Cron C. Clinical applications of the holmium:YAG laser in disorders of the paediatric airway. *J. Otolaryngol*., 28:337–343 (1999).
38. Jones JW, Schmidt SE, Richman BW, Miller CC III, Sapire KJ, Burkhoff D, and Baldwin JC. Holmium:YAG laser transmyocardial revascularization relieves angina and improves functional status. *Ann. Thorac. Surg*., 67:1596–1601; discussion 1601–1592 (1999).
39. Wells JD, Thomsen S, Whitaker P, Jansen ED, Kao CC, Konrad PE, and Mahadevan-Jansen A. Optically mediated nerve stimulation: Identification of injury thresholds. *Lasers Surg. Med*., 39:513–526 (2007).

40. Wells J, Kao C, Konrad P, Milner T, Kim J, Mahadevan-Jansen A, and Jansen ED. Biophysical mechanisms of transient optical stimulation of peripheral nerve. *Biophys. J.*, 93:2567–2580 (2007).
41. Niemz MH. *Laser-tissue interactions*. Springer, Berlin, 3rd edition (2004).
42. McCray JA and Trentham DR. Properties and uses of photoreactive caged compounds. *Annu. Rev. Biophys. Biophys. Chem.*, 18:239–270 (1989).
43. Eder M, Zieglgansberger W, and Dodt HU. Neocortical long-term potentiation and long-term depression: Site of expression investigated by infrared-guided laser stimulation. *J. Neurosci.*, 22:7558–7568 (2002).
44. Eder M, Zieglgansberger W, and Dodt HU. Shining light on neurons–elucidation of neuronal functions by photostimulation. *Rev. Neurosci.*, 15:167–183 (2004).
45. Thomsen S. Pathologic analysis of photothermal and photomechanical effects of laser-tissue interactions. *Photochem. Photobiol.*, 53:825–835 (1991).
46. Waldman G. *Introduction to light: The physics of light, vision, and color.* Prentice-Hall, Englewood Cliffs, NJ (1983).
47. Shusterman V, Jannetta PJ, Aysin B, Beigel A, Glukhovskoy M, and Usiene I. Direct mechanical stimulation of brainstem modulates cardiac rhythm and repolarization in humans. *J. Electrocardiol.*, 35(Suppl):247–256 (2002).
48. Norton SJ. Can ultrasound be used to stimulate nerve tissue? *Biomed. Eng., Online* 2:6 (2003).
49. Rylander CG, Dave DP, Akkin T, Milner TE, Diller KR, and Welch AJ. Quantitative phase-contrast imaging of cells with phase-sensitive optical coherence microscopy. *Opt. Lett.*, 29:1509–1511 (2004).
50. Jacques SL. Laser-tissue interactions. Photochemical, photothermal, and photomechanical. *Surg. Clin. North Am.*, 72:531–558 (1992).
51. Stein RJ, Santos S, Nagatomi J, Hayashi Y, Minnery BS, Xavier M, Patel AS, Nelson JB, Futrell WJ, Yoshimura N, Chancellor MB, and De Miguel F. Cool (TRPM8) and hot (TRPV1) receptors in the bladder and male genital tract. *J. Urol.*, 172:1175–1178 (2004).
52. Liapi A and Wood JN. Extensive co-localization and heteromultimer formation of the vanilloid receptor-like protein TRPV2 and the capsaicin receptor TRPV1 in the adult rat cerebral cortex. *Eur. J. Neurosci.*, 22:825–834 (2005).
53. Takumida M, Kubo N, Ohtani M, Suzuka Y, and Anniko M. Transient receptor potential channels in the inner ear: Presence of transient receptor potential channel subfamily 1 and 4 in the guinea pig inner ear. *Acta Otolaryngol.*, 125:929–934 (2005).
54. Yu S, Undem BJ, and Kollarik M. Vagal afferent nerves with nociceptive properties in guinea-pig oesophagus. *J. Physiol.*, 563:831–842 (2005).
55. Watanabe N, Horie S, Michael GJ, Keir S, Spina D, Page CP, and Priestley JV. Immunohistochemical co-localization of transient receptor potential vanilloid (TRPV)1 and sensory neuropeptides in the guinea-pig respiratory system. *Neuroscience*, 141:1533–1543 (2006).
56. Littlefield P, Izzo AD, Mundi J, Walsh JT, Jansen ED, Bendett M, Webb J, Ralph H, and Richter C-P. Characterization of single auditory nerve fibers in response to laser stimulation. In: *Optical interactions with tissue and cells XIX*. SPIE, San Jose, CA, 1st edition, pp. 68540F–68546F (2008).
57. Richter CP, Bayon R, Izzo AD, Otting M, Suh E, Goyal S, Hotaling J, and Walsh JT Jr. Optical stimulation of auditory neurons: Effects of acute and chronic deafening. *Hear Res.*, 242:42–51 (2008).
58. Agmon A and Connors BW. Thalamocortical responses of mouse somatosensory (barrel) cortex *in vitro*. *Neuroscience*, 41:365–379 (1991).
59. Kao CQ and Coulter DA. Physiology and pharmacology of corticothalamic stimulation-evoked responses in rat somatosensory thalamic neurons *in vitro*. *J. Neurophysiol.*, 77:2661–2676 (1997).
60. Salzman CD, Britten KH, and Newsome WT. Cortical microstimulation influences perceptual judgements of motion direction. *Nature*, 346:174–177 (1990).

61. Stepniewska I, Fang PC, and Kaas JH. Microstimulation reveals specialized subregions for different complex movements in posterior parietal cortex of prosimian galagos. *Proc. Natl. Acad. Sci. U.S.A.*, 102:4878–4883 (2005).
62. Tehovnik EJ, Tolias AS, Sultan F, and Slocum WM. Logothetis NK direct and indirect activation of cortical neurons by electrical microstimulation. *J. Neurophysiol.*, 96:512–521 (2006).
63. Weiner RL. Peripheral nerve neurostimulation. *Neurosurg. Clin. N. Am.*, 14:401–408 (2003).
64. Constantoyannis C, Kumar A, Stoessl AJ, and Honey CR. Tremor induced by thalamic deep brain stimulation in patients with complex regional facial pain. *Mov. Disord.*, 19:933–936 (2004).
65. Novakova L, Ruzicka E, Jech R, Serranova T, Dusek P, and Urgosik D. Increase in body weight is a non-motor side effect of deep brain stimulation of the subthalamic nucleus in Parkinson's disease. *Neuro. Endocrinol. Lett.*, 28:21–25 (2007).
66. Umemura A, Toyoda T, Yamamoto K, Oka Y, Ishii F, and Yamada K. Apraxia of eyelid opening after subthalamic deep brain stimulation may be caused by reduction of levodopa. *Parkinsonism Relat. Disord.*, 14(8):655–657 (2008).
67. Witt K, Daniels C, Reiff J, Krack P, Volkmann J, Pinsker MO, Krause M, Tronnier V, Kloss M, Schnitzler A, Wojtecki L, Botzel K, Danek A, Hilker R, Sturm V, Kupsch A, Karner E, and Deuschl G. Neuropsychological and psychiatric changes after deep brain stimulation for Parkinson's disease: A randomised, multicentre study. *Lancet Neurol.*, 7:605–614 (2008).
68. Fiore L, Corsini G, and Geppetti L. Application of non-linear filters based on the median filter to experimental and simulated multiunit neural recordings. *J. Neurosci. Methods*, 70:177–184 (1996).
69. Wagenaar DA and Potter SM. Real-time multi-channel stimulus artifact suppression by local curve fitting. *J. Neurosci. Methods*, 120:113–120 (2002).
70. Andreasen LN and Struijk JJ. Artefact reduction with alternative cuff configurations. *IEEE Trans. Biomed. Eng.*, 50:1160–1166 (2003).
71. Agnew WF, McCreery DB, Yuen TG, and Bullara LA. Histologic and physiologic evaluation of electrically stimulated peripheral nerve: Considerations for the selection of parameters. *Ann. Biomed. Eng.*, 17:39–60 (1989).
72. Chen LM, Turner GH, Friedman RM, Zhang N, Gore JC, Roe AW, and Avison MJ. High-resolution maps of real and illusory tactile activation in primary somatosensory cortex in individual monkeys with functional magnetic resonance imaging and optical imaging. *J. Neurosci.*, 27:9181–9191 (2007).

Chapter 25
Summary and Future

R. Rox Anderson

Decades of effort understanding, inventing, designing, developing and using photothermal treatments for human disease have clearly improved medical care. Lasers are the treatment of choice for common eye, skin, upper airway, gynecological and venous disorders. The microscopic precision of laser tissue ablation and selective photothermolysis are unprecedented in medicine. Optical diagnostics have also made significant impact. Is the pace of progress in this field likely to continue? Ophthalmology and dermatology are the only two medical specialties that depend heavily on optical treatments by photothermal and photochemical mechanisms, and ophthalmology also depends heavily on optical diagnostics. Many eye and skin problems fit for optical solutions remain unsolved. Lasers, light-activated drugs, optical nanoparticles, in vivo microscopies, and especially integrated optical diagnostic-therapeutic systems will continue to be developed to benefit problems including cancers, autoimmune diseases, infection, malformations, scars, metabolic disorders, etc.

Ironically, problems in less-accessible organs pose a greater need for the many advantages of photothermal therapy, photodynamic therapy and optical diagnostics, but the old analogy to first picking low-hanging fruit is very apt – we have so far used light where we can easily reach. This field is therefore poised for explosive progress when we figure out how to access much more of the body, as part of routine medical care. By the low-hanging-fruit analogy, our present catheters and endoscopes are equivalent to climbing the tree, as they are limited to moving along major passageways in the body. Miniature optical imaging systems are now the best candidates for access into these *and other* body sites. Strategies that successfully combine imaging for guidance, microscopy for diagnostic accuracy, and light for its wide variety of precision treatments in real time, could broadly transform medical practice. In addition, it is no longer justifiable to develop things that render medical care at greater overall cost, risk or inefficiency. Reducing cost, inefficiency and risk

R.R. Anderson (✉)
Wellman Center for Photomedicine, Massachusetts General Hospital,
Harvard Medical School, Boston, MA 02114, USA
e-mail: RRAnderson@partners.org

A.J. Welch, M.J.C. van Gemert (eds.), *Optical-Thermal Response of Laser-Irradiated Tissue*, 2nd ed., DOI 10.1007/978-90-481-8831-4_25,

will be major drivers in this field of bioengineering, because optical systems are inherently rapid, accurate, portable and often ideal for point-of-care use. Solving health problems in developing countries is especially challenging, and worthwhile. This chapter briefly discusses some foreseeable opportunities, with an emphasis on photothermal therapy.

25.1 Photothermal Therapies

The photothermal therapies include those using light, or light in combination with an exogenous chromophore such as a dye, drug or microparticle, to cause local heating. These are summarized in Table 25.1. Laser-induced hyperthermia, ablative and non-ablative laser surgeries, thermal tissue welding, fractional laser treatments using arrays of microbeams, selective photothermolysis at millimeter, micrometer and potentially nanometer spatial scales, and photothermally-driven cavitation/stress wave surgery (laser lithotripsy) are present widely-used examples of light-only photothermal therapies. Presently, laser removal of tattoos and drug-induced pigments

Table 25.1 Some existing and *emerging (italics)* photothermal therapies

Via tissue chromophores				
Melanin[1]	**Hemoglobins**[1]	**Lipid**[1]	**Water**	**Protein**
laser hair removal[2]	portwine stains	*fat removal,*	surgical lasers	*Amide bond*
pigmented lesions[3]	telangiectasia	*acne*	lithotripsy	*excitation for*
glaucoma[3]	superficial veins	*(sebaceous*	deep varicose veins	*precise surgery*
diabetic retinopathy	psoriasis, scars	*glands)*	fractional laser	
macular	vocal dysplasia		treatments	
degeneration?	viral papillomas		(microbeam arrays)	
	cancer		*robotic surgery*	
Exogenous photothermal chromophores				
Tattoo inks[1]	**Other micro or nanoparticles, e.g. gold (plasmonic), Ab-labelled gold, silver, nanoshells**[4]		**Dyes and drugs, e.g. Indocyanine, methylene blue**	
Tattoo removal[3]	*immunomodulation*		*tissue welding*	
	lymphomas		*vascular lesions*	
	solid tumors		*cancer*	
			acne	

[1]Examples of selective photothermolysis, in which confined photothermal excitation of localized chromophores creates selective damage through thermal coagulation and/or local thermally-driven cavitation

[2]Optical hair removal uses millisecond laser or flashlamp pulses that allow sufficient thermal conduction from pigmented hair shafts to non-pigmented target stem cells in the surrounding epithelium

[3]Examples of specific cell targeting, typically using nanosecond pulses to drive photothermal excitation of intracellular nanoscale organelles such as melanosomes. Mechanisms of damage involve microthermal, microcavitation and/or very local stress waves

[4]Gold plasmonic nanoparticles are highly efficient cell targeting agents when pumped by sub-nanosecond pulses, putatively through cavitation and stress waves coupled to photothermal excitation. Classical thermal damage mechanisms probably play a minor role

from the skin are the only routinely used clinical examples of photothermal therapy directed by an exogenous chromophore.

25.1.1 Ablative and Non-ablative Surgical Lasers

Continuous or pulsed infrared lasers at wavelengths absorbed primarily by water have been used for decades to vaporize and/or thermally coagulate tissue, and some surgeons still refer to a CO_2 laser as "the" laser. For such general surgical applications, lasers are have usually been less cost-effective than electrosurgical devices. However, new laser technologies that have not yet been applied to this general application, may have major advantages. Semiconductor, doped fiber lasers and fiber amplifiers developed for telecommunications offer greatly improved versatility, cost, reliability, size, efficiency, lifetime and ability to deliver power endoscopically compared with conventional devices. There may be a significant opportunity to replace or compete with standard radiofrequency devices used for cutting and coagulation in open and endoscopic surgery. Microbeam surgical applications have historically been limited by a lack of control systems, which are usually none or a simple pattern scanner. In the future, image-guided laser microbeam systems could be software-programmable treatment tools, and could also enable interactive robotic surgical platforms, e.g. for removing solid tissue tumors guided by a tissue imaging system in real time. Every major component needed for image-guided precise surgery have been demonstrated and are sitting "on the shelf" of the biomedical optics field.

25.1.2 Selective Photothermolysis Using Endogenous Chromophores

The ability to precisely damage certain "target" structures inside tissue is a major advantage of laser treatment. This can be accomplished by photochemical, photothermal, secondary mechanical, plasma-mediated, and other mechanisms. Since the mid-1980s, selective photothermolysis has been the major strategy for developing photothermal treatment lasers for skin, eye, and several endoscopic applications. However, the medical potential for endoscopic use of selective photothermolysis, particular for vascular targets, is far from realized. Specific targeting of vessels based on their size and oxygen saturation is possible, by choice of pulse duration, wavelength, and beam optics. In theory, a vein-selective laser could be constructed, for example. While creative uses for microvascular-selective lasers in many organs other than skin remain to be explored, even the classical examples of their use in skin can probably be greatly improved. Unfortunately, a "cookbook" approach is often taken in practice, for example treating babies with portwine stains using a standard set of laser exposure and skin cooling parameters. The general need for individualizing treatment to optimize individual response is enormous. Selective photothermolysis is widely used now for vascular malformations, glaucoma, tumors, scars, pigmented lesions, and aging-related changes. Pulsed dye and

KTP lasers have recently been adapted for angio-selective vocal fold surgery of papillomas, scars and superficial carcinomas.

Some very fundamental aspects of laser-tissue interactions underlying selective photothermolysis have yet to be understood and developed. For smaller targets, shorter pulse duration and higher power density are needed for adequate thermal confinement, because thermal relaxation time of the target decreases with the square of its size. Submicrosecond pulses used for various treatments often cause cavitation, stress waves, material fracture, multiphoton absorption, ionization and/or plasma formation in addition to very local thermal damage. These mechanisms can be useful at both gross and subcellular spatial scales. For example, plasma formation and high thermal stress assists stone fracture during laser lithotripsy. Thermally-driven microscopic cavitation around intracellular nanoparticles of tattoo ink or melanin are probably the dominant mechanism for nanosecond and picosecond selective photothermolysis. For biological targets below about 10 μm diameter, selective photothermolysis in its conventional single-pulse mode usually causes some combination of classical thermal denaturation and mechanical damage. Separation of these mechanisms by using pulse trains is theoretically possible, but has not been fully explored. An Arrhenius model for irreversible thermal denaturation suggests that multiple short thermal cycles, could be used to accumulate thermal damage. Selective photothermolysis from a train of optical pulses that could separate thermal from mechanical damage mechanisms would extend the range of applications in delicate tissues such as retina. In theory, the amount of thermal denaturation could be arbitrarily accumulated in light-absorbing targets down to the scale of large macromolecules that contain a strongly absorbing moiety. This and other strategies are fundamentally interesting.

The present medical applications of selective photothermolysis are all based on electronic absorption transitions (UV, visible, near infrared) of natural or exogenous chromphores such as melanin, hemoglobins, or tattoo inks. In the mid- and far-infrared spectrum, vibrational absorption transitions of macromolecules are strong and specific, but absorption by water tends to dominate. However, there may be tissues for which vibrational excitation provides enough absorption contrast for selective photothermal effects. Fat, other lipid-rich tissues, and bone have relatively less water content. It has recently been shown that vibrational absorption bands of CH_2 near 1210 and 1720 nm can be used to selectively "target" fat and sebaceous glands in skin respectively, which have high lipid content. Ablation of tissue at the far-infrared absorption maxima of amide bonds has also been described to enhance the efficiency and precision of laser surgery.

25.1.3 Selective Photothermolysis Using Exogenous Particles or Dyes

Dealing with skin tattoos has taught us that human cells can easily, safely and permanently be labeled with optical nanoparticles, then selectively killed at any time later using a pulse of light. Tattoos consist of indigestible, insoluble nanoscale solid

particles with strong optical absorption and/or scattering. Upon introduction into dermis, the foreign material is phagocytosed by dermal cells, which retain them "forever" as the tattoo itself sometimes says. Standard laser treatment of tattoos uses nanosecond domain (Q-switched laser) pulses that rapidly heat the particles, causing cavitation that ruptures the ink-containing cells. Although it is well established that picosecond pulses are far more efficient, these have not been made available for medical use. Potentially, femtosecond laser pulses would be even more efficient and perhaps less color-selective due to multiphoton absorption. This has not been well explored. With present ns pulses for tattoo removal, the laser wavelength must be chosen for absorption by the various tattoo ink colors. Some colors such as yellow are not removed by presently used lasers. In theory, ultrashort laser pulses may be nearly color-blind for tattoo removal.

The safety of laser tattoo removal and of tattoos themselves is impressive. Some tattoo inks are carcinogenic, toxic or allergic, but the majority of them are not. Tattoos now adorn about one quarter of the entire US adult population, despite a nearly complete lack of medical or regulatory oversight. This situation suggests that optically-pumped nanoparticles could almost surely be developed for selective photothermolysis of various target cells. Antibody-labelled nanogold particles were demonstrated to effect selective ps laser pulse killing of T lymphocyte subsets. More recently, laser-pumped plasmonic nanoparticles have been shown to target tumor cells.

25.1.4 Fractional Photothermolysis

Fractional photothermolysis is a recently developed approach for stimulating tissue remodeling. Fractional photothermolysis uses focused mid-infrared (nonablative) or far-infrared (ablative) laser microbeam exposures to create thousands or millions of microthermal damage zones (MTZs) within the tissue. Healing is rapid because each photothermal wound is very small (submillimeter), and the immediately surrounding tissue is stimulated to locally remove and remodel each photothermal wound. With a dense array of MTZs, this response "turns over" a large fraction of the tissue. Surprisingly, human skin heals rapidly and without apparent scarring, even when up to 50% of its volume is killed or removed in this manner. Fractional photothermolysis has recently been described and developed only for skin. Columns of thermal damage that are several hundred micrometers wide and up to several mm deep can be created. Depth can be precisely adjusted by varying the energy delivered per microbeam. Compared with conventional laser resurfacing, fractional treatment has major advantages including rapid healing, and lower risk of scarring or infection. Non-ablative (mid-infrared) and ablative (far-infrared) laser systems have recently become available for skin treatment of photoaging, scars, pigmentation disorders, and other maladies.

Fractional laser treatment is a deceptively simple new paradigm that will probably lead to major new capabilities. Ablative fractional lasers are the first medical devices ever, that allow removal of a large portion of tissue to essentially any

well-defined depth, followed by rapid healing. Depth is controlled primarily by the microbeam geometrical optics and energy delivered. Thus far, the fractional treatment laser systems deliver random or fixed patterns of microbeam exposures. "Smart" versions of fractional treatment lasers seem very suitable for robotic treatments. There are many tissue target structures that lack absorption contrast, a necessary part of selective photothermolysis. Potentially, feedback or image-guided fractional photothermolysis could offer software-programmable processing of tissue on the microscopic scale.

25.2 Imaging, and Photothermal Treatments

The recent advances in imaging scattered and/or stimulated light (fluorescence, multiphoton fluorescence, harmonic generation, Raman, Brillouin scattering, etc.) from in vivo tissue are impressive. Optoacoustic tissue imaging combines advantages of optical imaging (high absorption contrast) with those of ultrasound imaging (depth); this is in early development and looks particularly promising in the long run. Combined tissue imaging and treatment systems could be truly revolutionary for the practice of medicine and surgery. Pathology, the branch of medicine that deals mainly with microscopic diagnosis and analysis of disease mechanisms, might experience the same transition that radiology underwent about 3 decades ago when x-ray, MRI and ultrasound imaging tools were applied to guide procedures such as biopsies, catheterizations, cancer treatments, hemorrhage control, drug delivery, etc. Rather than analyzing static biopsy samples remotely from patient care, the practice of "interventional pathology" could be born when in vivo microscopy is used to guide procedures and/or make a primary diagnosis. The generally upward-spiraling cost of medical care could, in theory, be drastically decreased for some applications. However, this will not occur if academic laboratories, academic medicine, and industry fail to connect new technologies with early and appropriate high-value applications. Medical imaging is generally useless if it is not linked to prevention, prognosis or treatment. Optical technologies for diagnostic imaging, for guiding catheters or needles, for minimally-traumatic access, for surgical guidance, or for delivery of drugs or treatments are now poised to be integrated with treatment technologies. "See-and-treat" systems that use photothermal interactions would have many advantages.

25.3 Summary

Research and development of photothermal tissue interactions are promising for a variety of biomedical applications; some fundamental aspects of photothermal interactions still elude our understanding. In addition to stand-alone devices, integration of optical technologies into practical, affordable, robust systems is particularly promising. The host of applications include tissue engineering, drug development, detection of early response to interventions, gene expression and other rapid diagnostics, reliable screening, and many therapies specific to medical problems.

Index

A.J. Welch, M.J.C. van Gemert (eds.), *Optical-Thermal Response of Laser-Irradiated Tissue*, 2nd ed., DOI 10.1007/978-90-481-8831-4,

Q

R