Landolt-Börnstein Numerical Data and Functional Relationships in Science and Technology *New Series* / Editor in Chief: W. Martienssen

Group VIII: Advanced Materials and Technologies Volume 1

Laser Physics and Applications

Subvolume A: Laser Fundamentals

Part 2

Editors: H. Weber, G. Herziger, R. Poprawe

Authors:M.G. Benedict, M. Freyberger, F. Haug, N. Hodgson, N. Kerwien,B. Kuhlow, W. Martienssen, H. Paul, G. Pedrini, D. Ristau, W.P. Schleich,M. Scholl, H.J. Tiziani, E.D. Trifonov, K. Vogel



ISSN 1619-4802 (Advanced Materials and Technologies)

ISBN-10 3-540-28824-4 Springer Berlin Heidelberg New York ISBN-13 978-3-540-28824-4 Springer Berlin Heidelberg New York

Library of Congress Cataloging in Publication Data: Landolt-Börnstein: Numerical Data and Functional Relationships in Science and Technology, New Series. Editor in Chief: W. Martienssen. Group VIII, Volume 1: Laser Physics and Applications. Subvolume A: Laser Fundamentals. Part 2. Edited by H. Weber, G. Herziger, R. Poprawe. Springer-Verlag, Berlin, Heidelberg, New York 2006. Includes bibliographies. 1. Physics - Tables. 2. Chemistry - Tables. 3. Engineering - Tables. I. Börnstein, Richard (1852-1913). II. Landolt, Hans (1831-1910). QC 61.23 502'.12 62-53136

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in other ways, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer-Verlag. Violations are liable for prosecution act under German Copyright Law.

Springer is a part of Springer Science+Business Media. springeronline.com © Springer-Verlag Berlin Heidelberg 2006 Printed in Germany

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Product Liability: The data and other information in this handbook have been carefully extracted and evaluated by experts from the original literature. Furthermore, they have been checked for correctness by authors and the editorial staff before printing. Nevertheless, the publisher can give no guarantee for the correctness of the data and information provided. In any individual case of application, the respective user must check the correctness by consulting other relevant sources of information.

Cover layout: Erich Kirchner, Heidelberg Typesetting: Authors and Redaktion Landolt-Börnstein, Darmstadt Printing and Binding: AZ Druck, Kempten (Allgäu)

SPIN: 1155 2352 63/3020 - 5 4 3 2 1 0 – Printed on acid-free paper

Editors

Weber, Horst Technische Universität Berlin, Optisches Institut, Berlin, Germany

Herziger, Gerd Rheinisch-Westfälische Technische Hochschule, Aachen, Germany

Poprawe, Reinhart Fraunhofer-Institut für Lasertechnik (ILT), Aachen, Germany

Authors

Benedict, Mihály G. University of Szeged, Department of Theoretical Physics, Szeged, Hungary

Freyberger, Matthias Universität Ulm, Abteilung Quantenphysik, Ulm, Germany

Haug, Florian Universität Ulm, Abteilung Quantenphysik, Ulm, Germany

Hodgson, Norman Coherent Inc., Santa Clara, CA 95054, USA

Kerwien, Norbert Universität Stuttgart, Institut für Technische Optik, Stuttgart, Germany

Kuhlow, Berndt Technische Universität Berlin, Optisches Institut, Berlin, Germany

Martienssen, Werner Johann Wolfgang Goethe Universität Frankfurt am Main, Physikalisches Institut, Frankfurt am Main, Germany

Paul, Harry Humboldt-Universität zu Berlin, Institut für Physik, Berlin, Germany

Pedrini, Giancarlo Universität Stuttgart, Institut für Technische Optik, Stuttgart, Germany

Ristau, Detlev Laser Zentrum Hannover e.V., Department of Thin Film Technology, Hannover, Germany

Schleich, Wolfgang Peter Universität Ulm, Abteilung Quantenphysik, Ulm, Germany

Scholl, Marcus Verein Deutscher Ingenieure e.V., Kompetenzfeld Optische Technologien, Düsseldorf, Germany

Authors

Tiziani, Hans J. Universität Stuttgart, Institut für Technische Optik, Stuttgart, Germany

Trifonov, Evgeny Dmitrievich Herzen University of Russia, Department of Theoretical Physics, Saint-Petersburg, Russia

Vogel, Karl Universität Ulm, Abteilung Quantenphysik, Ulm, Germany

Landolt-Börnstein

Editorial Office

Gagernstraße 8 D-64283 Darmstadt, Germany fax: +49 (6151) 171760 e-mail: Redaktion.Landolt-Boernstein@springer.com

Internet http://www.landolt-boernstein.com

 \mathbf{VI}

Preface

The three volumes VIII/1A, B, C document the state of the art of "Laser Physics and Applications". Scientific trends and related technological aspects are considered by compiling results and conclusions from phenomenology, observation and experiments. Reliable data, physical fundamentals and detailed references are presented.

In the recent decades the laser source matured to an universal tool common to scientific research as well as to industrial use. Today the main technical goal is the generation of optical power towards shorter wavelengths, shorter pulses, higher efficiency and higher power for applications in science and industry. Tailoring the optical energy in wavelength, space and time is a requirement for the investigation of laser-induced processes, i.e. excitation, non-linear amplification, storage of optical energy, etc. According to the actual trends in laser research and development, Vol. VIII/1 is split into three parts: Vol. VIII/1A with its two subvolumes 1A1 and 1A2 covers laser fundamentals, Vol. VIII/1B deals with laser systems and Vol. VIII/1C gives an overview on laser applications.

In Vol. VIII/1A2 the following topics are treated in detail:

Part 5: Quantum optics

The quantum aspects of the electromagnetic field, represented by the photon, become of increasing importance, not only in scientific research, but also in applications. Examples are the stabilization of lasers and the lower limit of bandwidth, but also the entangled photons, quantum cryptography and their use in information technology.

This part compiles the basic elements of quantum aspects, starting with the quantization of the electromagnetic field, discussing the crucial experiments to prove the Bell inequality and Bose-Einstein condensation. The basic process of the laser, the mechanism of atom-field interaction, is presented in detail.

Part 6: Coherence and superradiance

In Part 6.1 the basic concepts of coherence including photon statistics are presented for classical and non-classical light. The parameters of interest to characterize the degree of temporal and spatial coherence are defined and experimental setups for their measurement are discussed. Part 6.2 deals with superradiance, also known as superfluorescence, amplified spontaneous emissions (ASE), or cooperative spontaneous emission, predicted already 1954 by Dicke. Although long time a phenomenon of more academic interest, it becomes now of importance for X-ray lasers, which operate mainly in the ASE-regime.

Part 7: Optical components

In three parts, the engineering aspects of modulators, thin-film technology, and beam shaping are compiled. For additional optical components see "Linear optics" (Part 3 of Vol. VIII/1A1).

The modulation of light is essential for most applications in data transmission as well as in material processing. The two relevant modulators, the electro-optical and the acousto-optical systems, are discussed in detail in Part 7.1 with many useful parameters in several tables.

Preface

Mirrors, polarizers and beam splitters are elements of major importance for the generation and handling of laser radiation. The quality of optical elements is in many cases the limiting factor for laser efficiency and output intensity. Therefore, thin-film technology plays a key role in laser engineering. In Part 7.2 the basis elements of thin-film systems are presented and the principles of production are compiled. The crucial quality parameters and their measurement are discussed in detail. The extensive references give access to detailed information.

In many cases, especially in material processing, the intensity structure of the laser output field has to be adapted to the special application. This requires beam shaping. Part 7.3 summarizes various elements of low-loss beam transforming with experimental examples.

Part 8: Optical resonators

Laser parameters as beam quality, output power, efficiency, misalignment sensitivity, intensity structure are determined by the optical resonator. In this part all types of optical resonators (stable, unstable, linear devices, ring resonators, wave-guide systems) are compiled together with experimental results and data relevant for a reliable construction. Special emphasis is laid on the real amplifying medium, which strongly influences the resonator by thermal effects and gain saturation.

Part 9: Interferometry

Although a classical field of metrology, it has become of increasing importance with the availability of coherent light sources and is on the way to become a powerful tool in many fields of optics. Part 9 starts with the basic facts of interference and coherence. All relevant interferometer types are described in detail. An extensive survey on the special procedures of measurement techniques including new methods as speckle interferometry, adaptive optics or digital holography is given. Many experimental results complete this part. A comprehensive list of references allows to gather detailed information.

August 2006

The Editors

VIII

Part 5 Quantum optics

5.1	Quantum optics F. Haug, M. Freyberger, K. Vogel, W.P. Schleich	3
5.1.1 5.1.1.1 5.1.1.2	Introduction	${3 \atop {4}}$
$5.1.2 \\ 5.1.2.1 \\ 5.1.2.1.1 \\ 5.1.2.1.2 \\ 5.1.2.1.3 \\ 5.1.2.2 \\ 5.1.2.3 \\ 5.1.2.3 \\$	Field quantization in Coulomb gauge Mode expansion Running waves Standing waves Energy of the radiation field Field quantization Single mode	5 6 6 7 8
5.1.3 5.1.3.1 5.1.3.2 5.1.3.3 5.1.3.4 5.1.3.5 5.1.3.6 5.1.3.6.1 5.1.3.6.2	Field states Pure and mixed states Photon number states Photon number states Coherent states Squeezed states Squeezed states Thermal states Measures of non-classicality Mandel Q-parameter Glauber–Sudarshan distribution Glauber–Sudarshan distribution	8 9 10 11 12 13 13 13
5.1.4 $5.1.4.1$ $5.1.4.2$ $5.1.4.2.1$ $5.1.4.2.2$ $5.1.4.2.3$ $5.1.4.3$ $5.1.4.3.1$ $5.1.4.3.2$	Atom-field interactionElectric field-dipole interactionSimple model for atom-field interactionHamiltonianDynamics of Jaynes-Cummings-Paul modelQuantum motion in an ion trapQuantum state engineeringResonant case: photon number state preparationFar off-resonant case: Schrödinger cat state preparation	$14 \\ 14 \\ 15 \\ 15 \\ 16 \\ 16 \\ 16 \\ 17$
$5.1.5 \\ 5.1.5.1 \\ 5.1.5.1.1 \\ 5.1.5.1.2 \\ 5.1.5.2 \\ 5.1.5.3$	Reservoir theory	18 18 18 18 19 19
5.1.6 5.1.6.1 5.1.6.2	One-atom maser	20 21 21

5.1.7 5.1.7.1 5.1.7.2 5.1.7.3 5.1.8 5.1.8.1	Atom-reservoir interaction Master equation Lamb shift Weisskopf-Wigner decay Resonance fluorescence Model	22 22 23 23 23 24 24
$\begin{array}{c} 5.1.8.2 \\ 5.1.9 \\ 5.1.9.1 \\ 5.1.9.1.1 \\ 5.1.9.1.2 \\ 5.1.9.2 \\ 5.1.9.2.1 \\ 5.1.9.2.2 \\ 5.1.9.2.2 \\ 5.1.9.3 \\ 5.1.9.4 \end{array}$	Spectrum and antibunching. Fundamental questions of quantum mechanics. Quantum jumps. Continuous versus discontinuous dynamics. Experimental observation. Wave-particle duality . Delayed choice experiments . Quantum-optical tests of complementarity Entanglement . Bell inequality .	24 26 26 27 28 28 28 29 30 31
5.1.10 5.1.10.1 5.1.10.2 5.1.10.2.1 5.1.10.2.2 5.1.10.2.3 5.1.10.2.4 5.1.10.3 5.1.10.3.1 5.1.10.3.2	New frontiersAtom optics in quantized fieldsBose–Einstein condensationHistoryBose–Einstein condensation in dilute atomic gasesGross–Pitaevskii equationExperiments with Bose–Einstein condensatesQuantum informationQuantum teleportationQuantum cryptographyReferences for 5.1	32 32 33 34 34 34 35 36 36 36 38 39

Part 6 Coherence and superradiance

6.1	Coherence	
	W. Martienssen, H. Paul	49
6.1.1	Historical remarks	49
6.1.2	Basic concepts	49
6.1.2.1	Classical light	49
6.1.2.2	Non-classical light	50
6.1.2.3	Field mode	50
6.1.2.4	Single-mode field	50
6.1.2.5	Electric field strength	50
6.1.2.6	Interference	50
6.1.2.7	Decoherence	51
6.1.3	Coherence theory	51
6.1.3.1	Field correlation functions	51
6.1.3.1.1	Definitions	51
6.1.3.1.2	Connections between correlation functions of different order	51
6.1.3.2	First-order coherence	52
6.1.3.2.1	Degree of coherence	52
6.1.3.2.2	Temporal coherence	52
6.1.3.2.3	Spatial coherence	53

$\begin{array}{c} 6.1.3.2.4\\ 6.1.3.2.5\\ 6.1.3.3\\ 6.1.3.3.1\\ 6.1.3.3.1.1\\ 6.1.3.3.1.2\\ 6.1.3.3.2\\ 6.1.3.3.2\\ 6.1.3.3.3\\ 6.1.3.4\\ 6.1.3.5\end{array}$	Filtering out coherent light from a chaotic sourceMeasurement of coherence lengthsLaser light versus chaotic lightGenerating mechanisms and radiation characteristicsChaotic lightLaser lightInterference between beams from independent sourcesCoherent interactionParticle interferenceHigher-order coherence	555 55 55 55 55 55 56 56 57 57
$\begin{array}{c} 6.1.4\\ 6.1.4.1\\ 6.1.4.2\\ 6.1.4.3\\ 6.1.4.3.1\\ 6.1.4.4\\ 6.1.4.4.1\\ 6.1.4.4.2\\ 6.1.4.4.2\\ 6.1.4.4.3\\ 6.1.4.5\\ 6.1.4.5.1\\ 6.1.4.5.2\\ 6.1.4.5.2.1\\ 6.1.4.5.2.1\\ 6.1.4.5.3.1\\ 6.1.4.5.3.1\\ 6.1.4.5.3.2\end{array}$	Intensity interferenceFormal descriptionMeasurementSpatial intensity correlationsStellar intensity interferometryTemporal intensity correlationsAmplitude stabilizationPhoton bunchingPhoton antibunchingExperiments with entangled photon pairsParametric down-conversionTwo-photon mixingHong-Ou-Mandel interferometerPhoton-pair interferenceInterferometric devicesFranson experiment	58 58 58 58 59 59 59 59 60 60 60 60 61 61 61
$\begin{array}{c} 6.1.5 \\ 6.1.5.1 \\ 6.1.5.2 \\ 6.1.5.3 \\ 6.1.5.4 \end{array}$	Photon counting statistics Measurement Photon distribution functions Measurements under different experimental conditions Variances of the photon number References for 6.1	62 62 63 63 64
6.2	Superradiance M.G. Benedict, E.D. Trifonov	67
6.2.16.2.26.2.2.16.2.2.2	Definitions and historical layout Superradiance theory Superradiance of a system with dimensions smaller than the radiation wavelength Superradiation of an extended multiatomic system	67 68 68 70
$\begin{array}{c} 6.2.3 \\ 6.2.3.1 \\ 6.2.3.2 \\ 6.2.3.3 \end{array}$	Superradiance experiments	73 74 75 77
6.2.4	Outlook	78 79

Part 7 Optical components

7.1.1 7.1.2	Introduction	85 86 87
7.1.2	Light propagation in crystals. Linear electro-optic effect. Modulator devices Phase modulation	86 87
710	Linear electro-optic effect Modulator devices Phase modulation	87
7.1.3 7.1.3.1 7.1.3.1.1 7.1.3.1.2 7.1.3.1.3 7.1.3.1.4 7.1.3.2 7.1.3.3 7.1.3.3.1 7.1.3.3.1.1 7.1.3.3.1.2 7.1.3.3.1.3 7.1.3.3.1.3 7.1.3.3.2	Polarization modulation (dynamic retardation) Amplitude modulation Design considerations Traveling-wave modulator Examples Crystal classes Crystal class 42m Crystal class 3m Crystal class 43m Selected electro-optic materials and modulator systems	 87 88 89 90 91 91 92 92 92 92 93 94
7.1.3.4	Electro-optic beam deflector	94
7.1.4 7.1.4.1	Kerr electro-optic effect modulators Kerr effect in isotropic media	96 96
7.1.5 7.1.5.1 7.1.5.2 7.1.5.2.1 7.1.5.2.2 7.1.5.2.2.1 7.1.5.2.2.2 7.1.5.2.2.3 7.1.5.2.2.3 7.1.5.2.2.4 7.1.5.3	Acousto-optic modulators	98 98 100 100 101 101 102 102 103 106
7.1.5.4	Acousto-optic deflector	106
7.1.6	Glossary References for 7.1	$\frac{108}{110}$
7.2	Thin-film technology D. RISTAU	111
7.2.1	Introduction	111
7.2.2	Basic principle of optical thin-film systems	111
7.2.3	Production of optical coatings	114
7.2.4	Quality parameters of optical laser components	116
7.2.5 7.2.5.1 7.2.5.2 7.2.5.3 7.2.5.4	Measurement of critical parameters of laser components Calorimetric measurement of absorption Measurement of total scattering Laser-induced damage thresholds	$117 \\ 117 \\ 119 \\ 120 \\ 121$

7.2.6	Examples for advanced laser components 122
7.2.7	Summary and future trends
	References for 7.2
7.3	Beam shaping
	M. Scholl
7.3.1	Introduction
7.3.2	Beam-shaping techniques and design procedures
7.3.2.1	Beam transformation
7.3.2.1.1	Coherent beams
7.3.2.1.2	Partially coherent beams and geometric optic approximation
7.3.2.1.3	Comparison of the geometric-optical solution with the solution from the
	coherent techniques
7.3.2.2	Beam integration
7.3.2.3	Beam integration with beamlet shaping
7.3.2.4	Beam shaping and coherence
7.3.2.5	Beam splitting
7.3.2.6	Manufacturing of beam-shaping elements
7.3.3	Conclusion
	References for 7.3

Part 8 Optical resonators

8.1	Optical resonators N. HODGSON	147
8.1.1	Introduction	147
8.1.1.1	Linear stable resonators	147
8.1.1.2 9.1.1.2	Linear unstable resonators	141
0.1.1.3 8 1 1 <i>1</i>	Wayoguido resonators	140
8.1.1.5	Reviewing the basic properties of all optical resonators	148 148
8.1.2	Classification of optical resonators	149
8.1.3	Unconfined stable resonators	151
8.1.3.1	Transverse mode structures	152
8.1.3.1.1	Gauss–Laguerre and Gauss–Hermite modes	152
8.1.3.1.2	Hybrid modes	156
8.1.3.2	Beam propagation of stable resonator modes	157
8.1.3.2.1	Fundamental mode	157
8.1.3.2.2	Higher-order modes	159
8.1.3.3	Beam quality and resonator parameters	161
8.1.3.4	Resonance frequencies	163
8.1.4	Aperture-limited stable resonators	164
8.1.4.1	Resonators with one aperture	164
8.1.4.2	Resonators with two apertures	166
8.1.5	Misalignment sensitivity	167
8.1.5.1	Fundamental-mode operation	169
8.1.5.2	Multimode operation	170

$\begin{array}{c} 8.1.6\\ 8.1.6.1\\ 8.1.6.1.1\\ 8.1.6.1.2\\ 8.1.6.2\\ 8.1.6.3\\ 8.1.6.3.1\\ 8.1.6.3.2\\ 8.1.6.4\\ 8.1.6.5\end{array}$	Unstable resonators . Beam propagation . Characterization of unstable resonators . Resonator schemes . Mode structures and losses . Beam quality . Circular symmetry . Rectangular symmetry . Unstable resonators with variable-reflectivity mirrors . Applications of unstable resonators .	172 172 172 174 176 178 178 178 179 180 183
8.1.7 8.1.7.1 8.1.7.2 8.1.7.2.1 8.1.7.2.2	Output power of stable resonators Calculation of the output power of stable resonators Optimum output coupling and maximum output power Homogeneous line broadening Inhomogeneous line broadening	184 184 187 188 190
8.1.8 8.1.8.1 8.1.8.2	Thermal lensing in solid-state lasers Fundamental-mode operation Transverse multimode operation	190 191 194
$\begin{array}{c} 8.1.9 \\ 8.1.9.1 \\ 8.1.9.2 \\ 8.1.9.3 \end{array}$	Ring resonators General properties of ring resonators Unstable ring resonators Nonplanar ring resonators	196 196 198 200
$\begin{array}{c} 8.1.10\\ 8.1.10.1\\ 8.1.10.2\\ 8.1.10.3\\ 8.1.10.3.1\\ 8.1.10.3.1.2\\ 8.1.10.3.1.3\\ 8.1.10.3.1.3\\ 8.1.10.3.2\\ 8.1.10.3.3\\ 8.1.10.4\end{array}$	Waveguide resonatorsMotivationEigenmodes of hollow rectangular waveguidesProperties of waveguide resonatorsWaveguide resonator configurationsCase I resonatorsCase II resonatorsCase III resonatorsCalculated round-trip losses of the lowest-loss resonator modeMode properties of a general waveguide resonatorWaveguide resonator mode and loss calculations	201 202 205 206 206 206 206 207 208 208
8.1.10.5	Properties of slab waveguide lasers References for 8.1	$\begin{array}{c} 210\\ 213 \end{array}$

Part 9 Interferometry

9.1	Interferometry H.J. Tiziani, N. Kerwien, G. Pedrini	21
9.1.1	Introduction	21
9.1.2	Basic principles of interference	22
9.1.2.1	Coherence	22
9.1.2.1.1	Temporal coherence	22
9.1.2.1.2	Spatial coherence	23
9.1.2.2	Two-beam interference	23
9.1.2.3	Interference in a plane-parallel plate	24
9.1.2.4	Vector effects of interference	26

9.1.3	Interferometry for optical testing	228
9.1.3.1	Basic interferometer types	228
9.1.3.1.1	Michelson interferometer	228
9.1.3.1.2	I wyman–Green interferometer	229
9.1.3.1.3 0 1 2 1 4	Fizeau interferometer	230
9.1.3.1.4 0 1 3 1 5	Shearing interferometry	230
9.1.3.1.5	Fabry-Perot interferometer	231
9.1.3.1.0	Quantitative electronic phase evaluation techniques	231
91321	The Fourier-transform technique	232
91322	Fringe analysis by phase shifting	232
9.1.3.2.3	Phase-locked interferometry	234
0.1.4	Interference for surface methology	024
9.1.4	Interferometry with extended range and reduced consitivity: Oblique insidence	234
9.1.4.1	interferometry with extended range and reduced sensitivity. Oblique incidence	<u>9</u> 24
91/11	Prismatic interferometer with oblique incidence	234
91419	Grating interferometer with oblique incidence	236
9142	Multiwavelength interferometry	236
9143	White-light interferometry	237
9.1.4.4	Polarization interferometry	239
0.1.5	Heters have been for an here the second distance and the	0.41
9.1.5	Dringing of betangdying interferometry for velocity and distance measurement	241
9.1.5.1	Absolute betonedyne interferometry	241
9.1.9.2	Double heterodyne interferometry (DHI)	2/2
0.1.0	Double neterodyne interretonietry (Din)	240
9.1.6	Interferometry with adaptive optics	244
9.1.6.1	Interferometry with a null corrector	244
9.1.6.2	Adaptive optics with optical light modulator	246
9.1.6.3	Adaptive optics with deformable membrane mirror	247
9.1.0.4	Adaptive optics for optical stitching using dynamically tilted reference wave	249
9.1.7	Speckle pattern interferometry	250
9.1.7.1	Some properties of speckles	251
9.1.7.2	Speckle applications	253
9.1.7.3	Speckle pattern interferometry for deformation measurements	253
9.1.7.4	Phase analysis in speckle interferometry	250
9.1.7.4.1	Phase analysis by phase stepping	200
9.1.7.4.2 0 1 7 4 3	Analysis of vibrating objects	200
9.1.7.4.5	Temporal speckle pattern interferometry (TSPI)	257
91751	Shape measurement with TSPI by using time-varying wavelength change	259
91752	Laser diode with external cavity for wavelength change	259
9.1.7.5.3	Application of TSPI for deformation measurement	260
9.1.7.5.4	Deformation measurements by TSPI and digital holography, a comparison	260
9.1.7.5.5	Vibration measurement with TSPI	262
018	Holographic interferometry	262
0181	Principle of holography	200 263
9182	Principle of holographic interferometry	265
9.1.8.3	Digital holography	266
9.1.8.3.1	Principle of digital holography	266
9.1.8.3.2	Configurations for recording and reconstruction of digital holograms	267
9.1.8.3.2.1	Lensless Fourier hologram	269

9.1.8.3.2.2	Fresnel hologram	270
9.1.8.3.2.3	Image-plane hologram	271
9.1.8.4	Digital holographic interferometry	272
9.1.8.4.1	Principle of digital holographic interferometry	272
9.1.8.4.2	Digital holographic interferometry for dynamic deformations	272
9.1.8.4.2.1	Dimensional measurements	272
9.1.8.4.2.2	Digital holographic interferometry for deformation and vibration analysis of 3D	
	objects	273
9.1.8.4.3	Pulsed digital holographic interferometry for endoscopic investigations	276
9.1.8.4.4	Temporal phase unwrapping of digital holograms	278
	References for 9.1	280
Index		. 285

5.1 Quantum optics

F. HAUG, M. FREYBERGER, K. VOGEL, W.P. SCHLEICH

5.1.1 Introduction

In 1924 Gregor Wentzel – then at the institute of Arnold Sommerfeld in Munich – analyzed [24Wen] the emission and absorption of light by two atoms at different positions. The only quantum theory available to him was the "Atommechanik" à la Bohr–Sommerfeld [25Bor1, 26Pau]. Wentzel was forced to propose a revolutionary idea: Probability amplitudes for the propagation of the light between two points in space are complex-valued phase factors, where the phase is given by the classical action expressed in units of \hbar . The path integral formulation of quantum mechanics was born not only before Richard Feynman, but even before quantum mechanics was discovered by Werner Heisenberg and Erwin Schrödinger.

Obviously, Wentzel's paper contains a remarkable approach. However, even more remarkable is the title of his paper: "Zur Quantenoptik", that is "On Quantum Optics". Therefore, Wentzel might well be considered the father of the expression "quantum optics" despite the fact that he used it in a completely different context. Nowadays we interpret the field of quantum optics [89Coh, 91Lou, 91Mey, 92Coh, 94Wal, 95Man, 96Scu, 99Pau, 01Sch] as the domain of optics where quantum mechanics is of importance, that is where the quantization of the radiation field is crucial.

5.1.1.1 A brief history of quantum optics

The electromagnetic field appears almost everywhere in physics. Historically it has played a major role in the development of our understanding of the physical laws of nature. First James Clerk Maxwell [1873Max] in 1864 found the famous system of equations which explains the whole classical realm of electromagnetic field effects. At the beginning of the last century Max Planck initiated quantum theory [1900Pla, 1901Pla] when he discovered the fundamental constant $h \equiv 2\pi\hbar$ in the laws of black-body radiation. In 1905 Albert Einstein explained the photoelectric effect [1905Ein] on the hypothesis of a corpuscular nature of radiation and in 1917 this paradigm helped him to describe the interaction between atoms and electromagnetic radiation introducing the coefficients of spontaneous and induced emission [17Ein].

However, none of this work had been based on a rigorous quantum theory of the electromagnetic field. Such a formalism was put forward later in the paper by Max Born and Pascual Jordan [25Bor3] in which they point out that Heisenberg's non-commuting objects were matrices and in the famous "Drei-Männer-Arbeit" by Born, Heisenberg and Jordan [25Bor2]. Nevertheless, the starting point of quantum optics might well be traced back to the article by Paul Adrian Maurice Dirac [27Dir] in 1927 in which he developed the quantum theory of the emission and absorption of radiation. An independent but equivalent approach was put forward by Enrico Fermi. For a beautiful exposition of the quantum theory of radiation and a wealth of applications we refer to his Ann Arbor summer school contribution [32Fer]. From this theory emerged quantum electrodynamics (QED). It became

the role model of all gauge theories of particle physics. In some sense quantum optics is the non-relativistic limit of QED.

In the mid-fifties Robert Hanbury Brown and Richard Twiss – working in the field of radio astronomy – were trying to determine the angular diameters of radio sources. They recognized that in these measurements radiation fluctuations play an essential role. In order to overcome these limitations they developed a new type of intensity interferometer [56Han]. This experiment was performed many years before the laser and represents the starting point for the theory of photon detection and coherence created by Roy J. Glauber at Harvard [63Gla1, 63Gla2].

The development of the laser [83Ber, 99Lam] brought a renaissance to the quantum theory of radiation. Indeed, all the tools of quantum optics were created in the framework of the quantum theory of the laser. Its main proponents were Hermann Haken and his school in Stuttgart [84Hak], Willis E. Lamb and Marlan O. Scully at Yale [74Sar], Melvin Lax [66Lax] and William H. Louisell [73Lou] at Bell Laboratories. Unfortunately, ordinary lasers do not show that many quantum effects. Therefore, the application of theoretical apparatus completed already during the early days of the laser – the high time of quantum electronics – had to wait. For a collection of some of the pioneering papers of this pre-quantum optics era we refer to [70Man].

The detection of photon antibunching [77Kim, 78Kim, 82Cre] and the Mollow-triplet [69Mol] in resonance fluorescence [74Sch, 75Wu, 76Har] ushered in the era of quantum optics. Finally, clean quantum manifestations of the radiation field had been detected and the theoretical tools developed almost 20 years earlier could be put to use. The subsequent observation of squeezed states [85Slu], the development of the one-atom maser [85Mes] and the controlled creation of entangled states capitalizing on the tools of nonlinear optics have emancipated quantum optics. Today it embraces topics such as Bose–Einstein condensation, degenerate fermionic quantum gases, quantum computing, quantum information and quantum communication.

It has become impossible to give a complete overview over the field in a few pages. Therefore, we have decided to concentrate on a few themes that define quantum optics. Space does not allow us to present detailed derivations. Nevertheless we try to outline their logic.

5.1.1.2 Outline of the review

Nowadays we gain more and more insight into the quantum aspects of the electromagnetic field. To this end we must understand the quantization of the field which leads to the concept of a quantum of radiation, the so-called "photon". The name "photon" was coined [95Lam] in 1926 by the chemist Gilbert N. Lewis when he speculated that the transmission of radiation from one atom to another was carried by a new particle [26Lew]. He specifically denied that this particle was Einstein's light quantum. The word "photon" caught on, but not Lewis' meaning. With this motivation in mind we summarize in Sect. 5.1.2 the essential ideas of field quantization in Coulomb gauge.

Specific non-classical features of the electromagnetic field arise for certain quantum states, like squeezed states. Modern quantum optics investigates characteristics which allow us to draw the border line between classical and non-classical field effects. We therefore review in Sect. 5.1.3 important states of radiation fields. Here coherent states play a prominent role since they are the building blocks of Glauber's coherence theory discussed in Sect. 6.1.3 of this volume.

In Sect. 5.1.4 we then turn to the interaction of an atom and a quantized electromagnetic field. Here we focus on the most elementary model of a two-level atom interacting with a single mode of the radiation field. We show how this model describes the experimental realization of photon number states and Schrödinger cat states.

Section 5.1.5 is devoted to a brief introduction into reservoir theory. Here we consider the interaction of a single quantum system – the master system – with a large ensemble of quantum systems which constitute the reservoir. We are only interested in the dynamics of the master system

and eliminate the reservoir. This approach allows us to derive the master equation for the damping of a cavity mode which has many applications in the context of decoherence [91Zur, 96Giu] and the emergence of the classical world.

The amazing system of a one-atom maser is the topic of Sect. 5.1.6. In contrast to the familiar laser or maser we achieve in this device microwave amplification by stimulated emission due to a single atom. The emitted radiation displays many non-classical features.

The difference between a classical field and a quantized field manifests itself in the interaction of the radiation field with an atom. A classical field can have zero amplitude. Obviously in this case it does not interact with the atom. On the other hand a quantized field *always* interacts with the atom, even if all the field modes are in their ground states. Then there are still vacuum fluctuations present which lead to various effects such as spontaneous emission and Lamb shift as discussed in Sect. 5.1.7.

The light emitted by a two-level atom driven by a classical field exhibits many interesting features such as photon antibunching and squeezing, as briefly summarized in Sect. 5.1.8.

Section 5.1.9 sketches a few quantum-optical experiments addressing fundamental questions of quantum mechanics. In particular, we review the debate on quantum jumps and discuss the wave-particle duality. We finally turn to the topics of entanglement and experimental verification of Bell inequalities. These tests rely on entangled photon pairs. For a review of other experiments using these unique tools of quantum optics we refer to Sect. 6.1.4.5 of this volume.

We conclude in Sect. 5.1.10 by turning to a few new frontiers of quantum optics. Here we highlight three timely topics: atom optics in quantized fields, Bose–Einstein condensation and quantum information represented by teleportation and quantum cryptography.

5.1.2 Field quantization in Coulomb gauge

In this section we lay the foundations for our review of quantum optics by quantizing the electromagnetic field [89Coh]. We expand the field in a complete set of normal modes which reduces the problem of field quantization to the quantization of a one-dimensional harmonic oscillator corresponding to each normal mode.

5.1.2.1 Mode expansion

The classical free electromagnetic field, that is the field in a region without charge and current densities, obeys the Maxwell equations (see also Vol. VIII/1A1, (1.1.4)-(1.1.7) with (1.1.8)-(1.1.9))

$$\nabla \times \boldsymbol{E} + \frac{\partial \boldsymbol{B}}{\partial t} = 0, \qquad \nabla \cdot \boldsymbol{B} = 0, \qquad \nabla \times \boldsymbol{H} - \frac{\partial \boldsymbol{D}}{\partial t} = 0, \qquad \nabla \cdot \boldsymbol{D} = 0$$
 (5.1.1)

expressed in the SI system with $B = \mu_0 H$ and $D = \varepsilon_0 E$. The magnetic permeability μ_0 connects the magnetic induction B with the magnetic field H and the electric permittivity ε_0 of free space connects the displacement D with the electric field E. In the case of a free field E and B may be obtained from

$$\boldsymbol{B} = \nabla \times \boldsymbol{A}, \quad \boldsymbol{E} = -\frac{\partial \boldsymbol{A}}{\partial t}, \quad (5.1.2)$$

where the vector potential \boldsymbol{A} obeys the Coulomb gauge condition $\nabla \cdot \boldsymbol{A} = 0$ and satisfies a wave equation.

Due to the linear structure of Maxwell's equations we can decompose the vector potential into normal modes according to

$$\boldsymbol{A}(\boldsymbol{r},t) = \sum_{\ell} \left(\frac{\hbar}{2\Omega_{\ell}\varepsilon_0 V_{\ell}}\right)^{1/2} \left(\boldsymbol{u}_{\ell}(\boldsymbol{r})\alpha_{\ell} \mathrm{e}^{-\mathrm{i}\Omega_{\ell}t} + \mathrm{c.c.}\right)$$
(5.1.3)

with complex mode amplitudes α_{ℓ} .

The mode functions $u_{\ell}(r)$ are solutions of the Helmholtz equation

$$\Delta \boldsymbol{u}_{\ell}(\boldsymbol{r}) + \frac{\Omega_{\ell}^2}{c^2} \boldsymbol{u}_{\ell}(\boldsymbol{r}) = 0$$
(5.1.4)

with the mode frequencies Ω_{ℓ} and the speed of light *c*. They have to fulfill the Coulomb gauge $\nabla \cdot \boldsymbol{u}_{\ell} = 0$. We choose \boldsymbol{u}_{ℓ} to be dimensionless and scale it such that the maximum of $|\boldsymbol{u}_{\ell}|^2$ is one. This choice allows us to define the effective mode volume V_{ℓ} of the mode ℓ via the orthogonality relation

$$\int \mathrm{d}V \boldsymbol{u}_{\ell}^{*}(\boldsymbol{r}) \boldsymbol{u}_{\ell'}(\boldsymbol{r}) = \delta_{\ell,\ell'} V_{\ell} \,. \tag{5.1.5}$$

Here we integrate over the region of space where the electromagnetic field and therefore u_{ℓ} is nonzero. The frequencies Ω_{ℓ} of the modes ℓ follow from the boundary conditions for u_{ℓ} . We consider two cases: running waves in free space and standing waves in a ideal resonator.

5.1.2.1.1 Running waves

We start by discussing periodic boundary conditions for the mode functions u_{ℓ} in the quantization volume V. In this case we can solve the Helmholtz equation (5.1.4) in terms of plane waves. Therefore, with these mode functions the vector potential reads

$$\boldsymbol{A}(\boldsymbol{r},t) = \sum_{\boldsymbol{k}} \sum_{\sigma=1}^{2} \left(\frac{\hbar}{2\Omega_{\boldsymbol{k}}\varepsilon_{0}V} \right)^{1/2} \left(\alpha_{\boldsymbol{k}\sigma}\boldsymbol{\epsilon}_{\boldsymbol{k}\sigma} \mathrm{e}^{\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{r}-\Omega_{\boldsymbol{k}}t)} + \mathrm{c.c.} \right) , \qquad (5.1.6)$$

where the effective volume V_{ℓ} of each mode is identical to the quantization volume V.

Instead of the index ℓ we now use a wave vector \mathbf{k} and two polarization vectors $\boldsymbol{\epsilon}_{k1}$ and $\boldsymbol{\epsilon}_{k2}$ to characterize the modes. The gauge condition $\nabla \cdot \mathbf{A} = 0$ implies that these polarization vectors have to be orthogonal to the wave vector \mathbf{k} , that is $\boldsymbol{\epsilon}_{k1} \cdot \mathbf{k} = \boldsymbol{\epsilon}_{k2} \cdot \mathbf{k} = 0$ for each wave vector \mathbf{k} . The frequencies $\Omega_{\mathbf{k}}$ are related to the wave vectors \mathbf{k} via the dispersion relation $\Omega_{\mathbf{k}} = c|\mathbf{k}|$. The periodicity condition leads to a discrete set of wave vectors $\mathbf{k} \equiv 2\pi/V^{1/3}(n_x, n_y, n_z)$, where $n_x, n_y, n_z = 0, \pm 1, \pm 2 \dots$ The Fourier amplitudes $\alpha_{\mathbf{k}\sigma}$ are complex numbers in the classical theory. With the help of (5.1.2) we can decompose \mathbf{E} and \mathbf{B} into normal modes and find

$$\boldsymbol{E}(\boldsymbol{r},t) = i \sum_{\boldsymbol{k}} \sum_{\sigma=1}^{2} \left(\frac{\hbar \Omega_{\boldsymbol{k}}}{2\varepsilon_{0} V} \right)^{1/2} \left(\alpha_{\boldsymbol{k}\sigma} \boldsymbol{\epsilon}_{\boldsymbol{k}\sigma} e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\Omega_{\boldsymbol{k}}t)} - \text{c.c.} \right)$$
(5.1.7)

for the electric field and a similar expression for the magnetic induction B.

5.1.2.1.2 Standing waves

In the context of cavity QED the mode functions are determined by the resonator. At the boundary of an ideal cavity the tangential component of the electric field E and the normal component of the

magnetic induction B have to vanish. Both conditions are fulfilled when the tangential component of each mode function u_{ℓ} vanishes at the boundary of the resonator. For open resonators we require in addition that the electromagnetic field and therefore u_{ℓ} vanishes at infinity. These boundary conditions determine the frequency Ω_{ℓ} of the mode ℓ . Furthermore, it is convenient to use real mode functions u_{ℓ} . For the vector potential we then find

$$\boldsymbol{A}(\boldsymbol{r},t) = \sum_{\ell} \left(\frac{\hbar}{2\Omega_{\ell}\varepsilon_{0}V_{\ell}}\right)^{1/2} \boldsymbol{u}_{\ell}(\boldsymbol{r}) \left(\alpha_{\ell}\mathrm{e}^{-\mathrm{i}\Omega_{\ell}t} + \mathrm{c.c.}\right)$$
(5.1.8)

and the electric field takes the form

$$\boldsymbol{E}(\boldsymbol{r},t) = \mathrm{i} \sum_{\ell} \left(\frac{\hbar \Omega_{\ell}}{2\varepsilon_0 V_{\ell}} \right)^{1/2} \boldsymbol{u}_{\ell}(\boldsymbol{r}) \left(\alpha_{\ell} \mathrm{e}^{-\mathrm{i}\Omega_{\ell} t} - \mathrm{c.c.} \right) \,.$$
(5.1.9)

From (5.1.2) we find a similar relation for the magnetic induction **B**.

5.1.2.1.3 Energy of the radiation field

With the help of (5.1.9) for the electric field and the analogous expression for the magnetic field we can express the field energy

$$H_{\rm f} \equiv \frac{1}{2} \int \mathrm{d}V \left(\varepsilon_0 \boldsymbol{E}^2(\boldsymbol{r}, t) + \boldsymbol{B}^2(\boldsymbol{r}, t)/\mu_0\right)$$
(5.1.10)

in terms of the mode amplitudes α_{ℓ} and α^*_{ℓ} and arrive at

$$H_{\rm f} = \sum_{\ell} \frac{\hbar \Omega_{\ell}}{2} \left(\alpha_{\ell} \alpha_{\ell}^* + \alpha_{\ell}^* \alpha_{\ell} \right) \,, \tag{5.1.11}$$

where we have already anticipated that in the next section the mode amplitudes become noncommuting operators.

5.1.2.2 Field quantization

The mode decomposition of the electromagnetic field discussed in the previous sections allows us to quantize the field. In the Schrödinger picture we have to replace the time-dependent amplitudes $\alpha_{\ell} e^{-i\Omega_{\ell}t}$ by the time-independent mode annihilation operators \hat{a}_{ℓ} . The complex conjugates $\alpha_{\ell}^* e^{i\Omega_{\ell}t}$ are replaced by the time-independent mode creation operators \hat{a}_{ℓ}^{\dagger} . These operators obey the commutation relations

$$[\hat{a}_{\ell}, \hat{a}_{\ell'}^{\dagger}] = \delta_{\ell\ell'} \,. \tag{5.1.12}$$

The representations of the vector potential, the electric field and the magnetic induction in terms of these operators follows from (5.1.6)–(5.1.9). For running waves we find

$$\hat{A}(\mathbf{r}) = \sum_{\mathbf{k}} \sum_{\sigma=1}^{2} \left(\frac{\hbar}{2\varepsilon_0 \Omega_{\mathbf{k}} V} \right)^{1/2} \left(\hat{a}_{\mathbf{k}\sigma} \boldsymbol{\epsilon}_{\mathbf{k}\sigma} \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}} + \mathrm{h.c.} \right) , \qquad (5.1.13)$$

$$\hat{\boldsymbol{E}}(\boldsymbol{r}) = i \sum_{\boldsymbol{k}} \sum_{\sigma=1}^{2} \left(\frac{\hbar \Omega_{\boldsymbol{k}}}{2\varepsilon_0 V} \right)^{1/2} \left(\hat{a}_{\boldsymbol{k}\sigma} \boldsymbol{\epsilon}_{\boldsymbol{k}\sigma} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} - h.c. \right) \equiv \hat{\boldsymbol{E}}^+(\boldsymbol{r}) + \hat{\boldsymbol{E}}^-(\boldsymbol{r}), \qquad (5.1.14)$$

$$\hat{A}(\boldsymbol{r}) = \sum_{\ell} \left(\frac{\hbar}{2\varepsilon_0 \Omega_{\ell} V_{\ell}} \right)^{1/2} \boldsymbol{u}_{\ell}(\boldsymbol{r}) \left(\hat{a}_{\ell} + \hat{a}_{\ell}^{\dagger} \right) \,, \tag{5.1.15}$$

$$\hat{\boldsymbol{E}}(\boldsymbol{r}) = i \sum_{\ell} \left(\frac{\hbar \Omega_{\ell}}{2\varepsilon_0 V_{\ell}} \right)^{1/2} \boldsymbol{u}_{\ell}(\boldsymbol{r}) \left(\hat{a}_{\ell} - \hat{a}_{\ell}^{\dagger} \right) \equiv \hat{\boldsymbol{E}}^{+}(\boldsymbol{r}) + \hat{\boldsymbol{E}}^{-}(\boldsymbol{r}) \,.$$
(5.1.16)

In the last step we have decomposed the electric field operator into the positive and negative frequency parts \hat{E}^+ and \hat{E}^- , respectively. A similar relation holds for the operator describing the magnetic induction \hat{B} .

The expression (5.1.11) for the field energy and the commutation relations (5.1.12) allow us to write the Hamiltonian

$$\hat{H}_{\rm f} = \sum_{\ell} \hbar \Omega_{\ell} (\hat{a}_{\ell}^{\dagger} \hat{a}_{\ell} + 1/2) \tag{5.1.17}$$

as a sum of independent harmonic oscillator Hamiltonians corresponding to each mode $\ell.$

5.1.2.3 Single mode

Since normal modes are independent from each other, we can often restrict the discussion to a single mode and suppress the mode index ℓ . The corresponding Hamiltonian then reads

$$\hat{H}_{\rm f} = \hbar \Omega (\hat{a}^{\dagger} \hat{a} + 1/2) = \hbar \Omega \left(\frac{1}{2} \hat{p}^2 + \frac{1}{2} \hat{x}^2 \right) \,, \tag{5.1.18}$$

where the quadrature operators

$$\hat{x} \equiv \frac{1}{\sqrt{2}}(\hat{a} + \hat{a}^{\dagger}) \quad \text{and} \quad \hat{p} \equiv \frac{1}{i\sqrt{2}}(\hat{a} - \hat{a}^{\dagger})$$

$$(5.1.19)$$

are equivalent to scaled position and momentum operators \hat{x} and \hat{p} of a massive particle in a harmonic potential. Since they obey the commutation relation $[\hat{x}, \hat{p}] = i$, their uncertainties $(\Delta x)^2 \equiv \langle \hat{x}^2 \rangle - \langle \hat{x} \rangle^2$ and $(\Delta p)^2 \equiv \langle \hat{p}^2 \rangle - \langle \hat{p} \rangle^2$ fulfill the Heisenberg inequality $\Delta x \Delta p \ge 1/2$.

We conclude by emphasizing that the quantum nature of the electromagnetic field enters through the amplitude operators \hat{a}_{ℓ} and \hat{a}_{ℓ}^{\dagger} . It is therefore not surprising that the components of the electric and the magnetic field do not commute. As a consequence they cannot be measured simultaneously with arbitrary accuracy. The uncertainty relations for the electric and magnetic field components were derived for the first time by Niels Bohr and Léon Rosenfeld [33Boh, 50Boh]. For a more detailed discussion we refer to [73Lou].

5.1.3 Field states

When we quantize a system the step of representing observables by operators is only one side of the coin. To describe the state of the system by a quantum state constitutes the other side.

In this section we briefly summarize important properties of several states of the electromagnetic field. Space does not allow us to go into great detail nor can we do justice to the wealth of quantum states. For our review we have chosen number states, coherent states, squeezed states and thermal

states. For a more complete treatment we refer to the literature [91Mey, 94Wal, 95Man, 96Scu, 01Sch]. For a discussion of unusual states such as phase states we recommend [68Car, 89Peg, 93Sch]. In later sections we describe how these states have been created in experiments.

Before we discuss those states in detail we briefly review the state concept of quantum mechanics.

5.1.3.1 Pure and mixed states

The state vector $|\psi\rangle$ contains the full information about a quantum system. However, in many cases we do not know every detail of our system. This lack of information could be for example because the system has too many degrees of freedom. Another reason could be due to the fact that our system is coupled to a reservoir and we cannot keep track of the motion of the individual constituents. A damped cavity field, discussed in Sect. 5.1.5, or an atom that undergoes spontaneous emission, analyzed in Sect. 5.1.7, represent two such systems coupled to a reservoir. In both cases we cannot describe the system by a state vector but by a density operator.

The density operator $\hat{\rho}$ of a pure state $|\psi\rangle$ reads

$$\hat{\rho} \equiv |\psi\rangle\langle\psi|. \tag{5.1.20}$$

A mixed state is an incoherent superposition

$$\hat{\rho} = \sum_{n} W_n |\psi_n\rangle \langle\psi_n| \tag{5.1.21}$$

of pure states $|\psi_n\rangle$ with probabilities W_n such that $\sum_n W_n = 1$.

The expectation value $\langle \hat{\mathcal{O}} \rangle$ of the observable $\hat{\mathcal{O}}$ then involves two averages, namely the quantummechanical average $\langle \psi_n | \hat{\mathcal{O}} | \psi_n \rangle$ and the one over the statistics W_n of the states $| \psi_n \rangle$, that is

$$\langle \hat{\mathcal{O}} \rangle = \sum_{n} W_n \langle \psi_n | \hat{\mathcal{O}} | \psi_n \rangle = \operatorname{Tr} \left(\hat{\mathcal{O}} \hat{\rho} \right) \,. \tag{5.1.22}$$

In the last step we have introduced the operation $\operatorname{Tr} \hat{A} \equiv \sum_{i} \langle i | \hat{A} | i \rangle$ of the trace of an operator \hat{A} , where we sum over an arbitrary complete set of states $|i\rangle$.

In the language of the trace the normalization condition of the probabilities W_n reads Tr $\hat{\rho} = 1$. Moreover, we find the inequality

$$\operatorname{Tr} \hat{\rho}^2 \le 1. \tag{5.1.23}$$

Since the equal sign arises only for pure states, the calculation of $\text{Tr} \hat{\rho}^2$ allows to identify a pure state from the density operator.

5.1.3.2 Photon number states

We start our walk through the gallery of single-mode field states by considering the eigenstates $|n\rangle$ of the number operator $\hat{n} \equiv \hat{a}^{\dagger}\hat{a}$. The corresponding eigenvalue equation reads

$$\hat{n}|n\rangle = \hat{a}^{\dagger}\hat{a}|n\rangle = n|n\rangle, \qquad (5.1.24)$$

where n = 0, 1, 2, ... denotes the number of excitations or the number of photons in the mode. The states $|n\rangle$ also carry the name Fock states.

The state $|0\rangle$, that is the vacuum state of the mode, is defined by $\hat{a}|0\rangle = 0$. We can climb the ladder of excitations up and down via the application of the creation and the annihilation operators \hat{a}^{\dagger} and \hat{a} on a Fock state $|n\rangle$ which yields

$$\hat{a}^{\dagger}|n\rangle = \sqrt{n+1} |n+1\rangle, \quad \hat{a}|n\rangle = \sqrt{n} |n-1\rangle.$$
(5.1.25)

Number states form a complete and orthonormal set of states and therefore provide the frequently used representation

$$|\psi\rangle = \sum_{n=0}^{\infty} \langle n |\psi\rangle |n\rangle \equiv \sum_{n=0}^{\infty} \psi_n |n\rangle$$
(5.1.26)

of a pure quantum state $|\psi\rangle$, or a mixed quantum state

$$\hat{\rho} = \sum_{m,n=0}^{\infty} \langle m | \hat{\rho} | n \rangle | m \rangle \langle n | \equiv \sum_{m,n=0}^{\infty} \rho_{mn} | m \rangle \langle n |$$
(5.1.27)

described by a density operator $\hat{\rho}$.

In Sect. 5.1.4.3 we briefly describe an experiment that has created in a controlled way photon number states in a cavity. A different method uses the photon pair in two-photon down conversion [99Zei].

5.1.3.3 Coherent states

In an attempt to show that there exist non-spreading wave packets in quantum mechanics Erwin Schrödinger [26Sch2] in 1926 studied the time evolution of a displaced ground state of a harmonic oscillator. He found that indeed the Gaussian probability density of the wave packet maintains its shape. In the concluding sentences of his paper he even expresses his optimism that similar wave packets may also be found for the hydrogen atom. Unfortunately, no such states exist since the spectrum of the hydrogen atom is not equidistant.

Almost 40 years later, Roy J. Glauber [63Gla1, 63Gla2] revisited these special states which he called "coherent states". They are the building blocks for the formalism of photon detection and coherence theory. For a summary of these topics we refer to [65Gla].

According to Glauber a coherent state $|\alpha\rangle$ is an eigenstate of the annihilation operator

$$\hat{a}|\alpha\rangle = \alpha|\alpha\rangle \tag{5.1.28}$$

with the complex amplitude $\alpha \equiv |\alpha| e^{i\theta}$.

An alternative representation of a coherent state is by the action of the displacement operator $\hat{D}(\alpha)$ on the vacuum, that is

$$|\alpha\rangle \equiv e^{\alpha \hat{a}^{\dagger} - \alpha^{*} \hat{a}} |0\rangle \equiv \hat{D}(\alpha) |0\rangle = e^{-|\alpha|^{2}/2} \sum_{n=0}^{\infty} \frac{\alpha^{n}}{\sqrt{n!}} |n\rangle.$$
(5.1.29)

Hence, the photon distribution

$$W_n \equiv |\langle n | \alpha \rangle|^2 = \frac{|\alpha|^{2n} e^{-|\alpha|^2}}{n!}$$
(5.1.30)

of a coherent state is a Poisson distribution with average photon number $\langle \hat{n} \rangle = |\alpha|^2$ and variance $(\Delta n)^2 \equiv \langle \hat{n}^2 \rangle - \langle \hat{n} \rangle^2 = |\alpha|^2$. Consequently, the relative fluctuations $(\Delta n)/\langle \hat{n} \rangle = \langle \hat{n} \rangle^{-1/2}$ vanish for a large average photon number.



Fig. 5.1.1. In its most elementary version the quantummechanical superposition of two coherent states of average photon number $\langle \hat{n} \rangle = \alpha^2$ and phase difference 2φ can be visualized by two circles of radius unity displaced by an amount α from the origin and having the angle φ between them and the real axis.

A coherent state is an eigenstate of \hat{a} . However, a superposition [91Sch]

$$|\psi_{\text{cat}}\rangle \equiv \mathcal{N} \left| |\alpha e^{i\varphi}\rangle + |\alpha e^{-i\varphi}\rangle \right| \tag{5.1.31}$$

of two coherent states of identical real-valued amplitudes α but different phases φ and $-\varphi$ depicted in Fig. 5.1.1 is not an eigenstate of \hat{a} . Such a state with an appropriate normalization factor \mathcal{N} is an example of a so-called Schrödinger cat state which was first introduced in Schrödinger's metaphor [35Sch].

In Sect. 5.1.4.3 we discuss a cavity QED experiment and mention an ion trap experiment which describes the preparation of such a state. Moreover, in Sect. 5.1.5.3 we explain why these states provide deeper insight into the emergence of the classical world from quantum mechanics.

5.1.3.4 Squeezed states

We obtain a squeezed state $|\alpha, \epsilon\rangle$ by applying the displacement operator $\hat{D}(\alpha)$ and the unitary squeeze operator

$$\hat{S}(\epsilon) \equiv e^{\frac{1}{2}\epsilon^* \hat{a}^2 - \frac{1}{2}\epsilon \hat{a}^{\dagger 2}} \tag{5.1.32}$$

with $\epsilon \equiv r e^{-2i\phi}$ to the vacuum

$$|\alpha, \epsilon\rangle \equiv D(\alpha)S(\epsilon)|0\rangle. \tag{5.1.33}$$

The nature of squeezed states stands out in the uncertainties $(\Delta X_1)^2 \equiv \langle \hat{X}_1^2 \rangle - \langle \hat{X}_1 \rangle^2$ and $(\Delta X_2)^2 \equiv \langle \hat{X}_2^2 \rangle - \langle \hat{X}_2 \rangle^2$ for the rotated quadratures

$$\hat{X}_1 \equiv \hat{x} \cos \phi - \hat{p} \sin \phi, \quad \hat{X}_2 \equiv \hat{p} \cos \phi + \hat{x} \sin \phi, \tag{5.1.34}$$

where the ordinary quadrature operators \hat{x} and \hat{p} are defined in (5.1.19). Indeed, we find

$$\Delta X_1 = e^{-r} / \sqrt{2}, \quad \Delta X_2 = e^r / \sqrt{2}.$$
 (5.1.35)

A squeezed state is a minimum uncertainty state since the product $\Delta X_1 \cdot \Delta X_2 = 1/2$ saturates the Heisenberg uncertainty relation. For r = 0 we have a coherent state. In this case the two



Fig. 5.1.2. Photon statistics W_n of a squeezed state for different choices of the squeeze parameter $s \equiv e^r$. All curves are plotted for the same value $\alpha = 7$ of the displacement parameter. The rearmost curve (no squeezing at all, s = 1) shows the ideal Poisson distribution associated with a coherent state. Curves that are further forward display oscillations in the photon statistics.

uncertainties $\Delta X_1 = \Delta X_2 = 1/\sqrt{2}$ are equal. However, for $r \neq 0$ the fluctuations in one quadrature are squeezed at the expense of the other. For r > 0 the fluctuations in the variable X_1 are squeezed, whereas for r < 0 the fluctuations in X_2 are squeezed. The angle ϕ indicates a rotation of the coordinate system. For $\phi = 0$ the squeezed state $|\alpha, r\rangle$ is a minimum uncertainty state for the quadratures \hat{x} and \hat{p} with $\Delta x = e^{-r}/\sqrt{2}$ and $\Delta p = e^r/\sqrt{2}$.

The photon number distribution $W_n \equiv |\langle \alpha, r | n \rangle|^2$ can be narrower or broader than the one for the corresponding coherent state with the same α . This narrowing of the photon distribution, that is the sub-Poissonian behavior, is one of the non-classical features of a squeezed state. Furthermore, W_n displays oscillations [87Sch] for larger squeezing as shown in Fig. 5.1.2.

Several experiments have demonstrated the generation of squeezed light making use of nonlinear optics. For more details we refer to the special issues [87Kim, 87Lou, 92Gia].

5.1.3.5 Thermal states

The most elementary example of a density operator $\hat{\rho}$ is that of a thermal state for an oscillator with frequency Ω . This density operator corresponds to a state for which we only know that it obeys a Boltzmann photon statistics $W_n = (1 - e^{-\beta}) e^{-n\beta}$, where $\beta \equiv \hbar \Omega / (k_{\rm B}T)$. Here $k_{\rm B}$ and T denote the Boltzmann constant and the temperature corresponding to the thermal state, respectively. In this case the density operator reads

$$\hat{\rho}_{\rm th} \equiv \sum_{n=0}^{\infty} W_n |n\rangle \langle n| = (1 - e^{-\beta}) \sum_{n=0}^{\infty} e^{-n\beta} |n\rangle \langle n|.$$
(5.1.36)

The parameter β and therefore the temperature T is intimately related to the average number of thermal photons

$$n_{\rm th} = \text{Tr}\,(\hat{n}\hat{\rho}_{\rm th}) = \sum_{n=0}^{\infty} nW_n = \frac{e^{-\beta}}{1 - e^{-\beta}} = \frac{1}{\exp\left(\frac{\hbar\Omega}{k_{\rm B}T}\right) - 1}\,.$$
(5.1.37)

When we express the Boltzmann factor $e^{-\beta}$ by $n_{\rm th}$ we find the expression

$$\hat{\rho}_{\rm th} = \frac{1}{n_{\rm th} + 1} \sum_{n=0}^{\infty} \left(\frac{n_{\rm th}}{n_{\rm th} + 1} \right)^n |n\rangle \langle n|$$
(5.1.38)

for the density operator of a thermal state.

5.1.3.6 Measures of non-classicality

One of our goals in this review is to highlight physical effects originating from the quantization of the radiation field. For this purpose we focus on states that cannot be described by a classical field.

5.1.3.6.1 Mandel Q-parameter

A convenient indicator of a non-classical field is the Mandel Q-parameter [95Man]

$$Q \equiv \frac{(\Delta n)^2 - \langle \hat{n} \rangle}{\langle \hat{n} \rangle} \equiv \sigma - 1, \qquad (5.1.39)$$

which is closely related to the normalized variance $\sigma \equiv (\Delta n)^2 / \langle \hat{n} \rangle$ of the photon distribution.

For a coherent state we find with $\langle \hat{n} \rangle = \alpha^2$ and $(\Delta n)^2 = \alpha^2$ the Mandel parameter Q = 0. A field in a coherent state is considered to be closest to a classical field since it saturates the Heisenberg uncertainty relation and has the same uncertainty in each quadrature component. Therefore, Q = 0defines a boundary between a classical and a quantum field. Indeed, for a thermal state we find $(\Delta n)^2 = n_{\rm th}^2 + n_{\rm th}$, corresponding to a photon distribution broader than a Poissonian and hence $Q = n_{\rm th} > 0$. For Q < 0 the photon distribution becomes narrower than that of a Poissonian. The corresponding state is non-classical. The most elementary examples of non-classical states are number states. Since they are eigenstates of the photon number operator \hat{n} the fluctuations in \hat{n} vanish and the Mandel Q-parameter reads Q = -1. A squeezed state can display a photon distribution that is either broader or narrower than a Poissonian leading a Mandel Q-parameter that is either positive or negative.

5.1.3.6.2 Glauber–Sudarshan distribution

Another indicator of a non-classical field state is the Glauber–Sudarshan distribution P. It arises in the expansion of the density operator into coherent states. Indeed, coherent states form an overcomplete set. Hence states and operators can be expanded into coherent states. In particular, the density operator $\hat{\rho}$ can be expressed in the diagonal form

$$\hat{\rho} = \int d^2 \alpha P(\alpha) |\alpha\rangle \langle \alpha|, \qquad (5.1.40)$$

where $P(\alpha)$ denotes the Glauber–Sudarshan phase space distribution [63Gla1, 63Gla2, 63Sud].

For a coherent state $|\alpha_0\rangle$ the *P*-distribution is obviously a Dirac delta function located at $\alpha = \alpha_0$, that is $P_{|\alpha_0\rangle} = \delta(\alpha - \alpha_0)$. In contrast, we find for a thermal state a Gaussian *P*-distribution, that is $P_{\rm th}(\alpha) = (\pi n_{\rm th})^{-1} \exp[-|\alpha|^2/n_{\rm th}]$ which is no longer singular. However, the *P*-distributions of a number state or a squeezed state are highly singular. They either involve a finite or an infinite number of derivatives of a delta function. Therefore, similar to the Mandel *Q*-parameter the Dirac delta function *P*-distribution of the coherent state defines a possible border between classical and quantum field states.

5.1.4 Atom–field interaction

So far, we have only considered the properties of the radiation field. In the present section we couple this quantized field to an atom. We therefore add more quantum degrees of freedom to our system. Indeed, we now have also to take into account the internal states of the atom as well as its center-of-mass motion. In this review we only motivate how to construct the appropriate interaction between these three quantum degrees and refer for a detailed derivation to the literature [01Sch].

5.1.4.1 Electric field-dipole interaction

An optical field does not change considerably over the size of an atom. Hence, the electric field E at the position $r_{\rm e}$ of the valence electron is almost identical to the electric field at the position $r_{\rm p}$ of the positively charged rest atom or to the one at the center-of-mass R. Therefore, we can make the dipole approximation in which the potential energy reads

$$H_{\boldsymbol{r}\cdot\boldsymbol{E}} \equiv -\boldsymbol{\wp}\cdot\boldsymbol{E}(\boldsymbol{R},t) = -e\boldsymbol{r}\cdot\boldsymbol{E}(\boldsymbol{R},t).$$
(5.1.41)

Here we have introduced the electric dipole moment $\boldsymbol{\wp} = e\boldsymbol{r} = e(\boldsymbol{r}_{\rm e} - \boldsymbol{r}_{\rm p})$ of an electron with charge e and the rest atom.

The total Hamiltonian of the atom in the electromagnetic field in dipole approximation then reads

$$\widetilde{H} \equiv H_{\rm f} + H_{\rm cm} + H_{\rm at} + H_{\boldsymbol{r}\cdot\boldsymbol{E}} \equiv H_{\rm f} + \frac{\boldsymbol{P}^2}{2M} + \frac{\boldsymbol{p}^2}{2\mu} + V(\boldsymbol{r}) - e\boldsymbol{r}\cdot\boldsymbol{E}(\boldsymbol{R},t), \qquad (5.1.42)$$

where $H_{\rm f}$ is the Hamiltonian of the radiation field (see (5.1.10)) and M and μ are the total and the reduced mass of the two parts of the atom. Moreover, we have introduced the momenta Pand p of the center-of-mass and the relative motion, respectively. The interaction between valence electron and rest atom is represented by the potential $V(\mathbf{r})$.

In the quantum version of this Hamiltonian we have to replace all observables by operators. We identify three quantum degrees of freedom: (i) the center-of-mass motion described by the conjugate variables $\hat{\boldsymbol{R}}$ and $\hat{\boldsymbol{P}}$ with commutation relation $[\hat{R}_j, \hat{P}_k] = i\hbar \delta_{jk}$; (ii) the relative motion described by the Hamiltonian $\hat{H}_{at} \equiv \hat{\boldsymbol{p}}^2/(2\mu) + V(\hat{\boldsymbol{r}})$ which contains the conjugate variables $\hat{\boldsymbol{r}}$ and $\hat{\boldsymbol{p}}$ with commutation relation $[\hat{r}_j, \hat{p}_k] = i\hbar \delta_{jk}$; (iii) the relative motion described by the Hamiltonian $\hat{H}_{at} \equiv \hat{\boldsymbol{p}}^2/(2\mu) + V(\hat{\boldsymbol{r}})$ which contains the conjugate variables $\hat{\boldsymbol{r}}$ and $\hat{\boldsymbol{p}}$ with commutation relation $[\hat{r}_j, \hat{p}_k] = i\hbar \delta_{jk}$; and (iii) the electric field operator $\hat{\boldsymbol{E}}$. The interaction Hamiltonian $\hat{H}_{\boldsymbol{r}\cdot\boldsymbol{E}}$ couples all three degrees of freedom.

We emphasize that for many applications this Hamiltonian suffices. However, this ad hoc procedure can also lead to inconsistent results as discussed in [94Wil].

5.1.4.2 Simple model for atom–field interaction

The most elementary model of the interaction of an atom with a quantized electromagnetic field consists of a two-level atom and a single mode of the cavity field. This model still contains enough physics to describe most phenomena in cavity QED and atom optics. When we neglect the center-of-mass motion, that is consider only the interaction between the quantized cavity field and the two levels we call this model the Jaynes–Cummings–Paul model [63Jay, 63Pau].

5.1.4.2.1 Hamiltonian

For the description of the internal dynamics of a two-level atom with excited state $|a\rangle$ and ground state $|b\rangle$ separated by an energy $\hbar\omega$ we recall the Pauli operators

$$\hat{\sigma} \equiv \frac{1}{2} \left(\hat{\sigma}_x - \mathrm{i}\hat{\sigma}_y \right) \equiv |b\rangle \langle a|, \quad \hat{\sigma}^{\dagger} \equiv \frac{1}{2} \left(\hat{\sigma}_x + \mathrm{i}\hat{\sigma}_y \right) \equiv |a\rangle \langle b|, \quad \hat{\sigma}_z \equiv |a\rangle \langle a| - |b\rangle \langle b|. \quad (5.1.43)$$

They allow us to represent the Hamiltonian $\hat{H}_{at} = \hbar \omega \hat{\sigma}_z/2$ of the atom and the transitions $\hat{\sigma} | a \rangle = |b\rangle$ and $\hat{\sigma}^{\dagger} | b \rangle = |a\rangle$ between the two levels in the energy eigenbasis. In this basis the atomic dipole operator $\hat{\boldsymbol{\wp}} = e\hat{\boldsymbol{r}}$ takes the form $\hat{\boldsymbol{\wp}} = \boldsymbol{\wp} \hat{\sigma}^{\dagger} + \boldsymbol{\wp}^* \hat{\sigma}$, where $\boldsymbol{\wp}$ is given by $\boldsymbol{\wp} \equiv e \langle a | \hat{\boldsymbol{r}} | b \rangle$. Please note, that the diagonal elements $\langle a | \hat{\boldsymbol{r}} | a \rangle$ and $\langle b | \hat{\boldsymbol{r}} | b \rangle$ vanish due to parity.

The electric field operator (5.1.16) of a single cavity mode characterized by a mode function $u(\hat{R})$, a mode frequency Ω and mode volume V reads

$$\hat{\boldsymbol{E}}(\hat{\boldsymbol{R}}) = \mathrm{i}\mathcal{E}_0 \boldsymbol{u}(\hat{\boldsymbol{R}}) \left(\hat{\boldsymbol{a}} - \hat{\boldsymbol{a}}^{\dagger}\right), \qquad (5.1.44)$$

where $\mathcal{E}_0 \equiv \sqrt{\hbar \Omega / (2\varepsilon_0 V)}$ denotes the vacuum electric field or the electric field per photon. With the help of (5.1.41) we find the interaction Hamiltonian

$$\hat{H}_{\boldsymbol{r}\cdot\boldsymbol{E}} = \hbar \left(\hat{g}\hat{\sigma}^{\dagger} - \hat{g}^{\dagger}\hat{\sigma} \right) \left(\hat{a} - \hat{a}^{\dagger} \right) = \hbar \left(\hat{g}\hat{\sigma}^{\dagger}\hat{a} - \hat{g}\hat{\sigma}^{\dagger}\hat{a}^{\dagger} - \hat{g}^{\dagger}\hat{\sigma}\hat{a} + \hat{g}^{\dagger}\hat{\sigma}\hat{a}^{\dagger} \right) , \qquad (5.1.45)$$

where the operator \hat{g} is given by $\hat{g}(\hat{R}) = -i\mathcal{E}_0(\boldsymbol{\wp} \cdot \boldsymbol{u}(\hat{R}))/\hbar$. This Hamiltonian contains operator products $\hat{\sigma}\hat{a}$ and $\hat{\sigma}^{\dagger}\hat{a}^{\dagger}$. They correspond to the annihilation of a photon together with a transition from the excited to the ground state and the creation of a photon together with the excitation of an atom. When we solve the Schrödinger equation perturbatively these operator products lead to higher-order contributions. Neglecting them amounts to the rotating-wave approximation.

In the rotating-wave approximation this simple model for the position-dependent interaction of a two-level atom with a single mode of the radiation field with frequency Ω is summarized by the Hamiltonian

$$\hat{H} = \frac{\hat{P}^2}{2M} + \hbar\Omega \,\hat{a}^{\dagger}\hat{a} + \frac{1}{2}\hbar\omega\hat{\sigma}_z + \hbar\left(\hat{g}\hat{\sigma}^{\dagger}\hat{a} + \hat{g}^{\dagger}\hat{\sigma}\hat{a}^{\dagger}\right) \,. \tag{5.1.46}$$

5.1.4.2.2 Dynamics of Jaynes–Cummings–Paul model

We can simplify the model even further when we keep the position of the atom fixed, that is we neglect the kinetic energy and replace the operator \hat{R} by a fixed vector R_0 . In this case, the Jaynes–Cummings–Paul model [63Jay, 63Pau] with the Hamiltonian

$$\hat{H}_{\rm JCP} \equiv \hbar \Omega \, \hat{a}^{\dagger} \hat{a} + \frac{1}{2} \hbar \omega \hat{\sigma}_z + \hbar g \left(\hat{\sigma} \hat{a}^{\dagger} + \hat{\sigma}^{\dagger} \hat{a} \right) \tag{5.1.47}$$

can be solved exactly. Here we have chosen the phases of the states $|a\rangle$ and $|b\rangle$ such that $\wp \cdot u(\mathbf{R}_0)$ is purely imaginary which allows us to introduce the vacuum Rabi frequency $g \equiv |\wp \cdot u(\mathbf{R}_0)| \mathcal{E}_0/\hbar$. For a detailed review of this model we recommend [90Sho, 93Sho].

Due to the particular atom-field coupling $\hat{\sigma}\hat{a}^{\dagger}$ and $\hat{\sigma}^{\dagger}\hat{a}$ the time-dependent solution of the Schrödinger equation is of the form

$$|\Psi(t)\rangle = \sum_{n=0}^{\infty} \left[\Psi_{a,n}(t) | a, n \rangle + \Psi_{b,n+1}(t) | b, n+1 \rangle \right] + \Psi_{b,0}(t) | b, 0 \rangle, \qquad (5.1.48)$$

where we have defined the states $|a, n\rangle \equiv |a\rangle |n\rangle$ and $|b, n\rangle \equiv |b\rangle |n\rangle$. In the interaction picture defined by $\hat{H}_0 = \hbar \Omega \, \hat{a}^{\dagger} \hat{a} + \frac{1}{2} \hbar \omega \hat{\sigma}_z$ the probability amplitudes

$$\Psi_{a,n}(t) = e^{-i\Delta t/2} \left\{ \left[\cos(\lambda_n t) + i\delta_n \sin(\lambda_n t) \right] \Psi_{a,n}(0) - i\varepsilon_n \sin(\lambda_n t) \Psi_{b,n+1}(0) \right\} ,$$

$$\Psi_{b,n+1}(t) = e^{i\Delta t/2} \left\{ -i\varepsilon_n \sin(\lambda_n t) \Psi_{a,n}(0) + \left[\cos(\lambda_n t) - i\delta_n \sin(\lambda_n t) \right] \Psi_{b,n+1}(0) \right\} ,$$

$$\Psi_{b,0}(t) = \Psi_{b,0}(0) \qquad (5.1.49)$$

contain the detuning $\Delta \equiv \Omega - \omega$ between the field and the atomic transition frequencies and the generalized Rabi frequencies $\lambda_n \equiv \sqrt{g^2(n+1) + (\Delta/2)^2}$. Moreover, we have introduced the abbreviations $\delta_n \equiv \Delta/(2\lambda_n)$ and $\varepsilon_n \equiv g\sqrt{n+1}/\lambda_n$.

5.1.4.2.3 Quantum motion in an ion trap

The Hamiltonian, (5.1.47), of the Jaynes–Cummings–Paul model can also be used to describe [92Blo, 94Cir] the center-of-mass motion of a two-level ion trapped in a harmonic potential and interacting with an appropriate classical field. In this application the operators \hat{a} and \hat{a}^{\dagger} are the annihilation and creation operators of the mechanical harmonic oscillator representing the center-of-mass motion of the ion in the trap. The coupling between the internal degrees of freedom and the motion is achieved by a classical electromagnetic field characterized by a coupling constant g. Even an anti-Jaynes–Cummings–Paul model [97Buz] and a nonlinear Jaynes–Cummings–Paul model [95Vog] can be formulated for this mechanical analog. Moreover, a whole gallery of non-classical states of motion has been realized experimentally [96Mee].

We conclude by noting that a Paul trap [90Pau] is a time-dependent harmonic oscillator. Nevertheless, the time-dependent nonlinear Jaynes–Cummings–Paul model allows for analytical solutions [96Bar, 97Sch].

5.1.4.3 Quantum state engineering

In the preceding section we have presented the exact solution of the Jaynes–Cummings–Paul model for arbitrary detuning. In the present section we briefly discuss the two extreme cases of $\Delta = 0$ and Δ very large. Both cases allow us to prepare special field states.

5.1.4.3.1 Resonant case: photon number state preparation

We start our discussion with the preparation of a Fock state $|N\rangle$ of N photons in a cavity. We assume that the resonator is initially in the vacuum state $|0\rangle$ and we inject N excited atoms one by one. Here we assume for the sake of simplicity that there is only one single atom in the cavity at a time. In this case, the interaction of each atom with the cavity field is governed by the Jaynes–Cummings–Paul model. For $\Delta = 0$ the time-dependent solution, (5.1.49), shows that for an appropriate interaction time an excited atom can deposit its excitation into the cavity field with unit probability. Since the Rabi frequency depends on the number of photons, the interaction time of every atom has to be different. Provided all N atoms transfer their excitation to the field we have, indeed, engineered the Fock state $|N\rangle$.

Such an experiment has been performed [00Var] using the microwave resonator of the one-atom maser in Garching. Starting from the vacuum state the experimentalists have created successively a one-photon and a two-photon state. They have probed these states with an additional atom and have measured the Rabi oscillations of this atom due to the so-prepared field as shown in Fig. 5.1.3.



Fig. 5.1.3. Fock state preparation and measurement. A sequence of two-level atoms passing an initially empty high-Q cavity (a) can prepare and probe a one-photon (b) or a two-photon Fock state (c) by giving the atomic excitation into the field and by measuring the resulting Rabi oscillations. Taken from [00Var].

5.1.4.3.2 Far off-resonant case: Schrödinger cat state preparation

Another interesting limit of the Jaynes–Cummings–Paul model arises for large detuning, that is for $|\Delta| \gg g\sqrt{n+1}$. Within this approximation the dynamics of the Jaynes–Cummings–Paul model can be described by the effective Hamiltonian [01Sch]

$$\hat{H}_{\text{eff}} \equiv -\frac{\hbar g^2}{\Delta} \left[\hat{\sigma}_z \, \hat{n} + \frac{1}{2} \left(\hat{\sigma}_z + \mathbb{1} \right) \right] \,. \tag{5.1.50}$$

Since the Hamiltonian only contains $\hat{\sigma}_z$ as well as the photon number operator \hat{n} it does not cause any transitions in the system. It conserves the populations in the atomic levels and the photon number. However, it introduces phase shifts in the field which depend on the state of the atom.

We make use of this fact [96Bru] to prepare the Schrödinger cat state, (5.1.31). For this purpose we add classical light fields at the entrance and the exit of the resonator. The first field prepares the atom in a coherent superposition of its internal states and the second field measures the soprepared dipole. In this Ramsey setup both fields have a fixed phase relation with each other. The quantized field in the resonator is initially in a coherent state $|\alpha\rangle$.

The photon number states accumulate phases whose sign depend on the internal states of the atom. The first field can be chosen such that the state of the combined system reads

$$|\Psi\rangle = \frac{1}{\sqrt{2}} \left[|\alpha e^{i\varphi}\rangle |a\rangle + |\alpha e^{-i\varphi}\rangle |b\rangle \right], \qquad (5.1.51)$$

where the phase φ is given by $\varphi = gt^2/\Delta$.

When we now measure the atom in an atomic superposition state $|\psi_{at}\rangle \equiv \psi_a |a\rangle + \psi_b |b\rangle$ the field state reads

$$|\psi_{\rm cat}\rangle \equiv \mathcal{N} \langle \psi_{\rm at} |\Psi\rangle = \frac{\mathcal{N}}{\sqrt{2}} \left[\psi_a^* |\alpha e^{i\varphi}\rangle + \psi_b^* |\alpha e^{-i\varphi}\rangle \right] , \qquad (5.1.52)$$

where \mathcal{N} is the normalization factor.

In this way the Schrödinger cat state, (5.1.31), of the radiation field has been prepared experimentally [96Bru]. A similar Schrödinger cat state for the vibratory motion of a harmonically trapped ion has also been generated [00Mya].

5.1.5 Reservoir theory

So far we have concentrated on pure states of the radiation field that interact in a unitary way with atoms. However, unitary time evolution cannot describe damping or amplification. For example, the state of an electromagnetic field in a leaky cavity cannot be a pure state. We have to resort to a density operator description [01Sch].

5.1.5.1 Master equation

Several models [96Scu] provide a quantum-mechanical description of a damped radiation field. The most intuitive approach is based on a cavity field interacting with an environment modeled by a stream of resonant two-level atoms. These atoms can carry energy away from the resonator. Since the disturbance caused by them is assumed to be small we need many atoms to create a significant change of the field state. It is therefore impossible to keep track of the internal states of every atom exiting the cavity. Moreover, we are only interested in the dynamics of the field. Consequently, we trace over the atomic states and arrive at an equation of motion for the density operator of the field. It is the trace operation that introduces irreversibility into this model.

5.1.5.1.1 Mathematics of the model

When we describe the interaction by the Jaynes–Cummings–Paul Hamiltonian the state of the complete system of atom and field is determined by (5.1.48). In this solution we take the trace over the atomic variables. As a result we find the density operator $\hat{\rho}_{\rm f}(t+\tau)$ of the field after the interaction time τ .

So far, we have only considered the dynamics of the field due to the interaction with a single atom. We now take into account the influence of a stream of atoms injected into the cavity with a rate r. Since each individual atom only introduces a small change we can approximate the dynamics of the field by the coarse-grained master equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\rho}_{\mathrm{f}}(t) \approx r \left[\hat{\rho}_{\mathrm{f}}(t+\tau) - \hat{\rho}_{\mathrm{f}}(t)\right].$$
(5.1.53)

Moreover, it is sufficient to retain terms quadratic in $g\tau$. The resulting equation of motion reads

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\hat{\rho}_{\mathrm{f}}(t) = \mathcal{L}(\hat{\rho}_{\mathrm{f}}) \equiv -\frac{1}{2}\,\mathcal{R}_{a}\left[\hat{a}\,\hat{a}^{\dagger}\,\hat{\rho}_{\mathrm{f}}(t) + \hat{\rho}_{\mathrm{f}}(t)\,\hat{a}\,\hat{a}^{\dagger} - 2\hat{a}^{\dagger}\,\hat{\rho}_{\mathrm{f}}(t)\,\hat{a}\right] \\
-\frac{1}{2}\,\mathcal{R}_{b}\left[\hat{a}^{\dagger}\,\hat{a}\,\hat{\rho}_{\mathrm{f}}(t) + \hat{\rho}_{\mathrm{f}}(t)\,\hat{a}^{\dagger}\,\hat{a} - 2\hat{a}\,\hat{\rho}_{\mathrm{f}}(t)\,\hat{a}^{\dagger}\right].$$
(5.1.54)

Here we have introduced the abbreviations $\mathcal{R}_a \equiv r \rho_{aa} (g\tau)^2$ and $\mathcal{R}_b \equiv r \rho_{bb} (g\tau)^2$, where ρ_{aa} and ρ_{bb} denote the populations in the two levels of the atoms. Furthermore, we have assumed a vanishing polarization $\rho_{ab} = 0$.

5.1.5.1.2 Methods of solution

Equations for the density operator of a subsystem such as (5.1.54) carry the name master equations. They are operator equations and usually involve a Liouville operator \mathcal{L} acting on the density operator. In the present case the Liouville operator contains only operators of the field mode. Moreover, the order of the individual operators is important. For this reason it is not straightforward to solve master equations. Several methods [85Gar, 89Ris, 91Gar] offer themselves:

 $a_{1,1} = l_{1,2} = s$ [89Coh, 91Mey, 92Coh, 94Wal, 96Scu, 01Sch]: An expansion of the density operator in terms of Fock states leads to equations of motions for the matrix elements forming a coupled system of ordinary linear differential equations.

a ..., [93Car]: Here a stochastic sequence of jump processes and non-unitary time evolutions based on a Schrödinger equation with a non-hermitean Hamiltonian is solved numerically. By averaging over all trajectories we obtain a numerical solution of the master equation.

5.1.5.2 Damping and amplification

For the case of the master equation, (5.1.54), these techniques show the following results. When more atoms enter the cavity in the $\dots s a$ rather than in the ground state, that is $\rho_{aa} > \rho_{bb}$ we find an $a \downarrow l \downarrow a \dots$ of the field. On average more atoms deposit their excitation than atoms take photons out of the cavity.

However, when more atoms enter the cavity in thes a rather than in the excited state, that is $\rho_{bb} > \rho_{aa}$ we find a a_{a} of the field. On average more atoms withdraw excitations than atoms deposit photons into the cavity. Independent of the initial state the cavity field eventually approaches a thermal state, (5.1.38), with a temperature T determined by the initial atomic populations. The average number of thermal photons $n_{\rm th} = [\exp[\hbar\omega/(k_{\rm B}T)] - 1]^{-1} = [\rho_{bb}/\rho_{aa} - 1]^{-1}$ in the limit of $t \to \infty$ and the decay rate $\kappa \equiv \mathcal{R}_b - \mathcal{R}_a = r(g\tau)^2(\rho_{bb} - \rho_{aa})$ allow us to cast the Liouville operator $\mathcal{L}_{\rm damp}$ of damping into the form

$$\mathcal{L}_{\rm damp}(\hat{\rho}_{\rm f}) \equiv -\frac{\kappa}{2} \left(n_{\rm th} + 1 \right) \left(\hat{a}^{\dagger} \hat{a} \hat{\rho}_{\rm f} + \hat{\rho}_{\rm f} \hat{a}^{\dagger} \hat{a} - 2\hat{a} \hat{\rho}_{\rm f} \hat{a}^{\dagger} \right) - \frac{\kappa}{2} n_{\rm th} \left(\hat{a} \hat{a}^{\dagger} \hat{\rho}_{\rm f} + \hat{\rho}_{\rm f} \hat{a} \hat{a}^{\dagger} - 2\hat{a}^{\dagger} \hat{\rho}_{\rm f} \hat{a} \right) .$$
(5.1.55)

We emphasize that \mathcal{L}_{damp} only contains field quantities such as the cavity decay constant κ and the average number n_{th} of photons together with field operators.

5.1.5.3 Decoherence

The superposition principle is one of the corner stones of quantum mechanics. Indeed, as P.A.M. Dirac states in his famous textbook [35Dir] "... any two or more states may be superposed to give a new state". The example of the Schrödinger cat state, (5.1.31), demonstrates the power of this principle. Both coherent states forming the cat are classical states. However, their superposition is a highly non-classical state [91Sch].

The non-classical features arising from the superposition principle stand out most clearly in $|\psi_{\text{cat}}\rangle$ for two coherent states of large amplitude α and a large angle 2φ between them, see Fig. 5.1.1. For this set of parameters we can distinguish the two individual states but still observe the interference between them. Since our classical world does not contain such superpositions of distinguishable states the question arises: What is the mechanism that erases the interference terms [91Zur, 96Giu] when we cross the border from the quantum to the classical world?

Niels Bohr has always insisted that a measurement apparatus has to be described in classical terms involving "irreversible amplification" [58Boh]. Irreversibility implies many degrees of freedom. It is the feature of the many degrees of freedom that erases the quantum interference.

We now illustrate this concept using the example of the measurement of a cavity field. For this purpose we extract a portion of the field from the resonator, that is we couple it to the modes of the measurement apparatus. Therefore, we arrive at the problem of a single mode coupled to



Fig. 5.1.4. Decoherence of a Schrödinger cat state coupled to a thermal reservoir, here demonstrated for the center-of-mass motion of an ion in a trap. The contrast of the interference fringes decreases as a function of time t. The decay depends on the square of the separation $|\Delta \alpha|$ of the individual constituents. The solid line is a fit to an exponential decay law. Taken from [00Mya].

many more field modes. The corresponding Liouville operator for this damping model is identical to (5.1.55). The number $n_{\rm th}$ of thermal photons in this case is determined by the state of the reservoir and κ is the decay constant of the cavity. Consequently, the dynamics of the cavity field in the interaction picture is determined by $\dot{\rho}_{\rm f} = \mathcal{L}_{\rm damp}(\hat{\rho}_{\rm f})$ with the Liouville operator $\mathcal{L}_{\rm damp}$, (5.1.55).

Deeper insight into the tantalizing question of the emergence of the classical world follows from this master equation when we consider as an initial state the Schrödinger cat

$$\hat{\rho}_{\text{cat}} \equiv |\psi_{\text{cat}}\rangle\langle\psi_{\text{cat}}| = |\mathcal{N}|^{2} \left[|\alpha e^{i\varphi}\rangle\langle\alpha e^{i\varphi}| + |\alpha e^{-i\varphi}\rangle\langle\alpha e^{-i\varphi}| + |\alpha e^{i\varphi}\rangle\langle\alpha e^{-i\varphi}| + |\alpha e^{-i\varphi}\rangle\langle\alpha e^{i\varphi}| \right]$$
(5.1.56)

shown in Fig. 5.1.1.

Indeed, a treatment using any of the three methods outlined in Sect. 5.1.5.1.2 shows that the diagonal elements $|\alpha e^{i\varphi}\rangle\langle\alpha e^{i\varphi}|$ and $|\alpha e^{-i\varphi}\rangle\langle\alpha e^{-i\varphi}|$ decay with the decay constant κ of the cavity. However, the off-diagonal terms $|\alpha e^{i\varphi}\rangle\langle\alpha e^{-i\varphi}|$ and $|\alpha e^{-i\varphi}\rangle\langle\alpha e^{i\varphi}|$, that is the interference terms, decay much faster. The decay constant of these contributions is determined by the product $\kappa (2\alpha \sin \varphi)^2$ of the decay rate of the cavity and the square of the separation $\Delta \alpha = 2\alpha \sin \varphi$ of the coherent states in phase space shown in Fig. 5.1.1. Hence, when they are far apart and we can easily distinguish them, the interferences decay on a much faster time scale and we are left with a mixed state $\hat{\rho} = (|\alpha e^{i\varphi}\rangle\langle\alpha e^{i\varphi}| + |\alpha e^{-i\varphi}\rangle\langle\alpha e^{-i\varphi}|)/2$ which only contains diagonal elements.

This mechanism of decoherence arises when we couple a system to a reservoir. It eliminates quantum-mechanical interference between macroscopically distinguishable states. The enhanced decay of interference structures has been observed using Schrödinger cat states in cavity QED [96Bru] and ion experiments [00Mya]. In particular, the quadratic dependence on the separation of the two contributing states has been verified as shown in Fig. 5.1.4.

5.1.6 One-atom maser

The combination of superconducting cavities with Rydberg atoms has created a new light source with highly unusual quantum-statistical properties. In this amazing maser [94Rai] a weak beam of excited two-level atoms traverses a cavity with a quality factor of 10^{10} and interacts resonantly

20

with a single mode of the radiation field. The flux of the atomic beam is so small that at most one atom at a time is in the cavity. The atom can deposit its excitation in the cavity and in this way amplify the field. The atom also serves a different purpose: It probes the field. A detector after the resonator measures the population of the internal levels.

5.1.6.1 Master equation

We describe this interaction by the resonant Jaynes–Cummings–Paul (JCP) Hamiltonian in the interaction picture. By taking the trace over the atomic degrees of freedom we find [01Sch] the coarse-grained equation of motion

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\hat{\rho}_{\mathrm{f}}(t) \equiv \mathcal{L}_{\mathrm{JCP}}(\hat{\rho}_{\mathrm{f}}) \equiv r[\hat{C}_{n}(\tau)\,\hat{\rho}_{\mathrm{f}}(t)\,\hat{C}_{n}(\tau) - \hat{\rho}_{\mathrm{f}}(t)] + r\frac{\hat{S}_{n-1}(\tau)}{\sqrt{\hat{n}}}\,\hat{a}^{\dagger}\hat{\rho}_{\mathrm{f}}(t)\,\hat{a}\,\frac{\hat{S}_{n-1}(\tau)}{\sqrt{\hat{n}}}\,,\qquad(5.1.57)$$

where we have introduced the abbreviations $\hat{C}_n(\tau) \equiv \cos(g\tau\sqrt{\hat{n}+1})$, $\hat{S}_n(\tau) \equiv \sin(g\tau\sqrt{\hat{n}+1})$ and the atomic injection rate r. Moreover, we have assumed that the atoms enter the cavity in the excited state, that is $\rho_{aa} = 1$ and $\rho_{bb} = \rho_{ab} = \rho_{ba} = 0$.

We assume, that the time interval between two successive atoms is much larger than the interaction time of a single atom. In this time interval we take into account the damping of the field mode by the Liouville operator $\mathcal{L}_{damp}(\hat{\rho}_{f})$, defined in (5.1.55). Hence the complete dynamics

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\hat{\rho}_{\mathrm{f}}(t) \equiv \mathcal{L}_{\mathrm{JCP}}(\hat{\rho}_{\mathrm{f}}) + \mathcal{L}_{\mathrm{damp}}(\hat{\rho}_{\mathrm{f}}) \tag{5.1.58}$$

of the one-atom maser [86Fil, 87Lug] is governed by the sum of the two Liouville operators \mathcal{L}_{JCP} and \mathcal{L}_{damp} . We emphasize that this equation of motion is an approximation which makes assumptions on the statistics of the injected atoms. Modifications of (5.1.58) due to different pump statistics are discussed in [94Ber].

5.1.6.2 Steady-state photon statistics

We first analyze the properties of the steady-state photon statistics [86Fil, 87Lug]

$$W_{n} \equiv \langle n | \hat{\rho}_{\rm f} | n \rangle = W_{0} \prod_{l=1}^{n} \left[\frac{n_{\rm th}}{n_{\rm th} + 1} + \frac{r \sin^{2}(g\tau \sqrt{l}\,)}{\kappa(n_{\rm th} + 1)l} \right]$$
(5.1.59)

determined by the diagonal elements of the density operator of the one-atom maser. Here W_0 is a normalization constant.

In Fig. 5.1.5 we show the mean photon number $\langle \hat{n} \rangle$ and the normalized variance σ defined in (5.1.39) in steady state using the stationary photon statistics, (5.1.59), as a function of the dimensionless interaction time $\theta \equiv \sqrt{r/\kappa} g\tau$. We recognize a steep increase of $\langle \hat{n} \rangle$ as θ approaches unity which is the maser threshold. A similar effect occurs in every laser. However, due to the presence of the sine function in (5.1.59) more thresholds appear, but they are not as pronounced as the first one.

Moreover, the normalized variance σ shown in Fig. 5.1.5 by a solid curve shows that the maser oscillates between sub-Poissonian and super-Poissonian statistics. These different domains have been observed [90Rem] in experiments.

The photon statistics of the one-atom maser becomes particularly interesting when there are no thermal photons present, that is for $n_{\rm th} = 0$ or T = 0. When we now choose the interaction time τ such that



Fig. 5.1.5. Normalized mean photon number $\langle \hat{n} \rangle$ and normalized variance σ of the photon statistics of the one-atom maser in steady state as a function of the dimensionless interaction time $\theta \equiv \sqrt{r/\kappa} g\tau$. The pump parameter is $r/\kappa = 200$ and the dotted line indicates the value $\sigma = 1$ corresponding to a coherent state with Poissonian statistics.

$$g\tau\sqrt{n_q+1} = q\pi\,,\tag{5.1.60}$$

where q and n_q are positive integers, the steady-state photon statistics, (5.1.59), vanishes starting from the photon number $n_q + 1$. This phenomenon carries the name "trapping state" [88Mey].

An intuitive explanation of trapping states recognizes that an excited atom interacting with the state $|n_q\rangle$ cannot deposit its excitation in the cavity because for this particular interaction time the atom undergoes an integer number of Rabi cycles and therefore leaves the cavity in the excited state. Trapping states have been observed experimentally [99Wei] in the one-atom maser.

5.1.7 Atom–reservoir interaction

Our previous studies have focused on the dynamics of a single mode driven by a reservoir of twolevel atoms. In the present section we consider the dynamics of a single two-level atom under the influence of infinitely many field modes forming the reservoir. This coupling of the atom to the field reservoir leads to spontaneous emission of the atoms as well as level shifts. We briefly discuss the equation of motion for the density operator $\hat{\rho}_{at}$ of the atom and analyze the consequences.

5.1.7.1 Master equation

In order to describe the atom–reservoir interaction we consider the Hamiltonian

$$\hat{H}_{\rm ar} \equiv \hat{H}_{\rm at} + \hat{H}_{\rm f} + \hat{H}_{\rm int} \equiv \frac{1}{2} \hbar \omega \hat{\sigma}_z + \sum_{\ell} \hbar \Omega_\ell \hat{a}_\ell^\dagger \hat{a}_\ell + \sum_{\ell} \hbar g_\ell (\hat{\sigma} \hat{a}_\ell^\dagger + \hat{\sigma}^\dagger \hat{a}_\ell) , \qquad (5.1.61)$$

where Ω_{ℓ} is the frequency of the ℓ th mode and $g_{\ell} \equiv g_{\ell}(\mathbf{R})$ denotes the coupling strength of the atom at the position \mathbf{R} to the ℓ th mode.

The resulting master equation in the interaction picture for the atomic density operator $\hat{\rho}_{at}$ reads [73Lou, 01Sch]

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\rho}_{\mathrm{at}} = -\frac{1}{\hbar} \left[\Delta \hat{H}, \hat{\rho}_{\mathrm{at}} \right]
- (\Gamma_{\mathrm{r}} + G_{\mathrm{r}}) \left[\hat{\sigma}^{\dagger} \hat{\sigma} \hat{\rho}_{\mathrm{at}} + \hat{\rho}_{\mathrm{at}} \hat{\sigma}^{\dagger} \hat{\sigma} - 2\hat{\sigma} \hat{\rho}_{\mathrm{at}} \hat{\sigma}^{\dagger} \right] - \Gamma_{\mathrm{r}} \left[\hat{\sigma} \hat{\sigma}^{\dagger} \hat{\rho}_{\mathrm{at}} + \hat{\rho}_{\mathrm{at}} \hat{\sigma} \hat{\sigma}^{\dagger} - 2\hat{\sigma}^{\dagger} \hat{\rho}_{\mathrm{at}} \hat{\sigma} \right]
+ 2\beta^{*} \hat{\sigma} \hat{\rho}_{\mathrm{at}} \hat{\sigma} + 2\beta \hat{\sigma}^{\dagger} \hat{\rho}_{\mathrm{at}} \hat{\sigma}^{\dagger},$$
(5.1.62)

where we have introduced the Hamiltonian

Ref. p. 39]

$$\Delta \hat{H} \equiv -\hbar \left(\Gamma_{\rm i} + \frac{1}{2} G_{\rm i} \right) \hat{\sigma}_z \,. \tag{5.1.63}$$

The explicit expressions [01Sch] for the complex-valued quantities $\Gamma \equiv \Gamma_{\rm r} + i \Gamma_{\rm i}$, $G \equiv G_{\rm r} + i G_{\rm i}$ and $\beta \equiv \beta_{\rm r} + i \beta_{\rm i}$ are rather complicated and can be even infinite [73Lou]. They are determined by the state of the reservoir through the expectation values of the reservoir field modes. For example, Γ depends on the average number $\langle \hat{n}_{\ell} \rangle$ of photons in the individual reservoir modes as well as the expectation values $\langle \hat{a}_{\ell}^{\dagger} \rangle$ and $\langle \hat{a}_{\ell} \rangle$. Since for a thermal reservoir we have $\langle \hat{a}_{\ell} \rangle = \langle \hat{a}_{\ell}^{\dagger} \rangle = 0$ it is only $\langle \hat{n}_{\ell} \rangle$ that determines Γ . The parameter β depends on $\langle \hat{a}_{\ell}^2 \rangle$ and $\langle \hat{a}_{\ell} \rangle$. For a thermal state both quantities vanish. However, for a squeezed vacuum we find $\langle \hat{a}_{\ell}^2 \rangle \neq 0$ which leads to a non-vanishing value of β . This fact has important consequences for the decay of the atomic excitation into a squeezed vacuum as discussed in [86Gar].

However, the most interesting coefficient is G since it does not contain any expectation values of the reservoir. Its origin can be traced back [01Sch] to the commutation relation between \hat{a}_{ℓ} and \hat{a}_{ℓ}^{\dagger} . It is therefore a consequence of the quantization of the radiation field.

5.1.7.2 Lamb shift

We now turn to the correction $\Delta \hat{H}$ to the Hamiltonian $\hat{H}_{at} = \frac{1}{2}\hbar\omega\hat{\sigma}_z$ of the free atom. The interaction of the atom with a reservoir of field modes leads to a level shift of frequency

$$\Delta\omega \equiv -2\left(\Gamma_{\rm i} + \frac{1}{2}\,G_{\rm i}\right)\tag{5.1.64}$$

which carries the name Lamb shift. It has been discovered experimentally in 1947 by Willis E. Lamb and his graduate student Robert C. Retherford in hydrogen [47Lam]. The Dirac equation predicts that the $2S_{1/2}$ and $2P_{1/2}$ energy levels are degenerated. However, the experiment showed clearly that this degeneracy is lifted. The Lamb shift is a manifestation of the quantization of the electromagnetic field [01Sch]. Indeed, we recognize from (5.1.64) that there are two shifts: The first arises from the imaginary part Γ_i of Γ . Since the parameter Γ is essentially determined by the average number of photons in the reservoir modes this contribution to the level shift is analogous to the familiar second-order Stark-effect of an atom in a static electric field.

In contrast, the contribution G_i arises from the commutation relations of the field operators. It is therefore a pure quantum effect of the field and is nonzero even when all modes of the reservoir are in the ground state, that is in the vacuum.

5.1.7.3 Weisskopf–Wigner decay

Apart from the shift of the levels the reservoir has another dramatic effect: It forces the atom to decay, that is the populations $\rho_{aa} \equiv \langle a | \hat{\rho}_{at} | a \rangle$ and $\rho_{bb} \equiv \langle b | \hat{\rho}_{at} | b \rangle$ in the two levels and the polarization $\rho_{ab} \equiv \langle a | \hat{\rho}_{at} | b \rangle$ change as a function of time. This phenomenon carries the name Weisskopf–Wigner decay.

We now consider all reservoir modes being in the vacuum state. In this case the coefficients Γ and β vanish. When we introduce the decay constant $\gamma \equiv 2G_r$ of the atom, (5.1.62) reduces to the master equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\rho}_{\mathrm{at}} = \mathcal{L}_{\mathrm{sp}}(\hat{\rho}_{\mathrm{at}}) \equiv -\frac{1}{2}\gamma \left[\hat{\sigma}^{\dagger}\hat{\sigma}\hat{\rho}_{\mathrm{at}} + \hat{\rho}_{\mathrm{at}}\hat{\sigma}^{\dagger}\hat{\sigma} - 2\hat{\sigma}\hat{\rho}_{\mathrm{at}}\hat{\sigma}^{\dagger}\right]$$
(5.1.65)

of spontaneous emission, where we have omitted the Lamb shift term $\left|\Delta \hat{H}, \hat{\rho}_{\rm at}\right|$.

When we now take matrix elements of the atomic density operator in the energy basis, that is project from left and right with energy states $|a\rangle$ and $|b\rangle$ on (5.1.65), we arrive at the equations of motion

$$\dot{\rho}_{aa} = -\gamma \rho_{aa}, \quad \dot{\rho}_{bb} = +\gamma \rho_{aa}, \quad \dot{\rho}_{ab} = -\frac{\gamma}{2} \rho_{ab}.$$
(5.1.66)

Hence the population of the excited state decays, but due to the conservation $\rho_{aa} + \rho_{bb} = 1$ of probability, the population in the ground state grows. Moreover, the polarization ρ_{ab} also decays and the atom finally ends up in its ground state.

5.1.8 Resonance fluorescence

We now consider a two-level atom driven by a classical monochromatic wave. The excited state of the atom can decay by spontaneous emission into vacuum modes of the electromagnetic field. This emission is called resonance fluorescence. Of particular interest are the quantum-statistical properties of the emitted light. For a detailed discussion of resonance fluorescence, see for example [91Mey, 92Coh, 93Car, 94Wal].

5.1.8.1 Model

The electromagnetic field radiated by an atom located at the origin is in the far field proportional to its dipole moment and can therefore be expressed in terms of the Pauli operators $\hat{\sigma}$ and $\hat{\sigma}^{\dagger}$ [93Car]. Knowledge of $\hat{\sigma}$ and $\hat{\sigma}^{\dagger}$ is therefore sufficient to study the properties of the emitted light in the far field. We therefore consider a two-level atom driven by a classical field and coupled to the reservoir discussed in Sect. 5.1.7.

The total Hamiltonian for this model system describing resonance fluorescence then reads

$$\hat{H}_{\rm rf} \equiv \hat{H}_{\rm ar} + \frac{1}{2}\hbar\Omega_{\rm R} \left(\hat{\sigma} e^{i\omega t} + \hat{\sigma}^{\dagger} e^{-i\omega t}\right) \,, \tag{5.1.67}$$

where we have added a resonant driving term to the Hamiltonian \hat{H}_{ar} , (5.1.61), of the atomreservoir interaction and Ω_{R} is the Rabi frequency associated with the driving field.

The corresponding master equation in the interaction picture takes the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\hat{\rho}_{\mathrm{at}} = -\frac{\mathrm{i}}{2}\,\Omega_{\mathrm{R}}\,\left[\hat{\sigma} + \hat{\sigma}^{\dagger}, \hat{\rho}_{\mathrm{at}}\right] + \mathcal{L}_{\mathrm{sp}}(\hat{\rho}_{\mathrm{at}})\,,\tag{5.1.68}$$

where the modes of the thermal reservoir are in the vacuum as described by (5.1.65).

5.1.8.2 Spectrum and antibunching

One of the quantities of interest is the spectrum of the emitted radiation. According to the Wiener– Khintchine theorem [95Man] this spectrum is determined by the autocorrelation function of the electric field. As mentioned the electric field follows from the dipole. As a consequence we need to


Fig. 5.1.6. Experimental three-peak Mollow spectrum for increasing laser intensity. We note the emergence of the side peaks. The elastic peak ideally represented by a delta function located on top of the central peak is not shown. Taken from [76Har].

calculate the autocorrelation function of the dipole [91Mey, 93Car]. For this purpose we take matrix elements of the master equation, (5.1.68), in the energy representation. The resulting equations are the Bloch equations.

The autocorrelation function is the expectation value of the product of the operators $\hat{\sigma}^{\dagger}(t)$ and $\hat{\sigma}(0)$ at two different times t and t = 0. In order to calculate such expectation values we make use of the quantum regression theorem which allows us to reduce two-time correlation functions to one-time expectation values. For more details we refer to [96Scu].

The result of this calculation shows that the steady-state spectrum of the fluorescence light consists of two contributions: a coherent part $S_{\rm coh}(\nu)$ and an incoherent part $S_{\rm inc}(\nu)$. The coherent part

$$S_{\rm coh}(\nu) \propto \frac{\Omega_{\rm R}^2 \gamma^2}{\left(\gamma^2 + 2\Omega_{\rm R}^2\right)^2} \,\delta(\nu - \omega) \tag{5.1.69}$$

contains a Dirac delta function. This is in analogy to a driven classical dipole that radiates at the frequency of the driving field.

The incoherent part of the fluorescence light displays two qualitatively different spectra depending on the value of the Rabi frequency. For $\Omega_{\rm R} < \gamma/4$ it has a single peak at the atomic transition frequency ω , whereas for $\Omega_{\rm R} > \gamma/4$ it consists of three peaks – the so-called Mollow-triplet, see Fig. 5.1.6. For $\Omega_{\rm R} \gg \gamma/4$ we can represent the spectrum

$$S_{\rm inc}(\nu) \propto \frac{1}{2\pi\gamma} \left[\mathcal{D}(\nu-\omega,\gamma/2) + \frac{1}{3} \mathcal{D}(\nu-\omega+\Omega_{\rm R},3\gamma/4) + \frac{1}{3} \mathcal{D}(\nu-\omega-\Omega_{\rm R},3\gamma/4) \right]$$
(5.1.70)

by the sum of three Lorentzians $\mathcal{D}(\xi, \Gamma) \equiv \Gamma^2 / (\xi^2 + \Gamma^2)$ centered at the frequencies $\nu = \omega$ and $\nu = \omega \pm \Omega_{\rm R}$ with different width and heights. The central peak at $\nu = \omega$ has a width of $\gamma/2$, whereas the width of the two side peaks at $\nu = \omega \pm \Omega_{\rm R}$ is $3\gamma/4$. Their heights are one third of the height of the central peak. This spectrum was predicted by Anatoly I. Burshtein [65Bur] and Benjamin R. Mollow [69Mol] and experimentally confirmed in a series of experiments [74Sch, 75Wu, 76Har], as shown in Fig. 5.1.6.

Resonance fluorescence also displays another interesting effect: The photons of the emitted radiation tend to be separated. This tendency is known as photon antibunching [76Car1, 76Car2] and was experimentally verified [77Kim, 78Kim]. This phenomenon has a simple explanation: After the atom has emitted a photon it is in the ground state and must first be excited again before it can emit another photon. The use of a single ion stored in a Paul trap [98Win] or an atom in a magneto-optical trap [93Hau] have led to extremely clear verifications of photon antibunching.

We conclude by mentioning that also squeezing in the resonance fluorescence has been detected [98Lu]. Moreover, antibunching of free electrons has been observed recently [02Kie, 02Spe].

5.1.9 Fundamental questions of quantum mechanics

The interpretation of quantum mechanics has been and still is the subject of a longstanding debate. For many years this topic has been confined to purely theoretical discussions. However, the enormous progress in the experimental tools of quantum optics has now promoted such gedanken experiments to real world experiments. In the present section we highlight the experiments on quantum jumps, illustrate quantum-optical tests of complementarity and present violations of Bell's inequalities. Space does not allow us to do justice to the field of entangled photon pairs and beam splitting experiments with correlated photons. For this topic we refer to Sect. 6.1.4.5 of this volume and to the overviews [88Bre, 91Bre].

5.1.9.1 Quantum jumps

Our understanding of the internal dynamics of atoms has undergone dramatic changes: from the static raisin model of Joseph J. Thomson [1904Tho] via the Bohr–Sommerfeld planetary concept to the atom in QED [49Kro]. However, even the Heisenberg–Schrödinger quantum atom is still full of surprises.

5.1.9.1.1 Continuous versus discontinuous dynamics

Electrons in atoms do not move on circles or ellipses around the nucleus, but declare themselves in transitions from one Bohr orbit to another, that is in quantum jumps. These quantum jumps form the building blocks of Werner Heisenberg's matrix mechanics [25Hei, 30Bor]. Erwin Schrödinger's wave mechanics takes the completely opposite point of view. It describes a quantum system in terms of a wave function which undergoes a continuous change in time governed by the familiar Schrödinger equation.

These on first sight contradictory pictures occupied the founders of quantum mechanics. On a visit to Copenhagen in 1926 Niels Bohr kept discussing the issue of quantum jumps with Schrödinger to such an extent that eventually Schrödinger fell ill and had to stay in bed. Even then Bohr kept pestering him. Finally Schrödinger exclaimed "If we are going to have to put up with these damn quantum jumps, I am sorry that I ever had anything to do with quantum theory". Bohr's answer: "But the rest of us are very thankful for it ..." [94Moo].

The ultimate resolution of this puzzle was provided by Schrödinger [26Sch1] showing that both approaches are different sides of the same coin. It is interesting to note that also Wolfgang Pauli in a letter to Pascual Jordan had arrived at the same conclusions but did not publish it since Schrödinger's paper had appeared in the meantime [79Pau]. In 1926 Jordan used the matrix mechanics to provide a quantum-mechanical theory of quantum jumps [27Jor].

5.1.9.1.2 Experimental observation

In 1952 Schrödinger returned to the question of quantum jumps in two papers [52Sch1, 52Sch2] entitled "Are there quantum jumps?". Schrödinger was wondering if quantum jumps could ever be observed in an experiment. His answer "No!" was based on the fact that at that time one could not imagine to experiment with single atoms. To quote from his paper: "... we are not with single particles, any more than we can raise Ichthyosauria in the zoo" [52Sch2].

Due to the invention of ion traps this impossibility argument no longer holds true. Indeed, Hans G. Dehmelt had suggested [75Deh] to demonstrate quantum jumps using a single ion stored in a Paul trap. For this purpose he investigated a three-level atom in a V-configuration with two excited levels and a common ground state. The transitions between ground and excited levels are driven by two resonant laser fields. The life times of the two upper levels are utterly different and we observe the fluorescence on the transition from the level with the fast decay.

The experimental result displayed in Fig. 5.1.7 shows the light emitted from a single ion driven in such a way. At random instances of time we find dark periods. During this time the atom has made a transition to the excited state with the slow decay.

We conclude by mentioning that many experiments [86Ber, 86Nag, 86Sau] have observed quantum jumps. Moreover, this phenomenon has important applications in the context of frequency measurements [00Zan] and the development of a quantum computer based on ions.



Fig. 5.1.7. Experimental observation of quantum jumps in In^+ . The fluorescence light from the laser-driven transition $1S_0-3P_1$ is observed. From time to time the atom jumps spontaneously from the $3P_1$ level to the meta-stable $3P_0$ level. During these periods no fluorescence is observed as can be seen from the dips in the fluorescence light in the insets. Picture kindly provided by Joachim von Zanthier, Max-Planck-Institut für Quantenoptik, Garching.

5.1.9.2 Wave–particle duality

Two hundred years ago Thomas Young demonstrated the wave nature of light by his now famous double-slit experiment [1802You, 64Bor]. At the turn of the last century the question emerged if this wave nature also manifests itself in the limit of low light intensities, in particular in the limit of single photons. For this reason many experiments with attenuated light sources have been performed. However, only recently clean quantum-optics experiments could answer this question with a definitive "Yes!". For example, the experiment of [86Gra, 89Asp] feeds the antibunched light emitted by an atom into a Mach–Zehnder interferometer, shown in Fig. 5.1.8. The phenomenon of antibunching guarantees that there is only one photon at a time in the apparatus. By repeating the experiment many times the familiar interference pattern emerges.



Fig. 5.1.8. Mach–Zehnder interferometer consisting of two beam splitters (BS) and two mirrors (M). We observe either interference (beam splitter BS 2 inserted) or obtain path information (beam splitter BS 2 removed). In the delayed choice mode of operation we insert BS 2 only after the photon has passed BS 1.

Young's double-slit experiment has also played a central role in the Bohr–Einstein dialog [83Whe] on wave–particle duality. This discussion centers around the question: Is it possible to construct an experimental setup that provides information about the path of the photon and at the same time displays interference? The Mach–Zehnder interferometer of Fig. 5.1.8 illustrates the point of question. Imagine a situation in which the second beam splitter is removed. A single photon entering the interferometer at the first beam splitter triggers only one of the two detectors. In this way we obtain information about the path of the photon. Detectors which provide such information carry the name "which-path" detectors, or in the German version "welcher Weg" detectors.

The fact that we have to either insert or take out the second beam splitter already indicates that we cannot obtain simultaneously in the very same experiment complete which-path *a* interference information. Niels Bohr felt that this exclusiveness of two observables is a central feature of quantum mechanics. He called [28Boh] mutual exclusive properties such as which-path and interference "complementary variables". His principle of complementarity is closely related to Heisenberg's uncertainty principle. However, the question which principle is the more fundamental is not resolved yet [96Eng].

An interesting manifestation of complementarity and wave-particle duality appears in resonance fluorescence [97Hoe]. The coherently scattered radiation displays a fixed phase relation with respect to the driving field reflecting the wave nature of light. On the other hand the same radiation exhibits photon antibunching which is a consequence of the quantum nature of light and matter. However, two different experimental setups are necessary to observe these complementary features. The first relies on a homodyne setup whereas the second one needs a correlation measurement.

5.1.9.2.1 Delayed choice experiments

A rather paradoxical situation occurs when we start from a Mach–Zehnder interferometer where initially the second beam splitter is removed. We insert it only after the photon has passed through the first beam splitter and is well on its way towards the detectors. Does this delayed choice experiment [31Wei, 41Wei, 79Whe] display interference? Loosely speaking, the photon has started its path as a particle. Indeed, when it passed the first beam splitter it had to make a choice to travel along only one of the two arms. However, by the insertion of the beam splitter it is forced to suddenly behave as a wave. For this purpose it would have had to travel on both paths. Classical intuition tells us that the past has already decided the future of the photon: When the photon entered the interferometer the second beam splitter was absent and the photon had to take one of the two paths. Therefore, we expect no interference to occur.

5.1.9.2.2 Quantum-optical tests of complementarity

Nature prevents us from obtaining complete which-path information when we try to observe also interference. But what is the mechanism that erases the information about one of the two complementary observables? Niels Bohr in his rebuttal to Albert Einstein's ingenious proposal of recoiling slits [49Boh] argues that the physical positions of the slits are only known within the uncertainty principle. This error contributes a random phase shift to the photon which erases the interference patterns. Such random phase arguments are appealing. Unfortunately, they are incomplete: In principle, and in practice, it is possible to design experiments [91Scu, 98Due] that provide whichpath information via detectors without disturbing the system in any noticeable way. In this case, the loss of coherence is due to the establishing of quantum correlations.

Figure 5.1.9 displays such a which-path detector. Here we consider a double-slit experiment for atoms with a high-Q microwave resonator behind each slit. The fields are resonant with an atomic transition. The interaction time is chosen in such a way that the atom makes a transition while it traverses the resonator. In this case the field state is modified by the deposition of a single photon. When there is only one atom in the apparatus at the same time it can only deposit one single photon. This event can take place either in the upper or the lower cavity. We denote the center-of-mass wave function for the path through the upper (lower) cavity by $\phi_u(\mathbf{r})$ ($\phi_l(\mathbf{r})$) and the initial field state in both cavities by $|\psi^{(i)}\rangle$. Hence, the initial state before the interaction reads

$$\Psi^{(i)}(\boldsymbol{r}) \rangle = \left[\phi_{u}(\boldsymbol{r}) + \phi_{l}(\boldsymbol{r})\right] |\psi_{u}^{(i)}\rangle |\psi_{l}^{(i)}\rangle |a\rangle.$$
(5.1.71)

The transition of the atom from the excited state $|a\rangle$ to the ground state $|b\rangle$ changes the field state in the respective cavity to $|\psi_{u}^{(f)}\rangle$ or $|\psi_{l}^{(f)}\rangle$. Thus, the final state

W(**r**)



Fig. 5.1.9. Cavity fields as which-path detectors in a doubleslit experiment. An atom in the excited state passing through the left double-slit can deposit its excitation in one of the two microwave cavities. The contrast of the interference structure on the screen depends on the initial field state in the cavities. A Fock state provides which-path information and no interference emerges (dashed line). For a coherent state no which-path information is available and we observe interference (solid line).

$$|\Psi^{(f)}(\boldsymbol{r})\rangle = \left[\phi_{u}(\boldsymbol{r})|\psi_{u}^{(f)}\rangle|\psi_{l}^{(i)}\rangle + \phi_{l}(\boldsymbol{r})|\psi_{u}^{(i)}\rangle|\psi_{l}^{(f)}\rangle\right]|b\rangle$$
(5.1.72)

of the system after the interaction reflects the fact that the center-of-mass motion is entangled with the field states. The probability $W(\mathbf{r})$ to find the atom at the position \mathbf{r} independent of the cavity fields is then

$$W(\boldsymbol{r}) = |\phi_{\rm u}(\boldsymbol{r})|^2 + |\phi_{\rm l}(\boldsymbol{r})|^2 + \left(\phi_{\rm u}(\boldsymbol{r})\phi_{\rm l}^*(\boldsymbol{r})\langle \psi_{\rm u}^{\rm (f)} | \psi_{\rm u}^{\rm (i)} \rangle \langle \psi_{\rm l}^{\rm (i)} | \psi_{\rm l}^{\rm (f)} \rangle + \text{c.c.}\right).$$
(5.1.73)

The size of the interference term is determined by the scalar product between the initial and the final quantum states of the two cavities. When we start from a coherent state of large average photon number the addition of one photon is not changing the field significantly and the scalar product is almost unity. In this case, the amplitude of the interference terms displays a maximum. However, when we start from a Fock state in both cavities, that is a state of well-defined photon number, the addition of one photon creates another Fock state which is orthogonal to the initial field state. As a consequence the interference terms vanish. This behavior is consistent with the principle of complementarity. In the case of the Fock state we can tell the path of the atom by observing the photon deposited in the cavity. In the case of a coherent state the broad photon distribution does not allow us to observe the addition of the photon and we can not obtain path information.

We conclude by mentioning that at no place in this argument have we made use of random phase disturbances. Nevertheless, a heated discussion objecting to this point of view [94Sto] has emerged but has been decided in favor of entanglement by an experiment [98Due].

5.1.9.3 Entanglement

The state description of classical mechanics is deeply rooted in our experience trained by a macroscopic world. Indeed, the state of any mechanical system in classical physics is completely determined when we specify or measure positions and momenta of all particles involved. As a consequence, a classical state describes deterministically all properties of each single particle in the total system. Moreover, this information on any single part of a classical ensemble does not depend on what measurements are to be performed on the other parts. This property reflects the typical local and realistic character of classical physics. Even though we have described it here for classical mechanics, it can be stated equally well for other classical branches of physics such as electrodynamics and thermodynamics.

Quantum theory strongly challenges this classical picture of our world. A quantum state is defined by a vector $|\Psi\rangle$ in Hilbert space. Characteristic properties of a quantum system can be entangled between its different subsystems. None of the subsystems alone usually suffices to measure these properties, since the results of the measurements depend on the observable that is measured at the other subsystems.

This amazing peculiarity of the quantum-mechanical state description was first clearly expressed and criticized in the famous paper [35Ein] by Albert Einstein, Boris Podolsky and Nathan Rosen (EPR). Their gedanken experiment shows that the quantum mechanics of entangled systems contains elements that are in conflict with a local and realistic view. EPR concluded that quantum mechanics must be an incomplete theory which one day should be replaced by a deeper theory based on further still unrecognized variables, the so-called hidden variables. The word "entanglement" itself was first used by Erwin Schrödinger in yet another famous paper [35Sch] of the year 1935.

5.1.9.4 Bell inequality

In 1964 John S. Bell studied the most general consequences of such a hidden-variable theory or Local Realistic Theory (LRT). He realized [64Bel] that the requirements of a LRT impose strong limitations on the possible correlations between the subparts of a two-particle system. The quantitative formulation of these limits is given by the famous Bell inequalities which are the topic of the present section [78Cla, 93Mer, 01Wer].

We imagine the most elementary composite system which consists of only two particles each of which has two distinguishable properties. That is, each particle can be described by a dichotomic variable. We stipulate that these correlated particles, after being prepared by a source Q, are sent to two separated observers A and B. A LRT assumes that the measurements of observers Aand B are fully described by observables $a(\alpha, \lambda)$ and $b(\beta, \lambda)$ which depend locally on the chosen measurement parameters α and β as well as on the hidden variables λ .

The value of such an observable is determined when we specify α or β and λ . Moreover, since we have simply assumed dichotomic properties we can denote these values by ± 1 . When each observer tunes his apparatus to two directions α_i and β_i (i = 1, 2) we have

$$|(a_1 + a_2)b_1 + (a_1 - a_2)b_2| = 2$$
(5.1.74)

with the abbreviations $a_i \equiv a(\boldsymbol{\alpha}_i, \lambda)$ and $b_i \equiv b(\boldsymbol{\beta}_i, \lambda)$.

Equation (5.1.74) holds true for any given hidden-variable set λ . If the source Q emits correlated particles described by a certain normalized distribution $\rho(\lambda)$ we immediately find

$$\left| \int_{\Lambda} d\lambda \,\rho(\lambda) \left[(a_1 + a_2) \,b_1 + (a_1 - a_2) \,b_2 \right] \right| \le 2 \tag{5.1.75}$$

or using the LRT correlation functions

$$\langle a_i b_j \rangle_{\text{LRT}} \equiv \int_{\Lambda} d\lambda \,\rho(\lambda) a(\boldsymbol{\alpha}_i, \lambda) b(\boldsymbol{\beta}_j, \lambda) \tag{5.1.76}$$

we arrive at the Bell inequality

$$|\langle a_1 b_1 \rangle_{\text{LRT}} + \langle a_2 b_1 \rangle_{\text{LRT}} + \langle a_1 b_2 \rangle_{\text{LRT}} - \langle a_2 b_2 \rangle_{\text{LRT}}| \le 2.$$
(5.1.77)

This inequality limits the correlations that can be described within a LRT. It is important to emphasize that inequalities of this type have nothing to do with quantum mechanics.

However, we may ask for the quantum-mechanical predictions for possible correlations between dichotomic observables of two systems. In fact, two-level systems (e.g. spin $\frac{1}{2}$ -systems) are dichotomic and the corresponding operator for a system A(B) is given by $\hat{a}(\alpha) \equiv \alpha \cdot \hat{\sigma}$ $(\hat{b}(\beta) \equiv \beta \cdot \hat{\sigma})$ with the usual Pauli operator $\hat{\sigma} \equiv (\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z)$, (5.1.43). The unit vector $\alpha(\beta)$ describes the measurement direction. If two systems A and B are prepared in a quantum state $|\Psi\rangle_{AB}$ the corresponding correlations are given by

$$\langle \Psi | \hat{a}(\boldsymbol{\alpha}_i) \hat{b}(\boldsymbol{\beta}_j) | \Psi \rangle \equiv \langle \hat{a}_i \hat{b}_j \rangle_{\text{QM}} \,. \tag{5.1.78}$$

Hence we can construct the same sum of correlation functions as in (5.1.77). When we calculate this sum for the a = l singlet state

$$|\Psi\rangle_{AB} = \frac{1}{\sqrt{2}} \Big(|\uparrow\rangle_A|\downarrow\rangle_B - |\downarrow\rangle_A|\uparrow\rangle_B\Big)$$
(5.1.79)

and the measurement directions $\alpha_1 = (0, 0, 1)$, $\alpha_2 = (1, 0, 0)$, $\beta_1 = (1/\sqrt{2}, 0, 1/\sqrt{2})$, and $\beta_2 = (-1/\sqrt{2}, 0, 1/\sqrt{2})$ we arrive at

$$\left| \langle \hat{a}_1 \hat{b}_1 \rangle_{\rm QM} + \langle \hat{a}_2 \hat{b}_1 \rangle_{\rm QM} + \langle \hat{a}_1 \hat{b}_2 \rangle_{\rm QM} - \langle \hat{a}_2 \hat{b}_2 \rangle_{\rm QM} \right| = 2\sqrt{2} \,. \tag{5.1.80}$$

Consequently, quantum mechanics allows for correlations that are stronger than the ones in any local realistic theory. The essential question now is whether dichotomic microsystems provided by nature show two-particle correlations that agree with the predictions of a LRT, (5.1.77) or with those of quantum mechanics, (5.1.80). A variety of such Bell-type experiments have been performed in the past with ever increasing accuracy [82Asp, 92Bre, 98Wei, 01Row]. They all support the quantum result, (5.1.80), and lead to the fundamental conclusion that quantum mechanics is a nonlocal and nonrealistic, but correct description of nature.

5.1.10 New frontiers

The material presented in the preceding sections defines the classical topics of quantum optics. We dedicate the final section of this review to a summary of recent developments in three newly emerging and rapidly moving fields of quantum optics: atom optics, Bose–Einstein condensation and quantum information.

5.1.10.1 Atom optics in quantized fields

One prediction of quantum mechanics is the wave nature of massive particles, that is we can associate a de Broglie matter wave with their center-of-mass motion. Due to its similarities to light optics this branch of atomic physics is called atom optics. In Sect. 5.1.4 we have derived the Hamiltonian, (5.1.46), describing the quantum dynamics of a two-level atom interacting with a quantized light field. Since this Hamiltonian also includes the center-of-mass motion we now briefly outline some consequences of this theory.

In the interaction of an atom with quantized light the roles of matter and light are interchanged. Indeed, classically atoms are treated as particles and light as waves. However, we now bring out the wave nature of matter and the particle nature of light. This combination has opened a new field namely atom optics in quantized light fields [90Kaz, 97Her, 99Fre, 01Mey] which is the marriage of the two fields atom optics and cavity quantum electrodynamics.

The state vector

$$|\Phi(t)\rangle = \sum_{n=0}^{\infty} \int dV \left[\Phi_{a,n-1}(\boldsymbol{R},t)|a,n-1\rangle + \Phi_{b,n}(\boldsymbol{R},t)|b,n\rangle\right] |\boldsymbol{R}\rangle$$
(5.1.81)

describes the combined system of the center-of-mass motion, internal states of a resonant two-level atom and the states of the electromagnetic field. Its time dependence follows from the Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} \Phi_n^{(\pm)}(\boldsymbol{R}, t) = \left[\frac{\hat{\boldsymbol{P}}^2}{2M} + U_n^{(\pm)}(\boldsymbol{R})\right] \Phi_n^{(\pm)}(\boldsymbol{R}, t)$$
(5.1.82)

with the help of the Hamiltonian, (5.1.46). Here we have introduced the dressed-state amplitudes $\Phi_n^{(\pm)} \equiv (\Phi_{b,n} \pm \Phi_{a,n-1})/\sqrt{2}$ and have defined the potentials

$$U_n^{(\pm)}(\mathbf{R}) \equiv \pm \boldsymbol{\wp} \cdot \boldsymbol{u}(\mathbf{R}) \,\mathcal{E}_0 \sqrt{n} \,. \tag{5.1.83}$$

Hence the probability amplitudes $\Phi_n^{(\pm)}$ satisfy a Schrödinger equation corresponding to a particle of mass M moving in the potentials $U_n^{(\pm)}$. These potentials are formed by the scalar product between



Fig. 5.1.10. A beam of resonant atoms propagating initially along the z-axis interacts with the light field in a rectangular cavity (left). Different Fock states deflect atoms in different directions and focus them at different points. This effect stands out most clearly in the contour plot (right) of the probability of finding an atom at the point with coordinates x and z. The Gaussian atomic beam centered at x = 0 leaves the cavity at z = 0. The field is in a coherent state of average photon number $\langle \hat{n} \rangle = 1$. The undeflected and unfocused partial wave associated with the cavity vacuum state represents the profile of the incident beam. The deflected partial waves associated with different photon states of the field focus at different points.

the dipole moment \wp and the mode function $u(\mathbf{R})$. Moreover, they scale with the vacuum electric field strength \mathcal{E}_0 and the square root of the photon number. Since the photon number n can only assume integer values we obtain a discrete set of potentials. Hence the granular structure of the radiation field manifests itself in a discrete superposition of potentials.

We can already gain a basic understanding of the corresponding effects by using a classical picture. The atom in each single potential feels a force that is proportional to the gradient of the potential. This force manifests itself in a deflection of atoms which is maximal when the atom starts at a node of the mode function. For a quantized field in a coherent state for example we have a discrete superposition of number states which translates into a discrete set of deflected beams [92Her] as shown in Fig. 5.1.10.

At an anti-node there is no force since the gradient vanishes. Nevertheless, quantum-mechanically there is an effect on the wave packet: The appropriate quadratic potential acts as a lens. Due to the superposition of potentials of discrete steepnesses we find a superposition of foci – a quantum lens [94Ave].

Unfortunately, so far no experimental results for the deflection of atoms in quantized fields are available yet. However, deflection [92Sle1], focusing [92Sle2] and interferometry [95Ras] of atomic de Broglie waves due to a *lass al* light field has been demonstrated experimentally. This case follows from (5.1.82) in the limit of a single photon number state $|n_0\rangle$, where we have only a single potential $U_{n_0}^{(\pm)}$ and n_0 is determined by the intensity of the classical light field.

5.1.10.2 Bose–Einstein condensation

Even more interesting is the case when the de Broglie wavelength of the atoms becomes of the order of their separation. Then the individual atoms loose their identity and their wave functions start to overlap. In the case of bosonic atoms there exists Bose–Einstein condensation where the atoms are in the ground state of the relevant Hamiltonian.

5.1.10.2.1 History

In 1924 Satyendra Nath Bose [24Bos] published his calculations on statistical properties of particles which are now known as bosons. More precisely, he presented an alternative derivation of Planck's radiation law. Albert Einstein [24Ein, 25Ein] extended this work to massive particles and predicted a phase transition in an ideal gas consisting of massive bosons – the famous Bose– Einstein condensation, where the ground state of the system becomes macroscopically occupied. Weakly interacting Bose gases were investigated by Nikolai N. Bogoliubov in 1947 [47Bog]. He showed that Bose–Einstein condensation was not much altered by weak interactions. However, the long-wavelength response of a Bose–Einstein condensate consisting of weakly interacting particles is completely different from an ideal gas.

In 1938 Fritz London suggested that Bose–Einstein condensation is responsible for the superfluidity of liquid ⁴He [38Lon]. However, due to the strong interactions between the atoms, only a small fraction of ⁴He atoms are condensed as has been observed in neutron scattering experiments later on. Superfluid ⁴He was therefore never accepted as a convincing proof for Bose–Einstein condensation. Other systems, where evidence for Bose–Einstein condensation was found, are excitons in optically pumped semiconductors [93Lin]. Experiments with spin-polarized atomic hydrogen started already in 1980 [80Sil], but Bose–Einstein condensation was achieved only recently [98Fri]. In 1995 clear evidence of Bose–Einstein condensation was found in dilute atomic gases consisting of alkali atoms [95And, 95Bra, 95Dav].

For a review on Bose–Einstein condensation in atomic gases, see the proceedings of the 1998 Varenna summer school [99Ing] and a collection of articles on ultracold matter in Nature (Nature **416** (2002) 205–246) [02Ang, 02Bur, 02Chu, 02Mon, 02Rol, 02Ude]. An overview of the theory of Bose-condensed atomic gases can be found in several review articles [98Par, 99Dal, 01Leg, 02Ste] and text books [01Mey, 01Pet, 03Pit].

5.1.10.2.2 Bose–Einstein condensation in dilute atomic gases

Roughly speaking, the condition for Bose–Einstein condensation is that the de Broglie wavelength of the atoms becomes comparable to the average distance between the atoms. When we want to have Bose–Einstein condensation in a dilute gas we therefore have to cool the atoms, that is increase their de Broglie wavelength.

In order to achieve Bose–Einstein condensation for gases with current technology two steps are necessary. First the atoms are trapped in a magneto-optical trap and cooled by lasers [92Dal, 01Bar, 02Met]. It turned out that this step is not sufficient to reach the condition for Bose– Einstein condensation. Therefore, the atoms are transformed into a purely magnetic trap and evaporative cooling is used to finally achieve Bose–Einstein condensation. Evaporative cooling requires collisions between the atoms. These collisions, however, open other decay channels for trapped atoms, for example formation of molecules via three-body collisions. Therefore, it is not obvious which elements are promising candidates for Bose–Einstein condensation. It turned out that ²³Na and ⁸⁷Rb are particularly good species to achieve Bose–Einstein condensation. These elements are now widely used to generate Bose–Einstein condensates consisting of several million atoms. Another element worth mentioning is ⁴He for which a Bose–Einstein condensate consisting of a dilute gas of helium atoms in the 2^3S_1 metastable state has been observed [01Per].

5.1.10.2.3 Gross–Pitaevskii equation

Many properties of Bose–Einstein condensates well below the transition temperature, in particular interference effects and dynamical properties, can be explained with the help of the Gross–Pitaevskii equation [61Gro, 61Pit, 63Gro]

Ref. p. 39]

$$i\hbar\frac{\partial}{\partial t}\psi(\boldsymbol{r},t) = \left(-\frac{\hbar^2}{2M}\Delta + V(\boldsymbol{r}) + \frac{4\pi\hbar^2 a}{M}|\psi(\boldsymbol{r},t)|^2\right)\psi(\boldsymbol{r},t)$$
(5.1.84)

or its time-independent version. Here $\psi(\mathbf{r}, t)$ is the macroscopic wave function of the condensate normalized to the number of particles, $V(\mathbf{r})$ is the trap potential (usually a harmonic potential), and a is the s-wave scattering length which takes into account the interaction between the atoms. For repulsive interaction a is positive whereas it is negative for attractive interaction. For homogeneous systems, that is $V(\mathbf{r}) = 0$, equations of this type have already been studied intensively in the 1960s. In order to describe the properties of condensates generated in current experiments the influence of trap potential $V(\mathbf{r})$ has to be taken into account. Based on the Gross-Pitaevskii equation, collective excitations, quantized vortices, and solitons in trapped Bose-Einstein condensates were predicted.

Generalizations of the Gross–Pitaevskii equation (5.1.84) are used to describe multi-component condensates and the manipulation of internal states of the atoms via external driving fields.

For more details on the application of the Gross-Pitaevskii equation to trapped Bose-condensed atoms, see [98Par, 99Dal, 01Leg, 01Mey, 01Pet, 02Ste, 03Pit].

5.1.10.2.4 Experiments with Bose–Einstein condensates

s. The first experiments by Eric A. Cornell, Carl E. Wieman and Wolfgang Ketterle have shown that there was a phase transition in their sample. However, there was no direct proof for coherent matter waves. In another experiment [97And] interference fringes in the overlap region of two condensates were observed, a clear indication of the coherence of matter waves. Since a Bose–Einstein condensate is a source for coherent matter waves, a potential application is an atom laser, a device analogous to an optical laser which instead of light emits matter waves. The first prototype of an atom laser was realized in 1996 [97Mew]. A radio-frequency field was used to change the internal state of the atoms in such a way that they were coupled out of the trap without loosing their coherence properties. Using similar ideas, several other groups demonstrated that atom lasers based on Bose–Einstein condensates can indeed provide an intense source for coherent matter waves [98And, 99Blo, 99Hag]. These results are summarized in Fig. 5.1.11.



Fig. 5.1.11. Atom lasers at MIT (5 mm), MPQ (2 mm), Yale (0.5 mm), and NIST (1 mm) (left to right). In the first three figures the atoms are coupled out of the trap and move downwards under the influence of gravity. In the right figure Raman pulses transfer momentum to the atoms when they are coupled out of the trap and gravity acts perpendicular to the figure. The numbers in brackets are the heights of each image. Taken from [00Ess].

In optics the invention of the laser gave rise to nonlinear optics. A well-known effect is four-wave mixing where coherent light with a new frequency is generated. Bose–Einstein condensates as a coherent source for matter waves opened the way for nonlinear atom optics. In 1999 Deng et al. [99Den] demonstrated that it is possible to mix three matter waves to produce a fourth wave with different momentum, that is four-wave mixing.

s. s. s. s. The theory of superfluidity predicts collective excitations, quantized vortices, and the existence of second sound in a condensate [93Gri, 99Dal, 01Pet]. Furthermore, from the nonlinear structure of the Gross–Pitaevskii equation, solitons are expected. These phenomena have been observed in various experiments after Bose–Einstein condensation in alkali vapors was achieved [96Jin, 96Mew, 99Bur, 99Mat, 00Den, 00Mad, 01Abo].

 $s \ la$. Bose–Einstein condensates can be used to demonstrate a phase transition from a superfluid to a Mott insulator [02Gre]. In order to observe such a quantum phase transition a Bose–Einstein condensate consisting of atoms with repulsive interaction is loaded into a threedimensional optical lattice generated by laser fields. Depending on the depth of the lattice potential, the atoms either form a condensate where the atoms can freely move though the lattice or a Mott insulator where they are localized at a lattice site and the motion of the atoms through the lattice is blocked. The depth of the lattice potential can easily be controlled by changing the intensity of the lasers.

We conclude our discussion of Bose–Einstein condensation by mentioning that Feshbach resonances [62Fes] can be used to control the *s*-wave scattering length and therefore the interaction between the atoms via external magnetic fields [98Ino]. For other experiments with Bose–Einstein condensates such as matter-wave amplification, interaction of light with Bose–Einstein condensates, multi-component condensates we refer to the literature at the end of Sect. 5.1.10.2.1 and the references therein.

5.1.10.3 Quantum information

5.1.10.3.1 Quantum teleportation

The strong and nonlocal correlations between quantum systems are not just a deep and striking fact of nature. They can be even exploited for highly non-classical tasks like quantum teleportation. It is crucial for any implementation to have a good source of entangled particles. Quantum-optical sources of polarization entangled photons [97Bou] as well as momentum entangled photons [98Bos] have been used successfully to perform quantum teleportations [93Ben] in the laboratory. Moreover, continuous variable teleportation [94Vai, 98Bra] has also been demonstrated experimentally [98Fur]. We shortly describe the basis of the corresponding protocol.

The sender, Alice, has a particle 1 in the state

$$|\psi\rangle_1 = \alpha |0\rangle_1 + \beta |1\rangle_1, \qquad (5.1.85)$$

where the orthogonal states $|0\rangle_1$ and $|1\rangle_1$ span a two-dimensional Hilbert space with complex amplitudes α and β that are normalized, $|\alpha|^2 + |\beta|^2 = 1$. In contrast to the preceding section we have now adopted the typical notation of quantum information theory by setting $|\uparrow\rangle \equiv |0\rangle$ and $|\downarrow\rangle \equiv |1\rangle$.

Alice's aim is to send the state $|\psi\rangle_1$ to Bob, the receiver, without sending particle 1 itself. Quantum teleportation achieves this by using an ancillary pair of particles 2 and 3 prepared in the entangled state

$$|\Psi^{-}\rangle_{23} = \frac{1}{\sqrt{2}} \left(|0\rangle_{2}|1\rangle_{3} - |1\rangle_{2}|0\rangle_{3} \right).$$
(5.1.86)



Fig. 5.1.12. Basic elements of quantum teleportation. An EPR source distributes the entangled particles 2 and 3 between Alice and Bob. Alice performs a Bell-type measurement on particle 2 and particle 1 whose state $|\psi\rangle$ will be teleported. By receiving Alice's measurement result via a classical channel Bob can in fact tune his particle to be finally in state $|\psi\rangle$.

We assume (see Fig. 5.1.12) that this pair of entangled particles stems from a common source. After preparation particle 2 is sent to Alice whereas particle 3 is kept by Bob.

The total state of all three particles reads

$$|\Psi\rangle_{123} = |\psi\rangle_1 \otimes |\Psi^-\rangle_{23} \tag{5.1.87}$$

and clearly particle 1 is neither correlated with particle 2 nor with particle 3. Formally we can, however, rewrite the total state

$$|\Psi\rangle_{123} = \frac{1}{2} \Big\{ |\Psi^{-}\rangle_{12} \Big(-\alpha |0\rangle_{3} - \beta |1\rangle_{3} \Big) + |\Psi^{+}\rangle_{12} \Big(-\alpha |0\rangle_{3} + \beta |1\rangle_{3} \Big) + |\Phi^{-}\rangle_{12} \Big(\alpha |1\rangle_{3} + \beta |0\rangle_{3} \Big) + |\Phi^{+}\rangle_{12} \Big(\alpha |1\rangle_{3} - \beta |0\rangle_{3} \Big) \Big\}$$
(5.1.88)

using the orthogonal Bell-basis states

$$|\Psi^{\pm}\rangle_{12} = \frac{1}{\sqrt{2}} \Big(|0\rangle_1 |1\rangle_2 \pm |1\rangle_1 |0\rangle_2 \Big), \quad |\Phi^{\pm}\rangle_{12} = \frac{1}{\sqrt{2}} \Big(|0\rangle_1 |0\rangle_2 \pm |1\rangle_1 |1\rangle_2 \Big)$$
(5.1.89)

for particles 1 and 2 owned by Alice. She can now perform a Bell-state measurement and thereby project the state $|\Psi\rangle_{123}$, (5.1.88), on one of the four states, (5.1.89). As a result of this measurement Bob's particle 3 will be prepared in one of the states directly related to Alice's result according to (5.1.88). For example, if Alice finds the result Φ^- , Bob's particle will be in state $\alpha |1\rangle_3 + \beta |0\rangle_3$ which looks similar to the original state, (5.1.85), but the basis states are flipped. This flip can be repaired by Bob with a simple unitary transformation $\hat{U}_{\Phi^-} = \hat{\sigma}_x = |0\rangle\langle 1| + |1\rangle\langle 0|$. This works similarly for all four measurement results that Alice can get. All that Alice has to do is to inform Bob via a classical channel about her measurement result. This enables him to choose the correct unitary transformation from the four possible ones: $\hat{U}_{\Psi^-} = 1$, $\hat{U}_{\Psi^+} = \hat{\sigma}_z = |0\rangle\langle 0| - |1\rangle\langle 1|$, $\hat{U}_{\Phi^-} = \hat{\sigma}_x = |0\rangle\langle 1| + |1\rangle\langle 0|$, and finally $\hat{U}_{\Phi^+} = \hat{\sigma}_y = i(|1\rangle\langle 0| - |0\rangle\langle 1|)$. After applying this transformation he obtains the state $|\psi\rangle_3 = \alpha |0\rangle_3 + \beta |1\rangle_3$ up to an overall phase and hence teleportation is completed.

This is a perfect example how quantum information can be sent with the help of entanglement. Moreover, it is also possible to transmit classical information (classical bits) by using entangled states between the communicating parties. The corresponding protocols of the so-called quantum dense coding [92Ben2, 96Mat] even show that this can be done more effectively than with any imaginable classical approach. This clearly shows the fascinating properties of quantum systems that can be prepared and controlled with quantum-optical means.

Another example for the far-reaching prospects of quantum information processing is quantum cryptography which we describe in the next paragraph.

37

5.1.10.3.2 Quantum cryptography

Secret messages are only as secure as the keys used to encrypt them. Classical cryptography has developed fascinating methods to encode a message using certain keys, but it has no means to ensure absolute security of those keys. Quantum cryptography closes this gap by showing how to generate and simultaneously exchange a secure key based on the laws of quantum physics.

The basic idea is that in quantum mechanics any measurement process changes the properties of a system. Consequently, whatever strategy an eavesdropper (normally called Eve) chooses, she will leave a trace on the quantum objects that are used for transmitting a key. This trace can be detected by the sender (typically known as Alice) and the intended recipient (usually called Bob).

In order to understand this in more detail we shall discuss a particular protocol for quantum key generation and distribution known as B92 protocol [84Ben, 91Eke, 92Ben1]. The quantum objects used for the transmission of the key are two-level systems (e.g. spin $\frac{1}{2}$) with basis states $|\uparrow\rangle$ and $|\downarrow\rangle$. They are defined by the eigenvalue equations

$$\hat{\sigma}_{z}|\uparrow\rangle = +|\uparrow\rangle, \quad \hat{\sigma}_{z}|\downarrow\rangle = -|\downarrow\rangle \tag{5.1.90}$$

of the $\hat{\sigma}_z$ Pauli operator with eigenvalues ±1. In addition we need the superposition states

$$|\rightarrow\rangle \equiv \frac{1}{\sqrt{2}} \left(|\uparrow\rangle + |\downarrow\rangle\right), \quad |\leftarrow\rangle \equiv \frac{1}{\sqrt{2}} \left(|\uparrow\rangle - |\downarrow\rangle\right)$$
(5.1.91)

which fulfill the eigenvalue equations

$$\hat{\sigma}_x | \to \rangle = + | \to \rangle, \quad \hat{\sigma}_x | \leftarrow \rangle = - | \leftarrow \rangle.$$
 (5.1.92)

of the $\hat{\sigma}_x$ Pauli operator. In the B92 protocol Alice now generates a random bit sequence and encodes it on her quantum objects by sending $|\uparrow\rangle$ if the bit reads a = 0 and by sending $|\rightarrow\rangle$ if the bit reads a = 1. Also Bob prepares a purely random list of bits. According to this list he measures Alice's quantum objects. He chooses the $\hat{\sigma}_z$ basis if his bit reads b = 1 and the $\hat{\sigma}_x$ basis if his bit reads b = 0. Clearly he can find the eigenvalue +1 for all four possible bit combinations a and b, but the value -1 he can only find with 50% probability if the bit combination was a = b = 0 or a = b = 1. Bob will never record a -1 when his bit is different from Alice's. For the values -1 the bits in the two lists are therefore completely correlated. Hence in the final step of the B92 protocol Bob simply sends a copy of his measurement. $s \ l \ s$ to Alice (..., the measurement basis). He may even send this over a public channel. Alice and Bob then only keep those bits in their lists that correspond to a -1 measurement result. The sequence of these $y \ ... \ s$ forms two identical and completely random keys.

Moreover, they can check their lists for an eavesdropper by sacrifying some of their key bits and comparing them. Let us assume that Eve makes her own measurement on Alice's quantum particles in the $\hat{\sigma}_x$ or $\hat{\sigma}_z$ basis and sends on the results to Bob. If, for example, Alice has prepared a $|\uparrow\rangle$ object (a = 0) and Eve wrongly chooses the $\hat{\sigma}_x$ basis, she might find $|\rightarrow\rangle$ and send it to Bob. Hence Eve has disturbed the properties of the original quantum object and left a trace by her quantum measurement process. If Bob applies a $\hat{\sigma}_z$ basis (b = 1) he might read of the value -1 with 50% probability, even though the bit combination was a = 0 and b = 1. Therefore, Alice and Bob can use sufficiently many of their key bits and check for a = 0 and b = 1 or a = 1 and b = 0 combinations. If the rate of such results is high, they know that an eavesdropper was present and consequently they have to discard the complete key. We emphasize that a general analysis of eavesdropping strategies is quite complicated [98Bih, 99Lo, 00Nie, 00Sho, 01May] and not even completely carried through for all possible protocols. Nevertheless, the technology for achieving quantum cryptography in practice [95Hug, 96Mul, 02Hug, 02Stu] is already highly developed and even stands at the edge of being commercialized.

References for 5.1

- 1802You Young, Th.: Philos. Trans. R. Soc. (London) 92 (1802) 12.
- 1873Max Maxwell, J.C.: A treatise on electricity and magnetism, 2 Vols., Oxford: Clarendon Press, 1873.
- 1900Pla Planck, M.: Verh. Dtsch. Phys. Ges. 2 (1900) 202.
- 1901Pla Planck, M.: Ann. Phys. (Leipzig) 4 (1901) 553.
- 1904Tho Thomson, J.J.: Philos. Mag. 7 (1904) 237.
- 1905Ein Einstein, A.: Ann. Phys. (Leipzig) **17** (1905) 132.
- 17Ein Einstein, A.: Phys. Z. **18** (1917) 121.
- 24Bos Bose, S.: Z. Phys. **26** (1924) 178.
- 24Ein Einstein, A.: Sitzungsber. K. Preuss. Akad. Wiss. **1924** (1924) 261.
- 24Wen Wentzel, G.: Z. Phys. **22** (1924) 193.
- 25Bor1 Born, M.: Atommechanik, Berlin: Springer-Verlag, 1925.
- 25Bor2 Born, M., Heisenberg, W., Jordan, P.: Z. Phys. **35** (1925) 557.
- 25Bor3 Born, M., Jordan, P.: Z. Phys. **34** (1925) 858.
- 25Ein Einstein, A.: Sitzungsber. K. Preuss. Akad. Wiss. **1925** (1925) 3.
- 25Hei Heisenberg, W.: Z. Phys. **33** (1925) 879.
- 26Lew Lewis, G.N.: Nature (London) **118** (1926) 874.
- 26Pau Pauli, W., in: Handbuch der Physik XXIII, Geiger, H., Scheel, K., (eds.), Berlin: Springer-Verlag, 1926, 1.
- 26Sch1 Schrödinger, E.: Ann. Phys. (Leipzig) **79** (1926) 734.
- 26Sch2 Schrödinger, E.: Naturwissenschaften **14** (1926) 664.
- 27Dir Dirac, P.A.M.: Proc. R. Soc. (London) A **114** (1927) 243.
- 27Jor Jordan, P.: Z. Phys. **40** (1927) 661.
- 28Boh Bohr, N.: Nature (London) **121** (1928) 580.
- 30Bor Born, M., Jordan, P.: Elementare Quantenmechanik, Heidelberg: Springer-Verlag, 1930.
- 31Wei von Weizsäcker, C.F.: Z. Phys. 70 (1931) 114.
- 32Fer Fermi, E.: Rev. Mod. Phys. 4 (1932) 87.
- 33Boh Bohr, N., Rosenfeld, L.: Mat. Fys. Medd. K. Dan. Vidensk. Selsk. 12 (8) (1933).
- 35Dir Dirac, P.A.M.: The principles of quantum mechanics, Oxford: Clarendon Press, 1935.
- 35Ein Einstein, A., Podolsky, B., Rosen, N.: Phys. Rev. 47 (1935) 777.
- 35Sch Schrödinger, E.: Naturwissenschaften **23** (1935) 807; 823; 844.
- 38Lon London, F.: Nature (London) **141** (1938) 643.

41Wei	von Weizsäcker, C.F.: Z. Phys. 118 (1941) 489.
47Bog 47Lam	Bogoliubov, N.N.: J. Phys. USSR 11 (1947) 23. Lamb, W.E., Retherford, R.C.: Phys. Rev. 72 (1947) 241.
49Boh 49Kro	Bohr, N., in: Albert Einstein: Philosopher-Scientist, Schilpp, P.A. (ed.), Evanston: Library of Living Philosophers, 1949, 200. Kroll, N.M., Lamb, W.E.: Phys. Rev. 75 (1949) 388.
50Boh	Bohr, N., Rosenfeld, L.: Phys. Rev. 78 (1950) 794.
52Sch1 52Sch2	Schrödinger, E.: Brit. J. Philos. Sci. 3 (1952) 109. Schrödinger, E.: Brit. J. Philos. Sci. 3 (1952) 233.
56Han	Hanbury Brown, R., Twiss, R.Q.: Nature (London) 177 (1956) 27.
58Boh	Bohr, N.: Atomic physics and human knowledge, New York: Wiley, 1958.
61Gro 61Pit	Gross, E.P.: Nuovo Cimento 20 (1961) 454. Pitaevskii, L.P.: Zh. Eksp. Teor. Fiz. 40 (1961) 646; Sov. Phys. JETP (English Trans.) 13 (1961) 451.
62Fes	Feshbach, H.: Ann. Phys. (New York) 19 (1962) 287.
63Gla1 63Gla2 63Gro 63Jay 63Pau 63Sud	 Glauber, R.J.: Phys. Rev. 130 (1963) 2529. Glauber, R.J.: Phys. Rev. 131 (1963) 2766. Gross, E.P.: J. Math. Phys. N.Y. 4 (1963) 195. Jaynes, E.T., Cummings, F.W.: Proc. IEEE 51 (1963) 89. Paul, H.: Ann. Phys. (Leipzig) 11 (1963) 411. Sudarshan, E.C.G.: Phys. Rev. Lett. 10 (1963) 277.
64Bel 64Bor	Bell, J.S.: Physics 1 (1964) 195. Born, M., Wolf, E.: Principles of optics, Oxford: Pergamon Press, 1964.
65Bur	Burshtein, A.I.: Zh. Eksp. Teor. Fiz. 49 (1965) 1362; Sov. Phys. JETP (English Trans.)
65Gla	22 (1966) 939. Glauber, R.J., in: Qantum optics and electronics, DeWitt C., Blandin, A., Cohen-Tannoudji, C. (eds.), New York: Gordon and Breach, 1965, 331.
66Lax	Lax, M., in: Brandeis University Summer Institute Lectures, Chretin, M., Gross, E.P., Deser, S. (eds.), Vol. 2., New York: Gordon and Breach, 1966.
68Car	Carruthers, P., Nieto, M.M.: Rev. Mod. Phys. 40 (1968) 411.
69Cah 69Mol	Cahill, K.E., Glauber, R.J.: Phys. Rev. A 177 (1969) 1882. Mollow, B.R.: Phys. Rev. 188 (1969) 1969.
70Man	Mandel, L., Wolf, E. (eds.): Selected papers on coherence and fluctuations of light, Vol. I and II, New York: Dover Publications, 1970.
73Lou	Louisell, W.H.: Quantum statistical properties of radiation, New York: Wiley, 1973.

74Sar	Sargent, M. III., Scully, M.O., Lamb, W.E.: Laser Theory, Massachusetts: Addison-Wesley, Reading, 1974.
74Sch	Schuda, F., Stroud jr., C.R., Hercher, M.: J. Phys. B 7 (1974) L198.
75Deh 75Wu	Dehmelt, H.: Bull. Am. Phys. Soc. 20 (1975) 60. Wu, F.Y., Grove, R.E., Ezekiel, S.: Phys. Rev. Lett. 35 (1975) 1426.
76Car1 76Car2 76Har	Carmichael, H.J., Walls, D.F.: J. Phys. B 9 (1976) 1199. Carmichael, H.J., Walls, D.F.: J. Phys. B 9 (1976) L43. Hartig, W., Rasmussen, W., Schieder, R., Walther, H.: Z. Phys. A 278 (1976) 205.
77Kim 78Cla 78Kim 78Whe	 Kimble, H.J., Dagenais, M., Mandel, L.: Phys. Rev. Lett. 39 (1977) 691. Clauser, J.F., Shimony, A.: Rep. Prog. Phys. 41 (1978) 1881. Kimble, H.J., Dagenais, M., Mandel, L.: Phys. Rev. A 18 (1978) 201. Wheeler, J.A., in: Mathematical foundations of quantum theory, Marlow, A.R. (ed.), New York: Academic Press, 1978.
79Pau 79Whe	Pauli, W., in: Wissenschaftlicher Briefwechsel mit Bohr, Einstein, Heisenberg u.a., Vol. I, Hermann, A., von Meyenn, K., Weisskopf, V.F. (eds.), New York: Springer, 1979. Wheeler, J.A., in: Problems in the foundations of physics, International School of Physics "Enrico Fermi", Course LXXII, Amsterdam: North-Holland, 1979.
80Sil	Silvera, I.F., Walraven, J.T.M.: Phys. Rev. Lett. 44 (1980) 164.
82Asp 82Cre	Aspect, A., Dalibard, J., Roger, G.: Phys. Rev. Lett. 49 (1982) 1804. Cresser, J.D., Häger, J., Leuchs, G., Rateike, M., Walther, H., in: Dissipative systems in quantum optics, Bonifacio, R., Lugiato, L. (eds.), Topics in Current Physics Vol. 27, Berlin: Springer-Verlag, 1982, 21.
83Ber 83Whe	Bertolotti, M.: Masers and lasers. An historical approach, Bristol: Adam Hilger Ltd, 1983. Wheeler, J.A., Zurek, W.H. (eds.): Quantum theory and measurement, Princeton: Princeton University Press, 1983.
84Ben 84Hak 84Hil	 Bennett, C.H., Brassard, G., in: Proc. IEEE Int. Conference on Computers, Systems and Signal Processing, Los Alamitos: IEEE Press, 1984. Haken, H.: Laser theory, Heidelberg: Springer-Verlag, 1984. Hillery, M., O'Connell, R.F., Scully, M.O., Wigner, E.P.: Phys. Rep. 106 (1984) 121.
85Gar 85Mes 85Slu	Gardiner, C.W.: Handbook of stochastic methods, Berlin: Springer-Verlag, 1985. Meschede, D., Walther, H., Müller, G.: Phys. Rev. Lett. 54 (1985) 551. Slusher, R.E., Hollberg, L.W., Yurke, B., Mertz, J.C., Valley, J.F.: Phys. Rev. Lett. 55 (1985) 2409.
86Ber 86Fil 86Gar 86Gra 86Nag 86Sau	 Bergquist, J.C., Hulet, R.G., Itano, W.M., Wineland, D.J.: Phys. Rev. Lett. 57 (1986) 1699. Filipowicz, P., Javanainen, J., Meystre, P.: Phys. Rev. A 34 (1986) 3077. Gardiner, C.W.: Phys. Rev. Lett. 56 (1986) 1917. Grangier, P., Roger, G., Aspect, A.: Europhys. Lett. 1 (1986) 173. Nagourney, W., Sandberg, J., Dehmelt, H.: Phys. Rev. Lett. 56 (1986) 2797. Sauter, Th., Neuhauser, W., Blatt, R., Toschek, P.E.: Phys. Rev. Lett. 57 (1986) 1696.

42	References for 5.1
87All	Alley, C.O., Jakubowicz, O.G., Wickes, W.C., in: Proceedings of the 2nd interna- tional symposium on foundations of quantum mechanics, Namiki, M., Ohnuki, Y., Mu-
97Uol	rayama, Y., Nomura, S. (eds.), Tokyo: Physical Society of Japan, 1987. Hollmuth, T. Wolthor, H. Zaione, A. Schleich, W.D. Phys. Rev. A 25 (1087) 2522.
87Kim	Kimble H I Walls D F (eds): I Opt Soc Am B 4 (10) (1987) 1449
87Lou	Loudon, R., Knight, P.L. (eds.): J. Mod. Opt. 34 (6/7) (1987) 709.
87Lug	Lugiato, L.A., Scully, M.O., Walther, H.: Phys. Rev. A 36 (1987) 740.
87Sch	Schleich, W.P., Wheeler, J.A.: J. Opt. Soc. Am. B 4 (1987) 1715.
88Bre	Brendel, J., Schütrumpf, S., Lange, R., Martienssen, W., Scully, M.O.: Europhys. Lett. 5 (1988) 223.
88Mey	Meystre, P., Rempe, G., Walther, H.: Opt. Lett. 13 (1988) 1078.
89Asp	Aspect, A., Grangier, P., Roger, G.: J. Opt. 20 (1989) 119.
89Bal	Baldzuhn, J., Mohler, E., Martienssen, W.: Z. Phys. B 77 (1989) 347.
89Coh	Cohen-Tannoudji, C., Dupont-Roc, J., Grynberg, G.: Photons and atoms. An introduc- tion to quantum electrodynamics, New York: Wiley, 1989.
89Peg	Pegg, D.T., Barnett, S.M.: Phys. Rev. A 39 (1989) 1665.
89Ris	Risken, H.: The Fokker-Planck equation, Berlin: Springer-Verlag, 1989.
90Kaz	Kazantsev, A.P., Surdutovich, G.I., Yakovlev, V.P.: Mechanical action of light on atoms, Singapore: World Scientific, 1990.
90Pau	Paul, W.: Rev. Mod. Phys. 62 (1990) 531.
90Rem	Rempe, G., Schmidt-Kaler, F., Walther, H.: Phys. Rev. Lett. 64 (1990) 2783.
90Sho	Shore, B.W.: The theory of coherent atomic excitations, New York: Wiley, 1990.
91Bre	Brendel, J., Lange, R., Mohler, E., Martienssen, W.: Ann. Phys. (Leipzig) 48 (1991) 26.
91Eke	Ekert, A.: Phys. Rev. Lett. 67 (1991) 661.
91Gar	Gardiner, C.W.: Quantum noise, Berlin: Springer-Verlag, 1991.
91Lou	Loudon, R.: The quantum theory of light, 2nd edition, Oxford: Clarendon Press, 1991.
91Mey	Meystre, P., Sargent, M. III.: Elements of quantum optics, Berlin: Springer-Verlag, 1991.
91Sch	Schleich, W.P., Pernigo, M., Fam Le Kien: Phys. Rev. A 44 (1991) 2172.
91Scu 91Zur	Zurek, W.H.: Phys. Today 44 (1991) 36.
02Bon1	Bonnott, C.H. Phys. Boy. Lett. 68 (1002) 3121
92Ben2	Bennett, C.H. Wiesner, S.J. Phys. Rev. Lett. 69 (1992) 2881
92Blo	Blockley, C.A., Walls, D.F., Risken, H.: Europhys. Lett. 17 (1992) 509.
92Bre	Brendel, J., Mohler, E., Martienssen, W.: Europhys. Lett. 20 (1992) 575.
92Coh	Cohen-Tannoudji, C., Dupont-Roc, J., Grynberg, G.: Atom-photon interactions, New York: Wiley 1992
92Dal	Dalibard, J., Raimond, JM., Zinn-Justin, J. (eds.): Fundamental systems in quantum optics. Les Houches 1990, session LIII. Amsterdam: North-Holland, 1992
92Gia	Giacobino E Fabre C (eds): Appl Phys R 55 (3) (1992) 189
92Her	Herkommer, A.M., Akulin, V.M., Schleich, W.P.: Phys. Rev. Lett. 69 (1992) 3298.
92Sle1	Sleator, T., Pfau, T., Balykin, V., Carnal, O., Mlynek, J.: Phys. Rev. Lett. 68 (1992) 1996.
92Sle 2	Sleator, T., Pfau, T., Balykin, V., Mlynek, J.: Appl. Phys. B 54 (1992) 375.
93Ben	Bennett, C.H., Brassard, G., Crépeau, C., Jozsa, R., Peres, A., Wootters, W.K.: Phys. Rev. Lett. 70 (1993) 1895.

93Car	Carmichael, H.J.: An open systems approach to quantum mechanics, Heidelberg: Springer-Verlag 1993
93Gri	Griffin, A.: Excitations in a Bose-condensed liquid, Cambridge: Cambridge University Press, 1993.
93Hau 93Lin 93Mer 93Sch	 Haubrich, D., Höpe, A., Meschede, D.: Opt. Commun. 102 (1993) 225. Lin, J.L., Wolfle, J.P.: Phys. Rev. Lett. 71 (1993) 1222. Mermin, N.D.: Rev. Mod. Phys. 65 (1993) 803. Schleich, W.P., Barnett, S.M. (eds.): Quantum phase and phase dependent measurements; Phys. Scr. T 48 (1993) 5.
93Sho	Shore, B.W., Knight, P.L.: J. Mod. Opt. 40 (1993) 1195.
94Ave 94Ber 94Cir 94Moo 94Rai	 Averbukh, I.Sh., Akulin, V.M., Schleich, W.P.: Phys. Rev. Lett. 72 (1994) 437. Bergou, J., Hillery, M.: Phys. Rev. A 49 (1994) 1214. Cirac, J.I., Blatt, R., Parkins, A.S., Zoller, P.: Phys. Rev. A 49 (1994) 1202. Moore, W.: A life of Erwin Schrödinger, Cambridge: Cambridge University Press, 1994. Raithel, G., Wagner, C., Walther, H., Narducci, L., Scully, M.O., in: Advances in Atomic, Molecular, and Optical Physics, Supplement 2: Cavity Quantum Electrodynamics, Berman, P.B. (ed.), Boston: Academic Press, 1994, 57.
94Sto	Storey, E.P., Tan, S.M., Collet, M.J., Walls, D.F.: Nature (London) 367 (1994) 626.
94Vai 94Wal 94Wil	 Vaidman, L.: Phys. Rev. A 49 (1994) 1473. Walls, D.F., Milburn, G.J.: Quantum optics, Berlin: Springer-Verlag, 1994. Wilkens, M.: Phys. Rev. A 48 (1994) 570.
95And	Anderson, M.H., Ensher, J.R., Matthews, M.R., Wieman, C.E., Cornell, E.A.: Science 269 (1995) 198.
95Bra 95Dav	Bradley, C.C., Sackett, C.A., Tollett, J.J., Hulet, R.G.: Phys. Rev. Lett. 75 (1995) 1687. Davis, K.B., Mewes, MO., Andrews, M.R., van Druten, N.J., Durfee, D.S., Kurn, D.M., Ketterle, W.: Phys. Rev. Lett. 75 (1995) 3969.
95Her	Herzog, T.J., Kwiat, P.G., Weinfurter, H., Zeilinger, A.: Phys. Rev. Lett. 75 (1995) 3034.
95Hug	Hughes, R.J., Alde, D.M., Dyer, P., Luther, G.G., Morgan, G.L., Schauer, M.: Contemp. Phys. 36 (1995) 149.
95Lam	Lamb, W.E.: Appl. Phys. B 60 (1995) 77.
95Man	Mandel, L., Wolf, E.: Optical coherence and quantum optics, New York: Cambridge University Press, 1995.
95Ras	Rasel, E.M., Oberthaler, M.K., Batelaan, H., Schmiedmayer, J., Zeilinger, A.: Phys. Rev. Lett. 75 (1995) 2633.
95 Vog	Vogel, W., de Matos Filho, R.L.: Phys. Rev. A 52 (1995) 4214.
96Bar	Bardroff, P.J., Leichtle, C., Schrade, G., Schleich, W.P.: Phys. Rev. Lett. 77 (1996) 2198.
96Bru	Brune, M., Hagley, E., Dreyer, J., Maitre, X., Maali, A., Wunderlich, C., Raimond J.M., Haroche, S.: Phys. Rev. Lett. 77 (1996) 4887.
96Eng 96Giu	Englert, BG.: Phys. Rev. Lett. 77 (1996) 2154. Giulini, D., Joos, E., Kiefer, C., Kupsch, J., Stamatescu, IO., Zeh, H.D.: Decoherence and the appearance of a classical world in quantum theory, Heidelberg: Springer-Verlag, 1996.
96Jin	Jin, D.S., Ensher, J.R., Matthews, M.R., Wieman, C.E., Cornell, E.A.: Phys. Rev. Lett. 77 (1996) 420.
96Mat 96Mee	Mattle, K., Weinfurter, H., Kwiat, P.G., Zeilinger, A.: Phys. Rev. Lett. 76 (1996) 4656. Meekhof, D.M., Monroe, C., King, B.E., Itano, W.M., Wineland, D.J.: Phys. Rev. Lett. 76 (1996) 1796.

43

44	References for 5.1
96Mew	Mewes, MO., Andrews, M.R., van Druten, N.J., Kurn, D.M., Durfee, D.S., Townsend, C.G., Ketterle, W.: Phys. Rev. Lett. 77 (1996) 988.
96Mul	Muller, A., Zbinden, H., Gisin, N.: Europhys. Lett. 33 (1996) 334.
96Scu	Scully, M.O., Zubairy, M.S.: Quantum optics, New York: Cambridge University Press, 1996.
97And	Andrews, M.R., Townsend, C.G., Miesner, HJ., Durfee, D.S., Kurn, D.M., Ketterle, W.: Science 275 (1997) 637.
97Bou	Bouwmeester, D., Pan, JW., Mattle, K., Eibl, M., Weinfurter, H., Zeilinger, A.: Nature (London) 390 (1997) 575.
97Buz	Buzek, V., Drobný, G., Kim, M.S., Adam, G., Knight, P.L.: Phys. Rev. A 56 (1997) 2352.
97Hoe	Höffges, J.T., Baldauf, H.W., Eichler, T., Helmfrid, S.R., Walther, H.: Opt. Commun. 133 (1997) 170.
97Her 97Mew	Herkommer, A.M., Schleich, W.P.: Comments At. Mol. Phys. 33 (1997) 145. Mewes, MO., Andrews, M.R., Kurn, D.M., Durfee, D.S., Townsend, C.G., Ketterle,
97Sch	 W.: Phys. Rev. Lett. 78 (1997) 582. Schrade, G., Bardroff, P.J., Glauber, R.J., Leichtle, C., Yakovlev, V., Schleich, W.P.: Appl. Phys. B 64 (1997) 181.
98And	Anderson, B.P., Kasevich, M. A.: Science 282 (1998) 1686.
98Bih	Biham, E., Boyer, M., Brassard, G., van de Graaf, J., Mor, T.: quant-ph/9801022.
98Bos	Boschi, D., Branca, S., De Martini, F., Hardy, L., Popescu, S.: Phys. Rev. Lett. 80 (1998) 1121.
98Bra	Braunstein, S.L., Kimble, H.J.: Phys. Rev. Lett. 80 (1998) 869.
98Due	Dürr, S., Nonn, T., Rempe, G.: Nature (London) 395 (1998) 33.
98Fri	Fried, D.G., Killian, T.C., Willmann, L., Landhuis, D., Moss, S.C., Kleppner, D., Grey- tak, T.J.: Phys. Rev. Lett. 81 (1998) 3811.
98Fur	Furusawa, A., Sorensen, J.L., Braunstein, S.L., Fuchs, C.A., Kimble, H.J., Polzik, E.S.: Science 282 (1998) 706.
98Ino	Inouye, S., Andrews, M.R., Stenger, J., Miesner, HJ., Stamper-Kurn, D.M., Ketterle, W.: Nature (London) 392 (1998) 151.
98Lu	Lu, Z.H., Bali, S., Thomas, J.E.: Phys. Rev. Lett. 81 (1998) 3635.
98Par	Parkins, A.S., Walls, D.F.: Phys. Rep. 303 (1998) 1.
98Wei	Weihs, G., Jennewein, T., Simon, C., Weinfurter, H., Zeilinger, A.: Phys. Rev. Lett. 81 (1998) 5039.
98Win	Wineland, D.J., Monroe, C., Itano, W.M., Leibfried, D., King, B.E., Meekhof, D.M.: J. Res. Natl. Inst. Stand. Technol. 103 (1998) 259.
99Blo	Bloch, I., Hänsch, T.W., Esslinger, T.: Phys. Rev. Lett. 82 (1999) 3008.
99Bur	Burger, S., Bongs, K., Dettmer, S., Ertmer, W., Sengstock, K., Sanpera, A., Shlyap- nikov, G.V., Lewenstein, M.: Phys. Rev. Lett. 83 (1999) 5198.
99Dal	Dalfovo, F., Giorgini, S., Pitaevskii, L.P., Stringari, S.: Rev. Mod. Phys. 71 (1999) 463.
99Den	Deng, L., Hagley, E.W., Wen, J., Trippenbach, M., Band, Y., Julienne, P.S., Simsarian, J.E., Helmerson, K., Rolston, S.L., Phillips, W.D.: Nature (London) 398 (1999) 218.
99Fre	Freyberger, M., Herkommer, A.M., Krähmer, D.S., Mayr, E., Schleich, W.P., in: Advances in Atomic, Molecular, and Optical Physics 41 , Bederson, B., Walther, H. (eds.), Bestern, Academic Press, 1999, 142
99Hag	Boston: Academic Press, 1999, 143. Hagley, E.W., Deng, L., Kozuma, M., Wen, J., Helmerson, K., Rolston, S.L., Phillips, W.D.: Science 283 (1999) 1706.

99Ing	Inguscio, M., Stringari, S., Wieman, C.E. (eds.): Bose-Einstein condensation in atomic gases; Proceedings of the International School of Physics "Enrico Fermi", Course CXL, Amsterdam: IOS Proce 1000
99Lam	 Amsterdam: 105 Fress, 1999. Lamb, W.E., Schleich, W.P., Scully, M.O., Townes, C.H.: Rev. Mod. Phys. 71 (1999) 263.
99Lo	Lo. H., Chan, H.F.: Science 283 (1999) 2050.
99Mat	Matthews, M.R., Anderson, B.P., Haljan, P.C., Hall, D.S., Wieman, C.E., Cornell, E.A.:
	Phys. Rev. Lett. 83 (1999) 2498.
99Pau	Paul, H.: Photonen. Eine Einführung in die Quantenoptik, Stuttgart: B. G. Teubner, 1999.
99Wei	Weidinger, M., Varcoe, B.T.H., Heerlein, R., Walther, H.: Phys. Rev. Lett. 82 (1999) 3795.
99Zei	Zeilinger, A.: Rev. Mod. Phys. 71 (1999) S288.
00Den	Denschlag, J., Simsarian, J.E., Feder, D.L., Clark, C.W., Collins, L.A., Cubizolles, J., Deng, L., Hagley, E.W., Helmerson, K., Reinhardt, W.P., Rolston, S.L., Schneider, B.I., Phillips, W.D.: Science 287 (2000) 97.
00 Ess	Esslinger, T., Bloch, I., Hänsch, T.W.: Physikalische Blätter 56(2) (2000) 47.
00Mad	Madison, K.W., Chevy, F., Bretin, V., Dalibard, J.: Phys. Rev. Lett. 84 (2000) 806.
00 Mya	Myatt, C.J., King, B.E., Turchette, Q.A., Sackett, C.A., Kielpinski, D., Itano, W.M.,
	Monroe, C., Wineland, D.J.: Nature (London) 403 (2000) 269.
00Nie	Nielsen, M.A., Chuang, I.L.: Quantum computation and quantum information, Cam-
0001	bridge: Cambridge University Press, 2000.
00Sho	Shor, P.W., Preskill, J.: Phys. Rev. Lett. 85 (2000) 441.
00Var	Varcoe, B.T.H., Brattke, S., Weidinger, M., Walther, H.: Nature (London) 403 (2000) 743.
00Zan	von Zanthier, J., Becker, Th., Eichenseer, M., Nevsky, A.Yu., Schwedes, Ch., Peik, E., Walther, H., Holzwarth, R., Reichert, J., Udem, Th., Hänsch, T.W., Pokasov, P.V., Skvortsov, M.N., Bagayev, S.N.: Opt. Lett. 25 (2000) 1729.
01Abo	Abo-Shaeer, J.R., Raman, C., Vogels, J.M., Ketterle, W.: Science 292 (2001) 476.
01Bar	Bardou, F., Bouchaud, J.P., Aspect, A., Cohen-Tannoudji, C.: Lévy statistics and laser cooling, Cambridge: Cambridge University Press, 2001.
01Leg	Leggett, A.J.: Rev. Mod. Phys. 73 (2001) 307.
01May	Mayers, D.: Journal ACM 48 (2001) 351.
01Mey	Meystre, P.: Atom optics, New York: Springer-Verlag, 2001.
01Per	Pereira Dos Santos, F., Léonard, J., Wang, J., Barrelet, C.J., Perales, F., Rasel, E., Unnikrishnan, C.S., Leduc, M., Cohen-Tannoudji, C.: Phys. Rev. Lett. 86 (2001) 3459.
01Pet	Pethick, C.J., Smith, H: Bose-Einstein condensation in dilute gases, Cambridge: Cambridge University Press, 2001.
01Row	Rowe, M.A., Kielpinski, D., Meyer, V., Sackett, C.A., Itano, W.M., Monroe, C., Wineland, D.J.: Nature (London) 409 (2001) 791.
01Sch	Schleich, W.P.: Quantum optics in phase space, Weinheim: VCH-Wiley, 2001.
01Wer	Werner, R.F., Wolf, M.M.: Quantum Inf. Comput. ${\bf 1}$ (2001) 1.
02Ang	Anglin, J.R., Ketterle, W.: Nature (London) 416 (2002) 211.
02Bur	Burnett, K., Julienne, P.S., Lett, P.D., Tiesinga, E., Williams, C.J.: Nature (London) 416 (2002) 225.
02Chu	Chu, S.: Nature (London) 416 (2002) 206.
02Gre	Greiner, M., Mandel, O., Esslinger, T., Hänsch, T.W., Bloch, I.: Nature (London) 415 (2002) 39.
02Hug	Hughes, R.J., Nordholt, J.E., Derkacs, D., Peterson, C.G.: New J. Phys. 4 (2002) 43.

46	References for 5.1
02Kie	Kiesel, H., Renz, A., Hasselbach, F.: Nature (London) 418 (2002) 392.
02Met	Metcalf, H.J., van der Straten, P.: Laser cooling and trapping, New York: Springer-Verlag, 2002.
02Mon	Monroe, C.: Nature (London) 416 (2002) 238.
02Rol	Rolston, S.L., Phillips, W.D.: Nature (London) 416 (2002) 219.
02Spe	Spence, J.C.H.: Nature (London) 418 (2002) 377.
02Ste	Stenholm, S.: Phys. Rep. 363 (2002) 173.
02Stu	Stucki, D., Gisin, N., Guinnard, O., Ribordy, G., Zbinden, H.: New J. Phys. 4 (2002) 41.
02Ude	Udem, T., Holzwarth, R., Hänsch, T.W.: Nature (London) 416 (2002) 233.

Pitaevskii, L., Stringari, S.: Bose-Einstein Condensation, Oxford: Clarendon Press, 2003. 03Pit

6.1 Coherence

W. MARTIENSSEN, H. PAUL

6.1.1 Historical remarks

In its original meaning, coherence was just a synonym of ability to interfere. Two light beams were said to be (mutually) coherent, when they could be made to interfere, i.e. to produce an observable interference pattern. This requires the presence of field correlations on which a strict coherence concept therefore has to be based. Noteworthy are studies by Taylor [09Tay], who showed that an interference pattern can be observed at extremely low intensities, without loss of visibility, provided the exposition time is made long enough. Great progress came from the development of photoelectric detectors with very short resolving times which made it possible to observe not only spatial but also temporal interference (beating) [62Jav], and, moreover, both spatial and temporal intensity correlations. In their pioneering work, Brown and Twiss [56Bro] observed intensity correlations in star light, thus determining the angular diameter of fixed stars. Photomultipliers allowed to detect single photons, and hence to count photons [67Are] and detect photon coincidence counts [57Reb, 57Twi]. With the advent of the laser [60Mai, 61Jav] the experimenter had a new source producing light with unprecedented properties at hand. Relevant theoretical papers are Einstein's light-quanta hypothesis [1905Ein], his analysis of energy fluctuations in blackbody radiation [09Ein], von Laue's analysis of the degrees of freedom of light beams [14Lau], and Glauber's introduction of the concept of coherent states into the quantum theory of radiation [63Gla, 65Gla].

6.1.2 Basic concepts

6.1.2.1 Classical light

Classical light is a radiation field that can be correctly described classically in the sense that all quantum-mechanical field correlation functions can be exactly reproduced by classical averages. This condition is fulfilled when a positive-definite P function, as introduced by Glauber [65Gla], exists for the field in quantum-mechanical description, i.e., when the density operator for the field can be represented, in the Dirac formalism, as $\rho = \int P(\alpha_1, \alpha_2, \ldots) |\alpha_1\rangle |\alpha_2\rangle \ldots \langle \alpha_2|\langle \alpha_1| d^2\alpha_1 d^2\alpha_2 \ldots$ with $P(\alpha_1, \alpha_2, \ldots) \ge 0$ ($|\alpha\rangle$: single-mode coherent state). Typical examples are chaotic, in particular thermal light and radiation from lasers as well as radio and radar transmitters. Chaotic light is produced by natural sources such as the sun or stars, or by any kind of conventional sources. The chaotic nature of the radiation is preserved when linear filters, e.g. apertures, spectral filters and polarizers, are applied.

6.1.2.2 Non-classical light

Non-classical light is a radiation field that has no positive-definite P representation. Another, however only sufficient, criterion is the appearance of negativities in the quantum-mechanical Wigner function of the field. Examples are fields with a fixed photon number, in particular the propagating wave packets generated in spontaneous emission (single photons) and parametric down-conversion (photon pairs, see Sect. 6.1.4.5), antibunched and squeezed light.

6.1.2.3 Field mode

A field mode is in its original meaning an eigenmode of a resonator, i.e. a monochromatic standingwave field whose spatial distribution and frequency are determined by the resonator geometry; in a generalized sense a polarized, monochromatic field described by a special solution of Maxwell's equations, e.g. a traveling plane wave or a spherical wave. Formally, a field mode can be treated equivalently to a harmonic oscillator.

6.1.2.4 Single-mode field

A single-mode field is an excitation of a single mode of an optical resonator, or, more generally, the field within a single laser pulse or in a spatial-temporal region that can be identified with a coherence volume (see Sect. 6.1.3.2). The common property of single-mode fields is that they can be described by a single complex amplitude and, hence, have only one degree of freedom of motion.

6.1.2.5 Electric field strength

The Fourier representation of the electric field strength $E(\mathbf{r},t) = \int_{-\infty}^{+\infty} E(\mathbf{r},\nu) e^{-2\pi i\nu t} d\nu$ can be written as

$$E(\mathbf{r},t) = E^{(+)}(\mathbf{r},t) + E^{(-)}(\mathbf{r},t) , \qquad (6.1.1)$$

$$E^{(+)}(\mathbf{r},t) = \int_0^{+\infty} E(\mathbf{r},\nu) \mathrm{e}^{-2\pi \mathrm{i}\nu t} \mathrm{d}\nu , \ E^{(-)}(\mathbf{r},t) = \int_0^{+\infty} E^*(\mathbf{r},\nu) \mathrm{e}^{2\pi \mathrm{i}\nu t} \mathrm{d}\nu , \qquad (6.1.2)$$

where $E^{(+)}$ and $E^{(-)}$ are called s_{-}, s_{-}, v_{-} and $\ldots, a_{-}, v_{-}, q_{-}, y_{-}, a_{-}, \ldots, l_{-}, \ldots, l_{-}, s_{-}, \ldots$. The former is known as $a_{-}aly_{-}, s_{-}, al$ [91Bor]. In the quantum-mechanical description E, and consequently also $E^{(+)}$ and $E^{(-)}$, become operators, where $E^{(+)}$ contains only photon annihilation operators and $E^{(-)}$ only photon creation operators.

6.1.2.6 Interference

Optical interference results from a superposition of the electric field strengths of different beams. The production of an interference pattern in conventional interference experiments (diffraction, Young's double slit experiment or any kind of interferometer) requires a fixed phase difference between the beams involved to be maintained during the whole observation time.

6.1.2.7 Decoherence

Decoherence is destruction, caused by interaction with the environment, of phase relations that are present in quantum-mechanical superposition states. In particular, due to decoherence quantummechanical superpositions of states corresponding to macroscopically distinguishable properties evolve, at an extremely short time scale, into mixtures, which substantiates the classical reality concept. Since decoherence destroys entanglement (see Sect. 6.1.4.5) as well, it is detrimental in any process that might be used in quantum information transmission and processing, especially quantum computing.

6.1.3 Coherence theory

6.1.3.1 Field correlation functions

6.1.3.1.1 Definitions

In its general form, the *n*-th-order correlation function of an electromagnetic field is defined as

$$G^{(n)}(\boldsymbol{r}_1, t_1; \dots; \boldsymbol{r}_{2n}, t_{2n}) = \langle E^{(-)}(\boldsymbol{r}_1, t_1) \dots E^{(-)}(\boldsymbol{r}_n, t_n) E^{(+)}(\boldsymbol{r}_{n+1}, t_{n+1}) \dots E^{(+)}(\boldsymbol{r}_{2n}, t_{2n}) \rangle$$

$$(n = 1, 2, \dots), \qquad (6.1.3)$$

where the bracket indicates an ensemble average in classical theory and the expectation value in quantum theory. In the quantum-mechanical description, $E^{(-)}$ and $E^{(+)}$ are noncommuting operators. In (6.1.3) they are written in the so-called normal order, which excludes vacuum contributions. The label *n* indicates the order of the correlation function (with respect to intensity). We follow here the notation that is customary in quantum optics. In classical optics, (6.1.3) is referred to as correlation function of 2*n*-th order (with respect to field strength). Since in conventional interference experiments (apart from polarization interferometers) the polarization is irrelevant, the electric field strength is treated as a scalar. Alternatively, (6.1.3) applies to a linearly polarized field.

Of practical relevance are only the first-order correlation function

$$G^{(1)}(\mathbf{r}_1, \mathbf{r}_2; \tau) = \langle E^{(-)}(\mathbf{r}_1, t) E^{(+)}(\mathbf{r}_2, t+\tau) \rangle$$
(6.1.4)

(the special value $G^{(1)}(\mathbf{r}, \mathbf{r}; 0)$ being the intensity at position \mathbf{r}) and the special second-order correlation function

$$G^{(2)}(\mathbf{r}_1, \mathbf{r}_2; \tau) = \langle E^{(-)}(\mathbf{r}_1, t) E^{(-)}(\mathbf{r}_2, t+\tau) E^{(+)}(\mathbf{r}_2, t+\tau) E^{(+)}(\mathbf{r}_1, t) \rangle$$
(6.1.5)

(the correlation between the intensities at space-time points \mathbf{r}_1 , t and \mathbf{r}_2 , $t + \tau$). In both cases the field is assumed to be stationary. In these circumstances, making use of the ergodic hypothesis, the classical ensemble averages can be replaced by time averages. The time difference τ is referred to as the delay time. The modulus of (6.1.4) can be inferred from the intensity distribution in a conventional interference experiment (see Sect. 6.1.3.2), whereas (6.1.5) can be measured with two separate detectors (see Sect. 6.1.4.1 and Sect. 6.1.4.2).

6.1.3.1.2 Connections between correlation functions of different order

The stochastic nature of a chaotic field has the consequence that the full statistical information on it is already contained in its first-order correlation function. In fact, any higher-order correlation

in chaotic fields can be expressed in terms of first-order correlation functions. In particular, the second-order correlation function (6.1.5) for chaotic light takes the form

$$G^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2; \tau) = G^{(1)}(\boldsymbol{r}_1, \boldsymbol{r}_1; 0) G^{(1)}(\boldsymbol{r}_2, \boldsymbol{r}_2; 0) + |G^{(1)}(\boldsymbol{r}_1, \boldsymbol{r}_2; \tau)|^2 \quad .$$
(6.1.6)

Equation (6.1.6) is the basis of stellar intensity interferometry (see Sect. 6.1.4.3).

6.1.3.2 First-order coherence

6.1.3.2.1 Degree of coherence

The quantitative description of interference is based on the first-order correlation function (6.1.4) which, apart from a factor, is just the interference term in the expression for the intensity at an observation point P. In a typical conventional interference experiment a primary beam is divided, by either beam splitting (with the help of a partly, normally semi-, reflecting mirror) or wave front splitting (provided by a screen with pinholes), into two partial beams which, after having propagated along different paths, are re-united on a detector surface. Tracing the two field strengths being superposed at P at a given instant back to the primary beam, one arrives, in general, at two space-time points \mathbf{r}_1 , t and \mathbf{r}_2 , $t + \tau$. This explains why the correlations between those field strengths determine the visibility of the interference pattern (see (6.1.8)).

The normalized correlation function

$$g^{(1)}(\boldsymbol{r}_1, \boldsymbol{r}_2; \tau) = \frac{G^{(1)}(\boldsymbol{r}_1, \boldsymbol{r}_2; \tau)}{\sqrt{G^{(1)}(\boldsymbol{r}_1, \boldsymbol{r}_1; 0)}\sqrt{G^{(1)}(\boldsymbol{r}_2, \boldsymbol{r}_2; 0)}}$$
(6.1.7)

is referred to as the p_1 , l_1 , and its modulus $|g^{(1)}|$ is called the p_1 and r_2 , with delay τ .

The degree of coherence determines the visibility v of the interference pattern through the relation [91Bor]

$$v = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} = \frac{2\mu}{\mu^2 + 1} \mid g^{(1)}(\boldsymbol{r}_1, \boldsymbol{r}_2; \tau) \mid ,$$
(6.1.8)

where μ^2 is the ratio of the intensities of the interfering beams at the observation point P, and $I_{\text{max}}, I_{\text{min}}$ are the maximum and minimum intensities in the interference pattern.

We speak of $(\ ll)$, when $|g^{(1)}|$ reaches its maximum value 1 (in that case the visibility takes its absolute maximum), of $a \ al$, when $0 < |g^{(1)}| < 1$, and of \dots , when $g^{(1)} = 0$.

6.1.3.2.2 Temporal coherence

Specializing to $\mathbf{r}_1 = \mathbf{r}_2$, the degree of coherence $|g^{(1)}(\mathbf{r}_1, \mathbf{r}_1; \tau)|$, as a function of τ , has its maximum (unity) at $\tau = 0$ and falls off for increasing $|\tau|$. Its effective width is called

 $\Delta t_{\rm c}$. Usually, the root mean square (r.m.s.) width is taken as effective width [65Man, 95Man]. Since the Fourier transform of (6.1.4), at $\mathbf{r}_1 = \mathbf{r}_2$, with respect to τ gives us the power spectrum (spectral density) of the field ($W_{-} - K_{-} - K_{-} - K_{-}$), $\Delta t_{\rm c}$ and the line width (r.m.s. width) $\Delta \nu$ satisfy the relation

$$\Delta \nu \cdot \Delta t_c \gtrsim 1. \tag{6.1.9}$$

It follows: In a light field with line width $\Delta \nu$ interferences between partial beams having path differences $c\Delta t$ can be observed only when

$$\Delta t < \Delta t_{\rm c} \quad \text{or} \quad \Delta \nu \Delta t < 1 \,.$$

$$(6.1.10)$$

The relations (6.1.10) are called $\ldots \ldots al \ldots \ldots \ldots \ldots s$.

The distance over which the light propagates during the coherence time is called $(l \dots al)$ $l \dots \Delta l_c$,

$$\Delta l_{\rm c} = c \Delta t_{\rm c} \approx c / \Delta \nu = \lambda_0^2 / \Delta \lambda , \qquad (6.1.11)$$

where λ_0 is the wavelength at the line center and $\Delta \lambda$ is the bandwidth expressed as a wavelength interval.

6.1.3.2.3 Spatial coherence

A rough estimate of $\Delta l_{c,trans}$ is given by the formula [65Man, 95Man]

$$\Delta l_{\rm c,trans} \lesssim \lambda_0 R / \Delta x$$
, (6.1.12)

where Δx is the linear dimension (side of a square or diameter of a circle) of the quasimonochromatic chaotic source and R is the distance of the observation plane from the source. Equivalently, light from such a source with surface ΔA is said to be (spatially) coherent within a cone, with its center in the midst of the emitting surface and its axis approximately orthogonal to it, whose angular aperture $\Delta \gamma_c$ and solid angle $\Delta \Omega_c$ satisfy the equations

$$\Delta x \sin \Delta \gamma_{\rm c} \lesssim \lambda_0 \quad \text{and} \quad \Delta A \cdot \Delta \Omega_{\rm c} \lesssim \lambda_0^2 ,$$

$$(6.1.13)$$

respectively. The angle $\Delta \gamma_c$, or the corresponding solid angle $\Delta \Omega_c$, is a natural measure of spatial coherence since, unlike $\Delta l_{c,\text{trans}}$, it is independent of R. It follows: Interferences within a light field emitted from a surface with diameter Δx can be observed only within a beam with the angular aperture $\Delta \gamma$, where

$$\Delta \gamma < \Delta \gamma_{\rm c} \quad \text{or} \quad \Delta x \, \sin \Delta \gamma < \lambda_0 \,. \tag{6.1.14}$$

The relations (6.1.14) are called $s_i a_i al_{a_i} a_j a_{a_i} a_{$

Equation (6.1.12) can be understood from the following argument: In a Young-type interference experiment, light from an extended chaotic source, after passing two pinholes in a screen, produces an interference pattern on a distant observation screen. Since different source points emit independently, they generate an individual interference pattern each. Those patterns are displaced with respect to each other. In order that they, nevertheless, add up to a visible interference pattern, the maximum displacement occurring must not reach the order of the fringe spacing. This requirement yields (6.1.12).

An accurate determination of the transverse coherence length is based on the $va = C_{\dots,\dots}$ -Z.... [34Cit, 38Zer] which relates the correlation function $G^{(1)}(\mathbf{r}_1, \mathbf{r}_2; 0)$ to the intensity distribution $I(\mathbf{r}')$ on the surface assumed to be a portion of a plane, of the spatially incoherent, quasimonochromatic light source in the form [65Man, 95Man]

$$G^{(1)}(\boldsymbol{r}_1, \boldsymbol{r}_2; 0) = \left(\frac{k_0}{2\pi}\right)^2 \int_{\sigma} I(\boldsymbol{r}') \frac{\mathrm{e}^{\mathrm{i}k_0(R_2 - R_1)}}{R_1 R_2} \mathrm{d}^2 \boldsymbol{r}' \,.$$
(6.1.15)

Here, R_i is the distance from the point \mathbf{r}' on the surface σ to the observation point \mathbf{r}_i (i = 1, 2), and k_0 is the wave number with respect to the line center. In the far-zone of the field, the integral

(6.1.15) can be solved analytically for the case of a uniformly emitting circular source of diameter Δx to yield [65Man, 91Bor, 95Man]

$$|g^{(1)}(\boldsymbol{r}_1, \boldsymbol{r}_2; 0)| = 2 |J_1(u)/u|, \ u = \frac{1}{2}k_0\Delta x |\boldsymbol{r}_2 - \boldsymbol{r}_1|/R.$$
(6.1.16)

Here, J_1 is the Bessel function of the first kind and the first order, and R is the distance between the surface σ and the observation plane which is assumed to be parallel to σ . Since the function $2J_1(u)/u$ drops from its maximum value 1 at u = 0 to 0.88 at u = 1, one can say that a circular area of diameter

$$d = 0.32 \lambda_0 R / \Delta x \tag{6.1.17}$$

is illuminated "almost coherently" [91Bor]. Equation (6.1.17) quantifies the estimate (6.1.12) for the special case of a circular source.

A circle with the transverse coherence length $\Delta l_{c,trans}$ as diameter is called *a a* [91Bor]. Actually, this concept traces back to M. von Laue [14Lau], who speaks of "elementary areas illuminated independently of each other" and identifies their number in a strictly monochromatic, polarized light beam with the number of the degrees of freedom of that beam. The segment of a cone whose base is the coherence area and whose height is the longitudinal coherence length Δl_c is referred to as the v_l [95Man]. In a classical picture, the instantaneous values of both the amplitude and the phase are, to a good approximation, constant over a coherence volume. This justifies to look upon the field within a coherence volume as a single-mode field. Due to the dependence of the coherence area on the distance between source and observation plane the coherence volume (at any observation plane) is also dependent on this distance. The mean photon number \bar{n} per coherence volume is called a y a a \ldots or \ldots a) of the corresponding mode. For radiation from black thermal sources (black-body (or aradiation) \bar{n} follows from Planck's radiation law to be given by

$$\bar{n} = \frac{1}{\frac{h\nu}{e^{kT} - 1}} \tag{6.1.18}$$

(T: temperature, h: Planck's constant, k: Boltzmann's constant) and hence depends strongly on wavelength and temperature, as is illustrated in Fig. 6.1.1.



Fig. 6.1.1. Mean photon number \bar{n} per mode (degeneracy parameter) for black-body radiation in dependence on wavelength and temperature. In the domain of visible radiation, for example, the figure shows the following: In order to achieve $\bar{n} = 10$, a black thermal source must be heated up to temperatures of some 100 000 K. At temperatures of the order of 1000 K typical of technical thermal sources, \bar{n} attains only the order of 10^{-15} , even at the temperature of the sun's surface (T = 5770 K) \bar{n} is still lower than 10^{-2} .

6.1.3.2.4 Filtering out coherent light from a chaotic source

The phrase "making (chaotic) light coherent" is used for making a coherently illuminated area or/and the coherence time in the field of a chaotic source sufficiently large, thus making the light more suitable for interference to be observed. In the case of spatial coherence this is achieved by reducing the effective size of the emitting surface. To this end, a secondary light source of small and perhaps adjustable dimensions is generated via imaging, with the help of a lens or a mirror, the primary source on a narrow-aperture diaphragm. In the case of temporal coherence it can be done by reducing the bandwidth with the help of a spectral filter.

6.1.3.2.5 Measurement of coherence lengths

The longitudinal coherence length Δl_c can be determined utilizing an interferometer, e.g. of Michelson's type. The path difference is gradually enlarged until the visibility vanishes. This critical path difference equals Δl_c . Similarly, the transverse coherence length $\Delta l_{c,\text{trans}}$ can be measured in a Young's type interference experiment by increasing the distance of the pinholes in the screen or, equivalently, bringing the screen nearer to the source, until the interference pattern becomes invisible.

6.1.3.3 Laser light versus chaotic light

6.1.3.3.1 Generating mechanisms and radiation characteristics

6.1.3.3.1.1 Chaotic light

Chaotic light is generated by uncorrelated (spontaneous) emission from individual microscopic emitters, usually atoms or molecules. In a classical picture, the individual wave trains have random phases. Hence their superposition at a given position and a given time will randomly be constructive or destructive. As a result, the intensity exhibits strong fluctuations which give rise to characteristic intensity correlations (see Sects. 6.1.4.3 and 6.1.4.4). Under laboratory conditions the coherence volume of extended chaotic sources is rather small. For black-body radiation and for thermal light the degeneracy parameter \bar{n} in the visible spectral range is by orders of magnitude smaller than unity (see Fig. 6.1.1).

6.1.3.3.1.2 Laser light

Laser light is generated by stimulated emission in a medium strongly pumped to produce population inversion, and placed within an optical resonator. Under these conditions, phase relations between the individual emitters are established so that the emitted waves interfere with each other. Only within a small aperture angle this interference is constructive and results in the emission of the coherent beam; in all other directions the interference turns out to be destructive, so that the entire emission is restricted to the coherent beam. The resonator gives rise to a very narrow bandwidth which can be reduced further, by orders of magnitude, by means of laser-frequency stabilization (locking the laser frequency to a narrow absorption line or a resonance line of a highly stabilized external resonator). Hence the coherence volume in cw laser fields is very large already at distances of a few meters; in pulsed lasers it coincides with the volume filled by a single pulse. The spectral density of the radiation power is extremely high, and the same holds for the degeneracy parameter. Qualitatively, laser light differs from chaotic light by its amplitude stabilization (see Sect. 6.1.4.4) resulting from a saturation mechanism in the generating process. Laser light, as well as the radiation from radar and radio transmitters, comes very close to the classical ideal of a monochromatic wave of fixed amplitude and phase.

6.1.3.3.2 Interference between beams from independent sources

Due to the small degeneracy parameter, any interference experiment performed with conventional light requires the observation time, or simply the exposition time of a photographic plate, to be much larger than the coherence time. The very large degeneracy parameter of laser light, on the contrary, makes it possible to detect interference even between beams from independently operated lasers (both cw and pulsed lasers), since the observation time can be chosen to be comparable with, or shorter than, the coherence time. For the first demonstration of this phenomenon [63Mag] pulsed lasers were used. Since the phases of pulses from different lasers are uncorrelated, the position of the interference pattern is random. This is in a marked contrast to conventional interference experiments, where the interfering beams are taken from the same source, so that the relative phases between them are determined solely by the geometry of the set-up.

Interference becomes possible even between strongly damped cw laser beams. This prediction relies on two basic theoretical results:

- 1. In the quantum-mechanical description, optimum interference (with visibility v = 1 for equal intensities) takes place between independent beams if, and only if, they are in a (quantum-mechanical) coherent state [63Gla], also named Glauber state, each [63Pau].
- 2. Absorption and reflection transform a Glauber state again into a Gauber state [64Bru, 65Bru]. In the actual interference experiment Radloff [71Rad], following a suggestion by Paul et al. [65Pau], utilized the intense laser beams off which the interfering beams were split by means of very weakly reflecting beam splitters, to control, via a shutter, the relative phase between the interfering beams.

6.1.3.3.3 Coherent interaction

Unlike chaotic light, laser light can be used to observe coherent interaction between light and matter. It is based on the fact that resonant intense coherent light induces on an atom or a molecule dipole oscillations whose phase is determined by the field phase. Hence in an atomic or molecular ensemble those individual dipole oscillations add up to a macroscopic (oscillating) polarization of the medium which, in turn, acts as a source of radiation. Examples of those processes are free induction decay, optical nutation, photon echo and self-induced transparency, as well as Coherent Stokes Raman Scattering (CSRS) and Coherent Antistokes Raman Scattering (CARS) (cf. [84She]).

6.1.3.4 Particle interference

Optical interference can be described equally well in the photon picture that dates back to Einstein [1905Ein]. Photons are particles with energy $E = h\nu$ (h: Planck's constant) and momentum $\mathbf{p} = (h/2\pi)\mathbf{k}$ (k: wave vector). Equation (6.1.9) is tantamount to the quantum-mechanical timeenergy uncertainty relation¹ $\Delta E \cdot \Delta t \ge h/4\pi$. Equation (6.1.13) can be based on Heisenberg's uncertainty relation for position and momentum of a particle, $\Delta x \cdot \Delta p \ge h/4\pi$. Applied to a photon, the latter states that a photon whose position is confined to the finite surface of the source with dimension Δx has a minimum momentum uncertainty with respect to the x direction, which defines a minimum angular aperture to be identified with $\Delta\gamma_c$ in (6.1.13).

Also massive particles possess wave properties (wave–particle dualism) resulting in interference phenomena. The analog of the optical wavelength is the de Broglie wavelength $\Lambda = h/mv$ (m: mass and v: velocity of the particle). It follows from this formula that interference can be observed only on particles with a small mass such as electrons, neutrons, atoms, molecules or clusters. While electron interference is known since 1927, atom optics has been established as a new discipline only recently. In fact, atom-optical analogs of optical elements, in particular beam splitters, are nowadays available.

Particle interference must always be interpreted as "interference of the particle with itself", as Dirac put it [58Dir]. In order to detect an interference pattern, the observation has to be carried out on an ensemble being in a pure quantum state with respect to the motional degrees of freedom, i.e. being described by a well-defined wave function. In the quantum-mechanical description, interference takes place, on a detector surface, between the probability amplitudes that correspond to different paths the particle is allowed to take. There exists a complementarity between the visibility of the interference pattern and the information on the particle's path. In particular, maximum visibility requires the path to be completely undetermined. In this context atoms or molecules are especially attractive because of their internal degrees of freedom that can be utilized to mark the particle's path, e.g. by excitation of an electronic level followed by spontaneous emission or via coherent excitation of hyperfine levels [98Due], which makes the interference pattern disappear.

6.1.3.5 Higher-order coherence

While first-order coherence underlies the conventional interference experiments resulting in the appearance of a visible interference pattern, Glauber [65Gla] generalized the coherence concept with regard to photon coincidence measurements to be carried out with two or more detectors. He termed a field M, M, M, when the normalized correlation functions

$$g^{(n)}(\mathbf{r}_{1}, t_{1}; \dots; \mathbf{r}_{n}, t_{n}; \mathbf{r}_{n+1}, t_{n+1}; \dots; \mathbf{r}_{2n}, t_{2n}) = \frac{G^{(n)}(\mathbf{r}_{1}, t_{1}; \dots; \mathbf{r}_{n}, t_{n}; \mathbf{r}_{n+1}, t_{n+1}; \dots; \mathbf{r}_{2n}, t_{2n})}{\sqrt{G^{(1)}(\mathbf{r}_{1}, t_{1}; \mathbf{r}_{1}, t_{1})} \dots \sqrt{G^{(1)}(\mathbf{r}_{2n}, t_{2n}; \mathbf{r}_{2n}, t_{2n})}}$$
(6.1.19)

have modulus 1 for, and only for, all $n \leq M$. Evidently, this condition implies the relations

$$G^{(n)}(\boldsymbol{r}_1, t_1; \dots; \boldsymbol{r}_n, t_n; \boldsymbol{r}_n, t_n; \dots; \boldsymbol{r}_1, t_1) = G^{(1)}(\boldsymbol{r}_1, t_1; \boldsymbol{r}_1, t_1) \dots G^{(1)}(\boldsymbol{r}_n, t_n; \boldsymbol{r}_n, t_n)$$
(6.1.20)

for all $n \leq M$. Physically, this means that the *n*-fold coincidence counting rate factorizes into the single counting rates which would be measured by each detector individually in the absence of all others. In other words, the detectors respond to the field in a statistically independent way. Hence

¹ Though the time–energy uncertainty relation is less fundamental than Heisenberg's uncertainty relation, since there exists no time operator, it is nevertheless very useful in spectroscopy [54Hei]. Then ΔE is interpreted as the accuracy of an energy measurement and Δt as the measurement time.

Glauber's higher-order coherence is characterized by the $a \ s \dots$ of $sys \ a \dots$ coincidences, when measurements on the radiation field are carried out with several independent detectors.

For chaotic light, (6.1.20) is not fulfilled already for n = 2 (see (6.1.6)). Hence this kind of light has only first-order coherence. Ideal laser light, on the other hand, is coherent in ∞ -th order.

The coherence condition (6.1.20) aims at fields that are laser-like only up to a certain order, and, hence, in a sense interpolate between chaotic and laser light. Since, however, no such fields are known to date, the concept of higher-order coherence has hitherto found only little practical use.

6.1.4 Intensity interference

6.1.4.1 Formal description

Intensity interference is based on measurement of intensity correlations. The latter are described by the special second-order correlation function $G^{(2)}(\mathbf{r}_1, \mathbf{r}_2; \tau)$ (see (6.1.5)). Apart from a factor, this function has the physical meaning of the (conditional) probability that a photodetector at position \mathbf{r}_2 will click at time $t + \tau$, on condition that a photodetector at \mathbf{r}_1 has clicked at t [65Gla]. Accordingly, the delayed coincidence counting rate is proportional to $G^{(2)}(\mathbf{r}_1, \mathbf{r}_2; \tau)$.

6.1.4.2 Measurement

The intensity correlation function for zero delay time can be measured by correlating electronically the photocurrents from two detectors placed at different positions. To this end, the photocurrents are multiplied and, after narrow-band amplification, averaged over time. A second technique is to use two separate photocounters and single out electronically those events where a click from one detector is followed, within a narrow time window, by a click from the other. To account for nonzero delay time, a delay is introduced between the photocurrents in the first case. In the second case use is often made of time-to-digital converters.

6.1.4.3 Spatial intensity correlations

6.1.4.3.1 Stellar intensity interferometry

Equation (6.1.6), for $\tau = 0$, means for astronomy that measuring intensity correlations yields precisely the same information (the modulus of the first-order correlation function) as Michelson's stellar interferometer. The pioneering work was devised and carried out by Brown and Twiss [56Bro, 64Bro], who observed a drop of the intensity correlations to a constant level with increasing distance $|\mathbf{r}_2 - \mathbf{r}_1|$ of the reflectors focusing light from the star on individual photomultipliers. This technique has two important advantages:

- 1. One gets rid of the disturbing influence of atmospheric scintillations, and
- 2. no mechanical connection of the reflectors is needed any longer, which allows to drastically enhance the detector distance.

6.1.4.4 Temporal intensity correlations

Delayed coincidences at a given position $(r_1 = r_2)$ are measured with the following set-up [57Reb, 57Twi]: The incoming light is divided, with the help of a semitransparent beam splitter, into two beams each of which is directed to a separate photodetector. With the detectors being in equivalent positions, one solves in this way the problem of placing two detectors at the same position.

6.1.4.4.1 Amplitude stabilization

In contrast to chaotic radiation, light from a single-mode laser operated well above threshold is amplitude-stabilized. This means that the intensity correlation function, to a good approximation, satisfies the relation

$$G^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_1; \tau) = \left[G^{(1)}(\boldsymbol{r}_1, \boldsymbol{r}_1; \tau) \right]^2 \quad \text{for all } \tau .$$
(6.1.21)

6.1.4.4.2 Photon bunching

It follows from (6.1.6), for $\mathbf{r}_1 = \mathbf{r}_2$, that the delayed coincidence counting rate for chaotic light is distinctly enhanced at $\tau = 0$ compared to its values at delay times that are comparable to, or larger than, the coherence time (for an experimental demonstration see e.g. [69Foo]). This indicates from the viewpoint of classical optics the presence of strong intensity (or, equivalently, amplitude) fluctuations, and from the viewpoint of quantum optics a tendency for photons to arrive in bunches. Accordingly, this effect has been termed

Similar to the spatial case, the modulus of the first-order correlation function $G^{(1)}(\mathbf{r}_1, \mathbf{r}_1; \tau)$ can be found from the measurement of temporal intensity correlations. In particular, the coherence time can be inferred from the drop of the intensity correlation to a constant level with increasing delay time τ . In view of (6.1.9), this opens a way to indirectly measure very narrow bandwidths. In fact, the scheme, unlike conventional spectroscopy, works the better, the narrower the bandwidth.

Unfortunately, the coherence time attainable with thermal sources is normally still short in comparison to the response time of photodetectors, which practically prohibits the observation of photon bunching. However, laser light scattered from a moving ground glass [64Mar], from moving centers suspended in a fluid, or from inhomogeneities due to temperature fluctuations in a fluid, has stochastic character and, moreover, possesses the desired small bandwidth. Diffusion coefficients [69Foo] and the thermal conductivity were determined in this way.

6.1.4.4.3 Photon antibunching

Whereas in classical statistics any autocorrelation function has an absolute maximum at zero delay time, quantum states of the electromagnetic field were specified whose delayed coincidence counting rate exhibits a dip at $\tau = 0$. This intrinsically quantum-mechanical effect is known as a_{1}, \ldots, a_{n} , and light showing it falls therefore in the category of a_{2} and b_{2} .

The antibunching effect is observed in resonance fluorescence on single atoms. The underlying physical mechanism is as follows: An atomic system driven by a strong resonant field (laser field) becomes excited at a time and hence emits a photon. However, it cannot emit a second photon immediately afterwards, since the laser field needs some time to excite it again. So one will never register coincidences at zero delay time, in contrast to the situation at larger delay times. One of the main experimental problems is to ensure that actually one atom, at most, is under observation at a time. In their pioneering work, Kimble et al. [77Kim] worked with a very thin beam of

sodium atoms, of which only a small section was observed. In these circumstances, it could not be excluded, however, that sometimes the observation volume was occupied by more than one atom. Nevertheless, a distinct dip in the intensity correlation function at $\tau = 0$ was observed. In later experiments, this handicap was overcome by studying resonance fluorescence on single ions stored in electromagnetic traps, or single molecules doped as impurities in solids or placed in a microcavity [98Kit].

6.1.4.5 Experiments with entangled photon pairs

6.1.4.5.1 Parametric down-conversion

One of the most puzzling quantum-mechanical features is entanglement (see also Sect. 5.1.9.3). Having no classical analog, it lies at the heart of the Einstein Podolsky Rosen (EPR) paradox [35Ein] and underlies any experiment designed to demonstrate a violation of Bell's inequalities (see also Sect. 5.1.9.4). By definition, a (pure) entangled state of two or more physical systems is described by an unfactorable wave function. Of special relevance are entangled states of two particles (labeled 1 and 2) for which a global variable $v = v_1 + v_2$, e.g. energy or spin, has a sharp value, whereas v_1 and v_2 are uncertain. Hence measuring v_1 fixes also v_2 , even when particle 2 has traveled far away.

A very effective, widely used technique to generate entanglement is spontaneous parametric down-conversion. In this process, an intense laser beam is sent, as a pump, in a nonlinear crystal, which leads to a sporadic "decay" of a pump photon into a photon pair named signal and idler photon. Energy conservation requires the sum of the signal and the idler frequency to coincide with the pump frequency, and the phase matching condition implies a similar requirement for the wave vectors resulting in different propagation directions for the two photons. Moreover, whereas the time at which such an event occurs is impredictable over a comparatively long period, the possible delay between the emission of the signal and the idler photon is extremely short.

As a result of the frequency condition mentioned, a new coherence length determined by the large coherence length of the pump and hence much longer than the one-photon coherence lengths for signal and idler, becomes manifest in suitable experiments, which has to be assigned to the two-photon wave packet (sometimes called ______) as a whole.

6.1.4.5.2 Two-photon mixing

6.1.4.5.2.1 Hong–Ou–Mandel interferometer

When two photons are mixed at a 50%: 50% beam splitter, they leave it both in one or the other of the two output channels, provided they have arrived simultaneously at the beam splitter [87Fea, 87Hon]. The pioneering experiment [87Hon] was performed on photon pairs produced in spontaneous parametric down-conversion. The coincidence counting rate of detectors placed in the two output channels was observed to exhibit a distinct dip reaching almost zero, for zero delay time. Actually, this destructive interference is not bound to entanglement [87Fea]. Experimentally, this was shown by Rarity et al. [96Rar], who mixed a photon produced in spontaneous parametric down-conversion and a highly attenuated pulse from the pump laser.

The Hong–Ou–Mandel experiment [87Hon] was recently modified [98Str] by inserting in the signal channel a rod of birefringent material followed by a linear polarizer. This set-up allows the signal photon to propagate along two ways that differ in their optical path lengths by $2\Delta L$. The latter quantity being chosen much greater than the coherence length for both the signal and the idler photon, the signal photon was divided into two well-separated wave packets. Positioning the

mirror such that the idler photon arrived precisely midway between the two signal wave packets, and afterwards varying ΔL with the help of a Pockels cell, the authors observed a sinusoidal change in the coincidence counting rate registered with a time window much greater than the delay between the two signal wave packets, and hence a transition from the dip in the Hong–Ou–Mandel experiment to a peak. Since signal and idler wave packets do not overlap at the beam splitter, this effect is clearly of non-classical origin.

6.1.4.5.3 Photon-pair interference

6.1.4.5.3.1 Interferometric devices

Photon pairs generated in spontaneous parametric down-conversion are sent into one entrance channel of a Michelson interferometer, and the transmitted light is detected by a "photon-pair detector" consisting of a semitransparent beam splitter with a separate detector in each output channel [89Moh, 90Kwi, 91Bre]. The coincidence counting rate observed with a very narrow time window, as a function of the path difference of the interferometer arms, displays interference fringes corresponding to a de Broglie wavelength that is given by the wavelength of the pump [90Kwi, 91Bre]. The corresponding second-order coherence time is much longer than the first-order coherence time of the signal and the idler photon. The described interference phenomenon results from the superposition of the probability amplitudes for the following two cases:

1. Both photons went through the short arm of the interferometer, and

2. both photons went through the long arm.

Those two cases cannot be distinguished due to the uncertainty of the emission time of the photon pair. Hence, what is observed is interference of the photon pair (biphoton) with itself. The fringe visibility of 87% measured in [91Bre] distinctly exceeded the critical value of 50% compatible with a classical model.

A similar investigation was recently carried out with a Young double-slit set-up [99Fon]. Further, it should be mentioned that Jacobson et al. [95Jac] proposed a modified Mach–Zehnder interferometer that is capable of measuring the de Broglie wavelength of a multiphoton wave packet.

6.1.4.5.3.2 Franson experiment

Following a proposal by Franson [89Fra], Kwiat et al. [93Kwi] performed the following experiment: The two photons of a pair generated in spontaneous parametric down-conversion are coupled each into a separate Mach–Zehnder interferometer. Exiting the interferometer, the photons fall on separate detectors whose coincidence counts are registered. The experimental parameters are chosen such that only two types of events are observed:

- 1. Both photons take the long way in the respective interferometer, and
- 2. both take the short way.

Since the emission time of the photon pair is uncertain, the detectors cannot distinguish those two cases, which gives rise to the occurrence of interference fringes in the coincidence counting rate, as a function of an additional path difference introduced in one of the interferometers by translating the right-angle prism. The measured fringe visibility of about 80% indicated a strong violation of Bell's inequality.

6.1.5 Photon counting statistics

6.1.5.1 Measurement

The field is made to impinge on a single photodetector. Confining the illumination, with the help of a pinhole placed in front of the detector, to a coherence area ΔA_c and registering the photoelectrons (clicks) over a time interval that equals the coherence time Δt_c , one counts the photons contained in a coherence volume.

6.1.5.2 Photon distribution functions

For chaotic light the probability p_n to detect just n photons is given by the distribution

$$p_n = \frac{\overline{n}^n}{(\overline{n}+1)^{n+1}} \quad (\ s = \dots \ s_{n-1} \ s_{n-1} \ s_{n-1} \), \tag{6.1.22}$$

where \overline{n} is the mean photon number per coherence volume. For (ideal) laser light the corresponding probability is given by the distribution

$$p_n = \frac{\exp(-\overline{n})\,\overline{n}^n}{n!} \quad (\dots, ss_1, \dots, s_{n-1}, \dots) \,. \tag{6.1.23}$$

Both distributions are shown in Fig. 6.1.2 for two values of \bar{n} .



Fig. 6.1.2. Bose–Einstein distribution (1) and Poisson distribution (2) for a mean photon number of $\bar{n} = 1.5$ and $\bar{n} = 10$. The figure indicates that already for moderate mean photon numbers the two distributions differ distinctly in both their shape and the positions of their maxima, whereas for $\bar{n} \leq 1$ they are hard to discriminate experimentally.
6.1.5.3 Measurements under different experimental conditions

Measurements of this kind are also performed under different experimental conditions. When the illuminated area $\Delta A_{\rm ill}$ is chosen smaller than $\Delta A_{\rm c}$ or/and the observation time $\Delta t_{\rm obs}$ shorter than Δt_c , the distributions (6.1.22) and (6.1.23) are still valid. (Only the mean photon numbers, referring to the actual observation volume, change.) For growing $\Delta A_{\rm ill}$ (> $\Delta A_{\rm c}$) or/and $\Delta t_{\rm obs}$ (> Δt_c) the photon distribution for chaotic light tends to a Poisson distribution, while the distribution for laser light remains Poissonian.

The Bose–Einstein distribution was first measured on laser light scattered from a suspension of polysterene globules [67Are]. Probability densities for the intensity of pseudothermal light [65Man] were measured already before [66Mar1, 66Mar2].

6.1.5.4 Variances of the photon number

The variances of the photon number follow from (6.1.22) and (6.1.23) to be

$$\Delta n^2 = n^2 - \overline{n}^2 = \overline{n}^2 + \overline{n} \quad \text{for chaotic light} , \qquad (6.1.24)$$

and

 $\Delta n^2 = \overline{n} \quad \text{for (ideal) laser light} . \tag{6.1.25}$

Evidently, the Bose-Einstein distribution is broader, for equal average photon numbers \overline{n} , than the Poisson distribution. Hence the former falls in the category of $s_1 \dots ss_n a_n s_n s_n s_n$. Antibunched light, on the other hand, normally exhibits a photon distribution that is narrower than the Poisson distribution – accordingly one speaks of $s_1 \dots ss_n a_n s_n s_n s_n$, and the variance of the photon number is smaller than \overline{n} . Equivalently, the so-called $Q_1 a_1 a_1 \dots$ introduced by Mandel [79Man], see also Sect. 5.1.3.6.1, $Q = (\Delta n^2 - \overline{n})/\overline{n}$, being equal to zero for Poissonian statistics and positive for super-Poissonian statistics, becomes negative for sub-Poissonian statistics.

Einstein, A.: Ann. Physik (Leipzig) 17 (1905).

References for 6.1

09Ein Einstein, A.: Phys. Z. 10 (1909) 185. Taylor, G.I.: Proc. Cambridge Philos. Soc. 15 (1909) 114. 09Tay 14Lau Laue, M. von: Ann. Phys. (Leipzig) 44 (1914) 1197. 21Mic Michelson, A.A., Pease, F.G.: Astrophys. J. 53 (1921) 249. 34Cit Cittert, P.H. van: Physica 1 (1934) 201. 35Ein Einstein, A., Podolsky, B., Rosen, N.: Phys. Rev. 47 (1935) 777. 38Zer Zernike, F.: Physica 5 (1938) 785. 54Hei Heitler, W.: The quantum theory of radiation, 3rd ed., London: Oxford University Press, 1954. 56Bro Brown, R.Hanbury, Twiss, R.Q.: Nature (London) 178 (1956) 1046. Rebka, G.A., Pound, R.V.: Nature (London) 180 (1957) 1035. 57Reb Twiss, R.Q., Little, A.G., Brown, R.Hanbury: Nature (London) 180 (1957) 324. 57Twi 58Dir Dirac, P.A.M.: The principles of quantum mechanics, 4th ed., London: Oxford University Press, 1958. 60Mai Maiman, T.H.: Nature (London) 187 (1960) 493. 61Jav Javan, A., Bennett, W.R., Herriott, D.R.: Phys. Rev. Lett. 6 (1961) 106. 62.Jav Javan, A., Ballik, E.A., Bond, W.L.: J. Opt. Soc. Am. 52 (1962) 96. Glauber, R.J.: Phys. Rev. 131 (1963) 2766. 63Gla 63Mag Magyar, G., Mandel, L.: Nature (London) **198** (1963) 255. Paul, H., Brunner, W., Richter, G.: Ann. Phys. (Leipzig) 12 (1963) 325. 63Pau Brown, R.Hanbury: Sky Telesc. 28 (1964) 64. 64Bro Brunner, W., Paul, H., Richter, G.: Ann. Phys. (Leipzig) 14 (1964) 384. 64Bru 64Mar Martienssen, W., Spiller, E.: Am. J. Phys. 32 (1964) 919. 65Bru Brunner, W., Paul, H., Richter, G.: Ann. Phys. (Leipzig) 15 (1965) 17. Glauber, R.J., in: Qantum optics and electronics, DeWitt C., Blandin, A., Cohen-65Gla Tannoudji, C. (eds.), New York: Gordon and Breach, 1965. 65Man Mandel, L., Wolf, E.: Rev. Mod. Phys. 37 (1965) 231. Paul, H., Brunner, W., Richter, G.: Ann. Phys. (Leipzig) 16 (1965) 93. 65Pau 66Mar1 Martienssen, W., Spiller, E.: Phys. Rev. 145 (1966) 285. 66Mar2 Martienssen, W., Spiller, E.: Phys. Rev. Lett. 16 (1966) 531. 67Are Arecchi, F.T., Giglio, M., Tartari, U.: Phys. Rev. 163 (1967) 186.

1905Ein

- 69Foo Foord, R., Jakeman, E., Jones, R., Oliver, C.J., Pike, E.R.: IERE Conf. Proc. 14 (1969).
- 71Rad Radloff, W.: Ann. Phys. (Leipzig) **26** (1971) 178.
- 77Kim Kimble, H.J., Dagenais, J.M., Mandel, L.: Phys. Rev. Lett. **39** (1977) 691.
- 79Man Mandel, L.: Opt. Lett. 4 (1979) 205.
- 84She Shen, Y.R.: The principles of nonlinear optics, New York: Wiley, 1984.
- 87Fea Fearn, H., Loudon, R.: Opt. Commun. **64** (1987) 485.
- 87Hon Hong, C.K., Ou, Z.Y., Mandel, L.: Phys. Rev. Lett. 59 (1987) 2044.
- 89Fra Franson, J.D.: Phys. Rev. Lett. **62** (1989) 2205.
- 89Moh Mohler, E., Brendel, J., Lange, R., Martienssen, W.: Europhys. Lett. 8 (1989) 511.
- 90Kwi Kwiat, P.G., Vareka, W.A., Hong, C.K., Nathel, H., Chiao, R.Y.: Phys. Rev. A **41** (1990) 2910.
- 91Bor Born, M., Wolf, E.: Principles of optics, 6th ed., Oxford: Pergamon Press, 1991.
- 91Bre Brendel, J., Mohler, E., Martienssen, W.: Phys. Rev. Lett. 66 (1991) 1142.
- 93Kwi Kwiat, P.G., Steinberg, A.M., Ciao, R.Y.: Phys. Rev. A 47 (1993) R2472.
- 95Jac Jacobson, J., Björk, G., Chuang, I., Yamamoto, Y.: Phys. Rev. Lett. **74** (1995) 4835.
- 95Man Mandel, L., Wolf, E.: Optical coherence and quantum optics, Cambridge: Cambridge University Press, 1995.
- 96Rar Rarity, J.G., Tapster, P.R., Loudon, R., in: Quantum interferometry, DeMartini, F., Denardo, G., Shih, Y. (eds.), Weinheim: VCH, 1996.
- 98Due Dürr, S., Nonn, T., Rempe, G.: Nature (London) **395** (1998) 33.
- 98Kit Kitson, S.C., Jonsson, P., Rarity, J.G., Tapster, P.R.: Phys. Rev. 58 (1998) 620.
- 98Str Strekalov, D.V., Pittman, T.B., Shih, Y.H.: Phys. Rev. A 57 (1998) 567.
- 99Fon Fonseca, E.J.S., Monken, C.H., Pádua, S.: Phys. Rev. Lett. 82 (1999) 2868.

6.2 Superradiance

M.G. BENEDICT, E.D. TRIFONOV

6.2.1 Definitions and historical layout

SuperRadiance (SR) is the cooperative spontaneous emission of radiation by a collection of atoms, molecules or nuclei. The spontaneous transition from an excited state to the ground state is the result of interaction with the vacuum fluctuations of the electromagnetic field. The theory of spontaneous emission from a single atom was first proposed by Dirac [27Dir] and then developed by Wigner and Weisskopf [30Wei]. It yields the usual exponential decay with the natural radiation lifetime γ^{-1} , see (5.1.66) and (6.2.1). In most cases one observes this type of radiation also from a system of several atoms, because the spontaneous decays of the sources are not correlated. In 1954 Dicke [54Dic] recognized that if the phases of the atomic states do not suffer random changes, then the interaction of the atoms with each other through their common radiation electromagnetic field results in a correlation between the atomic dipole moments. In this way a macroscopic polarization is created which is proportional to the number of atoms N, the multi-atomic system radiates collectively, and the radiation is coherent. The decay time of this collective emission is proportional to 1/N, and the radiation intensity is proportional to the square of the number of atoms N^2 . This is the most characteristic property of SR. Furthermore, the shape of the emitted electromagnetic pulse is no longer exponential, but has a peak (or some ringing) after a certain delay. These are the reasons why SR is one of the efficient methods used to generate intensive and short electromagnetic pulses.

SR is effective only if other interactions, such as collisions, thermal noise etc., which are usually also present besides the interaction with the radiation electromagnetic field, do not disturb the phases of the atomic states during emission. If this condition is not fulfilled, one has ordinary spontaneous emission, or amplification of the latter: *Amplified Spontaneous Emission* (ASE), an incoherent effect which is sometimes also called *superluminescence*. We use here the term *super-radiance* for any cooperative emission process, although the cooperative spontaneous emission of a macroscopic system of atoms being initially in a fully excited but uncorrelated state is termed often by the special name *SuperFluorescence* (SF).

The first experimental observation of SR followed not until nearly two decades after the prediction by Dicke. In 1973 Skribanowitz et al. [73Skr] were able to realize experimental conditions where the phase memory time of the radiation centers was longer than the cooperative spontaneous emission time. Their experiment showed the effect in a pencil-shaped extended sample consisting of HF molecules.

Several reviews and books [74Aga, 75All, 81Sch, 82Gro, 82Vre, 85Har, 86Leo, 93And, 96Ben] treat SR in great detail from different points of view, mainly theoretically. Experimental investigations on extended systems are reviewed in [82Vre]. The work [85Har] gives a detailed description of microwave SR experiments in a cavity, not considered here. In the book [96Ben] one of the chapters is devoted exclusively to the discussion of the experiments performed with superradiant systems, and in several other chapters emphasis is laid on the experimental justification of the different aspects of the effect. It contains also a rather comprehensive bibliography of both experimental and theoretical works in the field.

6.2.2 Superradiance theory

6.2.2.1 Superradiance of a system with dimensions smaller than the radiation wavelength

A two-level atom is a simple and effective model of a real atom interacting resonantly with the electromagnetic field. The upper and lower states, $|e\rangle$ and $|g\rangle$, are supposed to be separated by an energy interval $\hbar\omega_0$. If the atom is initially in its upper state, (and the field in the vacuum state), then the probability P_e of finding the atom in the state $|e\rangle$ obeys the equation [27Dir, 30Wei]:

$$\frac{\mathrm{d}P_{\mathrm{e}}}{\mathrm{d}t} = -\gamma P_{\mathrm{e}} , \quad \text{with} \quad \gamma = \frac{1}{3\pi\varepsilon_0} \frac{|\boldsymbol{d}|^2 \omega_0^3}{\hbar c^3} , \qquad (6.2.1)$$

where $\boldsymbol{d} = \langle \mathbf{e} | \hat{\boldsymbol{D}} | \mathbf{g} \rangle$ is the off-diagonal matrix element of the dipole operator $\hat{\boldsymbol{D}}$, see below. Equation (6.2.1) gives the usual exponential law of spontaneous emission. It is valid for an arbitrary quantum system provided its dimensions are less than the emission wavelength.

In order to treat the problem of several atoms one introduces an effective quantum-mechanical energy operator, $\hat{H}_0^i = \hbar \omega_0 \hat{\mathcal{R}}_3^i$, and the raising and lowering operators $\hat{\mathcal{R}}_+^i$ and $\hat{\mathcal{R}}_-^i$ for the *i*-th atom with:

$$\widehat{\mathcal{R}}_{3}^{i} := \frac{1}{2} (|\mathbf{e}\rangle \langle \mathbf{e}| - |\mathbf{g}\rangle \langle \mathbf{g}|)^{i} , \quad \widehat{\mathcal{R}}_{+}^{i} := (|\mathbf{e}\rangle \langle \mathbf{g}|)^{i} , \quad \widehat{\mathcal{R}}_{-}^{i} := (|\mathbf{g}\rangle \langle \mathbf{e}|)^{i} , \quad (6.2.2)$$

which are seen to obey the commutation relations of angular momentum operators:

$$[\widehat{\mathcal{R}}_3^i, \widehat{\mathcal{R}}_{\pm}^j] = \pm \delta_{ij} \widehat{\mathcal{R}}_{\pm}^i , \quad [\widehat{\mathcal{R}}_{\pm}^i, \widehat{\mathcal{R}}_{\pm}^j] = 2\delta_{ij} \widehat{\mathcal{R}}_3^i .$$
(6.2.3)

The corresponding operators for the total system are:

$$\widehat{\mathcal{R}}_{\alpha} = \sum_{i}^{N} \widehat{\mathcal{R}}_{\alpha}^{i} \quad (\alpha = +, -, 3) .$$
(6.2.4)

The total dipole moment operator can be expressed as: $\widehat{D} = \sum_{i} \widehat{D}^{i} = d(\widehat{\mathcal{R}}_{+} + \widehat{\mathcal{R}}_{-})$. The energy operator of the N identical two-level atoms is $\widehat{H}_{0} = \hbar \omega_{0} \widehat{\mathcal{R}}_{3}$. As it is easily seen, it has N + 1 different eigenvalues:

$$E_M = M\hbar\omega_0$$
 with $M = N/2, N/2 - 1, \dots, -N/2$, (6.2.5)

where $M = (N_+ - N_-)/2$ is half of the difference of the number of atoms in the upper and lower levels, N_+ and N_- . In each subspace determined by a degenerate energy value there is exactly one linear combination that is *symmetric* with respect to the permutation of the atoms:

$$|\mathbf{e}, \mathbf{e}, \mathbf{e}, \dots, \mathbf{e}\rangle \qquad M = N/2 ,$$

$$\frac{1}{\sqrt{N}}(|\mathbf{g}, \mathbf{e}, \dots, \mathbf{e}\rangle + |\mathbf{e}, \mathbf{g}, \dots, \mathbf{e}\rangle + \dots |\mathbf{e}, \mathbf{e}, \dots, \mathbf{g}\rangle) \qquad M = N/2 - 1 ,$$

$$\binom{N}{2}^{-\frac{1}{2}}(|\mathbf{g}, \mathbf{g}, \mathbf{e}, \dots, \mathbf{e}\rangle + |\mathbf{g}, \mathbf{e}, \mathbf{g}, \dots, \mathbf{e}\rangle + \dots |\mathbf{e}, \mathbf{e}, \dots, \mathbf{g}, \mathbf{g}\rangle) \qquad M = N/2 - 2 ,$$

$$\vdots$$

$$|\mathbf{g}, \mathbf{g}, \mathbf{g}, \dots, \mathbf{g}\rangle \qquad M = -N/2 . \qquad (6.2.6)$$

These states all belong to the highest eigenvalue, $\frac{N}{2}\left(\frac{N}{2}+1\right)$, of the operator $\hat{\mathcal{R}}^2 = \left(\frac{1}{2}(\hat{\mathcal{R}}_-\hat{\mathcal{R}}_+ + \hat{\mathcal{R}}_-) + (\hat{\mathcal{R}}_3)^2\right)$. $\hat{\mathcal{R}}^2$ commutes with both the energy operator of the unperturbed system of atoms \hat{H}_0 and the dipole interaction operator

$$\widehat{H}_{\rm int} = -\widehat{\boldsymbol{\mathcal{E}}}\widehat{\boldsymbol{D}} , \qquad (6.2.7)$$

 $\begin{array}{c} {\rm Landolt-B\" ornstein} \\ {\rm New \ Series \ VIII/1A2} \end{array}$

$$\widehat{\boldsymbol{\mathcal{E}}} = i \sum_{\boldsymbol{k}\sigma} \boldsymbol{\epsilon}_{\boldsymbol{k}\sigma} \sqrt{\hbar \omega_k / 2\varepsilon_0 V} \left(\widehat{a}_{\boldsymbol{k}\sigma} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} - h.c. \right) = \widehat{\boldsymbol{\mathcal{E}}}^+ + \widehat{\boldsymbol{\mathcal{E}}}^-$$
(6.2.8)

is the electric field operator, written as the sum of its positive and negative frequency parts as in (5.1.14), and which is to be taken at the position of the atoms. Therefore, if the system starts from the symmetric state, where all the atoms are inverted $(M = \frac{1}{2}N)$, then it will remain during the evolution in the totally symmetric N + 1-dimensional subspace given by (6.2.6). Starting from the uppermost state with M = N/2, the system first emits one photon and gets into the state with M = N/2 - 1, where all atoms but one are excited, then to the state with M = N/2 - 2, where all atoms but two are excited, etc. Because of the symmetric superposition of the individual atomic states, we cannot say which of the atoms has emitted the photon, we can only state that the number of excited atoms has become less by one. The states given by (6.2.6) are typical examples of entangled states. It is important to note that besides those given in (6.2.6) there exist other common eigenstates of $\hat{\mathcal{R}}^2$ and $\hat{\mathcal{R}}_3$, belonging to smaller eigenvalues of $\hat{\mathcal{R}}^2$, but these are not totally symmetric and the transitions between levels belonging to different eigenvalues of $\hat{\mathcal{R}}^2$ are forbidden. This latter effect is called *subradiance* [85Pav, 96Ben, 96DeV].

Calculating the matrix elements of the total dipole moment operator \hat{D} between the states given in (6.2.6), we can find the transition probabilities for the allowed transitions and we obtain [71Bon, 74Aga, 96Ben] a master equation for the probability P_M of observing the ensemble of atoms in the symmetric state of energy E_M :

$$\frac{\mathrm{d}P_M}{\mathrm{d}t} = \gamma(\frac{1}{2}N + M + 1)(\frac{1}{2}N - M)P_{M+1} - \gamma(\frac{1}{2}N + M)(\frac{1}{2}N - M + 1)P_M .$$
(6.2.9)

Superradiance corresponds to the initial condition: $P_{\frac{N}{2}}(t=0) = 1$, and $P_M(t=0) = 0$, $M \neq \frac{1}{2}N$, and the evolution of the atomic system appears as a cascade emission down the "ladder" of the N+1equidistant levels given in (6.2.6). While the system reaches level M, it emits n = N/2 - M photons, therefore the solution of (6.2.9) gives information about the number of the emitted photons. As it is seen from the coefficients in (6.2.9) the emission of the photons is getting exponentially faster and faster while M decreases from N/2 till around 0, and then the emission slows down again while M goes to -N/2. An approximate analytic solution [71Bon, 74Aga, 96Ben] of (6.2.9) yields the following expression for the radiated intensity given as the number of emitted photons per unit time:

$$I(t) = \gamma \frac{1}{4} N^2 \operatorname{sech}^2 [\gamma \frac{1}{2} N(t - t_{\rm D})], \qquad (6.2.10)$$

where $t_{\rm D} = (\gamma N)^{-1} \ln N$. From (6.2.10) the main features of the superradiant pulse are readily seen:

- 1. The initial intensity of the pulse is the same as that of ordinary spontaneous emission of N atoms: $I(0) = \gamma N$.
- 2. Its duration measured at half maximum is of the order $(\gamma N)^{-1}$, i.e. N times shorter than the radiation decay time of a single atom.
- 3. The maximum of the intensity is proportional to the square of the total number of atoms N, and appears after a delay time of $t_{\rm D}$.

These properties keep their validity for a macroscopically extended system, as well, with the exception that the value of the delay time is modified, see (6.2.19) below. From the solution of (6.2.9) one can also find the photon statistics of the emitted pulse. In the central part of the SR pulse – when M = 0 – it is nearly Poissonian, and hence the radiation is coherent, while at the beginning and at the end it shows larger fluctuations in photon number [71Bon, 96Ben].

The picture outlined above for the small system is an idealized one. In reality, the near-field dipole–dipole coupling between the atoms will lead to a site-dependent perturbation of the relevant

atomic levels. This effect destroys the coherence of the identical atomic emitters, and the radiated intensity remains proportional only to N, even in the absence of other dephasing effects [72Fri].

There are two ways to circumvent the problem above. The first one is to place the atomic system into a resonant cavity. Spontaneous emission of atoms is a consequence of the interaction with fluctuations of the huge number of modes of the electromagnetic vacuum surrounding the atoms. A cavity of appropriate dimension modifies the mode structure of the field around the atoms so that only a few or just one resonant mode is present, and therefore the density of the possible final states is considerably reduced. The dynamics of the system is then described by a modified master equation. Cooperative emission of a system of Rydberg atoms in a cavity has been observed by Haroche and coworkers, for a review see [85Har]. The same effect has been demonstrated for a system of nuclear spins in a magnetic field [90Baz, 96Ben]. This enhancing effect of a cavity had been first considered in [54Blo].

The other possibility to retain the cooperativity of the atoms is to use an extended system, which is discussed in the following Sect. 6.2.2.2.

6.2.2.2 Superradiation of an extended multiatomic system

In a cylindrical, pencil-shaped extended system the collective interaction will be effective only with the photon modes, for which the wave vectors are nearly parallel with the axis of the cylinder. The atoms have symmetric positions with respect to this mode, and then the influence of near-field dipole–dipole coupling is negligible. In this case, however, only these modes will give a contribution to the probability of the spontaneous decay, and their density is proportional to the diffraction solid-angle λ^2/A , where λ is the wavelength and A is the cross-section of the sample. Therefore, the emission will go in this narrow cone in both directions, and the decay rate – being reduced by the same factor – will be be of the order of $\gamma N \lambda^2/A$. The time scale of superradiance for an extended pencil-shaped sample is the inverse of this collective rate. It is customary to introduce an additional numerical factor and to define *superradiation time* as:

$$T_{\rm R} = \left(\frac{3}{8\pi}\gamma N\,\lambda^2/A\right)^{-1} = \left(\frac{3}{8\pi}\gamma N_0\lambda^2 L\right)^{-1} = 2\,\varepsilon_0\frac{\hbar c}{N_0 d^2\omega_0 L}\,,\tag{6.2.11}$$

where N_0 is the number of inverted atoms per unit volume in the sample of length L. Note, however, that some authors use a different factor, and define superradiation time as $T'_R = 3T_R!$ We see that for an extended system the duration of the superradiant pulse decreases by a factor proportional to the number of atoms in the volume $\lambda^2 L$, rather than by a factor equal to the total number of atoms N which is the case in a small sample.

In order to describe SR in an extended system in more detail one has to take into account the spatial positions of the atoms, as well as the propagation of the emitted electromagnetic field. One can derive the relevant equations by a quantum-electrodynamical calculation, see e.g. [79Haa, 79Pol, 81Sch, 82Gro, 96Ben]. To have a tractable description, one introduces *smoothly varying* fields by averaging the microscopic operators over a volume much smaller than the relevant wavelength, but still containing several atoms: $\hat{\mathcal{R}}_{\alpha}(\mathbf{r}) = \frac{1}{\Delta V N_0} \sum_{\mathbf{r}_i \in \Delta V} \hat{\mathcal{R}}^i_{\alpha}$. One considers a onedimensional problem: a cylinder of length $L \gg \sqrt{A} \gg \lambda$, by treating propagation only along the axis of the cylinder, the x axis, and neglecting the transversal variation of the fields. Then the standard approximations – used also in laser theory – can be exploited. Slowly varying, scalar electric field operator amplitudes \hat{E}^{\pm} and atomic operators \hat{R}_{\pm} and \hat{Z} are introduced, so that factors of the form $e^{\pm i(kx-\omega_0 t)}$ are separated from the variables:

$$\widehat{\mathcal{E}}^{\pm}(x,t) = \frac{\hbar}{2dT_{\rm R}}\widehat{E}^{\pm}(x,t)\mathrm{e}^{\pm\mathrm{i}(kx-\omega t)} , \quad \widehat{\mathcal{R}}_{\pm}(x,t) = \frac{1}{2}\widehat{R}_{\pm}(x,t)\mathrm{e}^{\pm\mathrm{i}(kx-\omega t)} , \quad \widehat{Z} = 2\widehat{\mathcal{R}}_3 . \quad (6.2.12)$$

The coefficient of \hat{E}^{\pm} , containing $T_{\rm R}$ – the superradiation time introduced in (6.2.11) – has been chosen so that the slowly varying electric field amplitude is dimensionless, as well as \hat{R}_{\pm} , and \hat{Z} . Then using the Rotating Wave Approximation (RWA) and the Slowly Varying Envelope Approximation (SVEA) one arrives at the following system:

$$\left(\frac{\partial}{\partial x} + \frac{1}{c}\frac{\partial}{\partial t}\right)\widehat{E}^{\pm} = \pm \frac{\mathrm{i}}{L}\widehat{R}_{\mp} ,$$

$$\frac{\partial}{\partial t}\widehat{R}_{\pm} = \frac{\mathrm{i}}{T_{\mathrm{R}}}\widehat{E}^{\mp}\widehat{Z} , \quad \frac{\partial}{\partial t}\widehat{Z} = \frac{\mathrm{i}}{2T_{\mathrm{R}}}\left(\widehat{R}_{+}\widehat{E}^{+} - \widehat{E}^{-}\widehat{R}_{-}\right) .$$
(6.2.13)

It is impossible to solve these nonlinear Quantum Maxwell–Bloch (QMB) equations. In order to obtain results to be compared with experiments, one has to simplify the problem to a semiclassical treatment, which amounts to replace the operators in equations (6.2.13) by corresponding classical complex variables. With the convention: $\hat{E}^+ \to E^*, \ \hat{E}^- \to E, \ \hat{R}_+ \to R, \ \hat{R}_- \to R^*, \ \hat{Z} \to Z,$ one obtains the undamped semiclassical Maxwell–Bloch (MB) equations, which are essentially the same as those in (1.1.45) but for the dimensionless classical quantities used in the present section. The necessity of retaining at least some elements of a quantum treatment is seen, however, from the following fact: If the variables in (6.2.13) are replaced by their classical counterparts, and the initial conditions correspond to total inversion density: Z = 1, zero polarization: $R = R^* = 0$, and no field present along the whole sample, then the system remains in that state, and does not start to radiate, as initially all the derivatives are equal to zero. This shows that quantum fluctuations are especially important in the initial stage of the process. This quantum effect manifests itself in the experiments, as well, the characteristic properties of the emitted pulses show fluctuations in repeated observations. The quantum nature of the initial stage can be incorporated into the semiclassical theory by a method worked out by Glauber and Haake, and by Polder and Schuurmans [79Haa, 79Pol, 81Haa, 81Sch]. The quantum correlation function of the atomic polarization operators in the initial state $|\{e\}\rangle$ (i.e. all the atoms in the upper state $|e\rangle$) gives:

$$\langle \{\mathbf{e}\}|\hat{R}_{+}(x,t=0)\hat{R}_{-}(x',t=0)|\{\mathbf{e}\}\rangle = \frac{4L}{N}\delta(x-x')$$
, (6.2.14)

which is the consequence of the relation $\langle \{e\} | \hat{\mathcal{R}}^i_+ \hat{\mathcal{R}}^j_- | \{e\} \rangle = \delta_{ij}$ and the smoothing procedure. Considering the variables in (6.2.13) as classical stochastic quantities one solves the system for a large number of random initial conditions with correlation functions corresponding to (6.2.14). The average over the realizations is to be compared with the average experimental pulse forms.

Two comments on the one-dimensional approximation leading to (6.2.13) are in order:

- 1. As there are no end mirrors like in a laser, diffraction effects at the output of the long cylinder cannot be neglected. The reduction of diffraction would necessitate a sample with a large Fresnel number $\mathcal{F} = \mathcal{A}/\lambda L$. This, on the other hand, would contradict to the one-dimensional model. The optimal choice for the geometry is therefore to choose a Fresnel number close to 1 [79Pol, 82Gro, 96Ben]. Theories that take into account diffraction in more detail have been considered in [81Mat, 89Ave, 89Sch, 91Ave, 96Ben].
- 2. The form of SVEA in (6.2.12) assumes the propagation only in the positive x direction. There should be other terms with phase ($\omega t + kx$) describing waves propagating in the opposite direction. In the observations of SR both pulses appear, but according to the experiments [81Vre] the two pulses can be considered to be independent, at least in gases. In solids with a greater density of active atoms there seems to be a correlation between the counterpropagating pulses, for an experiment see [84Flo]. Theoretical proposals explaining this correlation can be found in [92Jan, 94Tri, 94Mal, 96Ben].

In order to take into account relaxation effects, one can introduce phenomenological damping terms with time constants T_1 and T_2 describing population decay and polarization dephasing, respectively. The distribution of the atoms according to their resonant frequencies – inhomogeneous broadening – can be also included by a line shape function $g(\nu) = g(\omega_0 - \omega_c)$, where ω_c is a central

frequency assumed to be identical with the carrier frequency of the field. $g(\nu)$ is usually a Gaussian of the form $g(\nu) = (T_2^*/\sqrt{\pi}) \exp(-\nu^2 T_2^{*2})$. The resulting equations are then the same as the ones of semiclassical laser theory. We stress, however, that *superradiance in the pure sense is possible* only if all these additional dephasing effects are absent, i.e. the process is faster than T_1 , T_2 and T_2^* .

In SR the atoms of the system are initially in their excited states. Therefore, an extended superradiant sample is a coherent amplifier, because during propagation the field amplitude is growing. Owing to its principal similarity to superradiance, we also consider briefly the coherent attenuator which differs from SR in the initial conditions. In the attenuator all atoms are in the ground state in the beginning, and there is an injected external pulse.

In the undamped MB equations (see Sect. 1.1.7.3) the variables allow the following parameterization: $Z = \cos \theta$, $|R| = \sin \theta$. This means, that the values of the quasiclassical atomic variables can be illustrated by a point on the "Bloch sphere". Then, in the absence of phase modulation (when *E* is real), the classical counterparts of (6.2.13) yield $\theta(x,t) = \frac{1}{T_{\rm R}} \int_{-\infty}^{t} E(x,t') dt'$, which means that a pulse with electric field *E* generates a polarization and depopulates the upper level of the amplifier. Important information on coherent propagation of light pulses – in particular that of a superradiant pulse – is provided by a general result, the area theorem of McCall and Hahn [69McC, 75All]. The pulse area is defined as:

$$\mathcal{A}(x) := \theta(x, t = \infty) = \frac{1}{T_{\rm R}} \int_{-\infty}^{\infty} E(x, t') \,\mathrm{d}t' \,. \tag{6.2.15}$$

This dimensionless quantity obeys a simple differential equation:

$$\frac{\mathrm{d}\mathcal{A}}{\mathrm{d}x} = \pm \frac{\alpha}{2} \sin \mathcal{A} , \quad \text{with} \quad \alpha = \frac{2\pi g(0)}{LT_{\mathrm{R}}} , \qquad (6.2.16)$$

where the + and - signs correspond to the amplifier and attenuator, and α is the gain or the absorption coefficient, respectively. Equation (6.2.16) admits the following solution:

$$\tan(\mathcal{A}/2) = \tan(\mathcal{A}_0/2) \exp[\pm(\alpha/2)(x - x_0)].$$
(6.2.17)

In the case of an attenuator the asymptotic value $\mathcal{A}(\infty)$ is necessarily $2n\pi$, $n = 0, 1, 2, \ldots$, where n is determined by \mathcal{A}_0 . In the case of an amplifier the asymptotic value is $(2n + 1)\pi$ depending again on \mathcal{A}_0 . In SR, where initially no field is present, the asymptotic value of the pulse area is π .

The undamped Maxwell–Bloch system has analytical solutions for the attenuator. The simplest one is the particularly stable " 2π secanthyperbolic pulse" [69McC, 75All]:

$$E = \frac{2T_{\rm R}}{T} \operatorname{sech}\left[\frac{1}{T}\left(t - \frac{x}{v}\right)\right] \,. \tag{6.2.18}$$

The duration T and the velocity v of this pulse are in the relation: $c/v = 1 + \alpha cT/2$. This ratio can be of the order of 10^2 , the velocity of the coherent pulse is reduced by two orders of magnitude if compared with the incoherent propagation. The stability of the secanthyperbolic solution is explained in the mathematical theory of solitons. Experiments as well as theory show that all incoming pulses for which the area at the input is $\pi < A_0 < 3\pi$ are transformed into this type of pulse after sufficiently long propagation in an attenuator. McCall and Hahn have coined the name Self-Induced Transparency (SIT) to this effect, and demonstrated its remarkable properties in ruby [69McC]. Several other works on SIT are cited in [75All].

For superradiance, i.e. for an amplifier with the initial condition E(x, t = 0) = 0, there is no analytic solution of the MB equations even in the absence of the relaxation terms. We know, however, that during the propagation the *intensity* – the area below the square of the amplitude – must grow, while *the area* below the slowly varying field amplitude *cannot be larger than* π , if initially the field was zero. Therefore the slowly varying *field envelope* must change its sign several times. The sign change of the envelope means that the field oscillates in opposite phase than before, and therefore a series of pulses occur in the output radiation. This effect, called ringing, has been first predicted in [69Bur]. A characteristic property of an initially completely inverted superradiant system is that the evolution of the process begins after a certain delay time. In the experiments this delay of the radiated pulses shows fluctuations, which is the consequence of the intrinsic quantum nature of the effect. The mean value of the time instants when the first peak of the SR pulse appears after the excitation is denoted by $T_{\rm D}$. The analysis of the MB equations allows one to calculate its value, as well as the width of this first peak, $T_{\rm W}$. The calculation based on the correlation function (6.2.14) yields also the relative standard deviation $\Delta T_{\rm D} = (\langle T_{\rm D}^2 \rangle - \langle T_{\rm D} \rangle^2)^{1/2} / \langle T_{\rm D} \rangle$ of the delay time. One has [76Mac, 79Haa, 79Pol, 81Mac, 81Sch]:

$$T_{\rm D} = \frac{1}{4} T_{\rm R} \left[\ln \frac{R_0}{2\pi} \right]^2 , \quad T_{\rm W} = T_{\rm R} \ln \left| \frac{R_0}{2\pi} \right| , \quad \Delta T_{\rm D} = 2.3/(\ln N) , \qquad (6.2.19)$$

where R_0 is the average initial value of |R| in the MB equations, which – according to (6.2.14) – is given by $R_0 = 2/\sqrt{N}$. The initial fluctuations of $|R| = \sin \theta$ can be considered as the fluctuations of θ , and the small initial mean value is $\theta_0 \approx \sin \theta_0 = 2/\sqrt{N}$, while the initial phase of R is completely random. The experiment confirming this value of the initial "tipping angle" θ_0 [79Vre] will be described below.

We summarize now the conditions to observe superradiance in an extended system. The relaxation times must be longer than the cooperative emission time, $T_{\rm R}$, and if the atoms are to emit cooperatively, they must act as a single system, therefore the so-called Arecchi–Courtens cooperation time $T_{\rm c} := T_{\rm R}|_{L=cT_{\rm c}} = \sqrt{T_{\rm R}L/c}$ must be shorter than $T_{\rm R}$, otherwise the system breaks into a number of independently radiating segments. Accordingly the characteristic times have to satisfy:

$$L/c < T_{\rm c} < T_{\rm R} < T_{\rm D} < T_1, T_2, T_2^*$$
 (6.2.20)

The duration of the excitation $T_{\rm p}$ must also be shorter than the evolution of the process $T_{\rm D}$.

Sometimes the relation $T_{\rm R} < T_1, T_2, T_2^*$ is satisfied, but $T_{\rm D}$ is longer than the relaxation times. Then by injecting a small area pulse into the sample one can increase the initial polarization (tipping angle), and thus reduce $T_{\rm D}$ to make it shorter than the characteristic times of damping. In that case one speaks of *induced* or *triggered* superradiation [86Var, 87Mal].

If the time constant of the phase relaxation is shorter than the process: T_2 is smaller than $T_{\rm R}$ (this is the usual case), then the dominating term in the time derivative of R will be $-\frac{1}{T_2}R$, which has been omitted so far. Then, after a very short transition period the polarization becomes proportional to the field amplitude. In this way we arrive at the rate equation model of the atomic subsystem [75All, 96Ben], and the result is the incoherent process called Amplified Spontaneous Emission (ASE). It can be shown [96Ben] that inhomogeneous broadening leads to the same effect.

6.2.3 Superradiance experiments

The observations of SR can be classified according to the size of the sample compared with the emitted wavelength. Small sample experiments with Rydberg atoms in cavities [85Har], as well as with nuclear spins [90Baz] belong to one group. The other, much larger group consists of observations in extended pencil-shaped samples. A selective list of SR experiments is given in Table 6.2.1.

Material	Wavelength	Remarks	Ref.
HF	252, 126, 84, 63, 50 μm	first demonstration	[73Skr]
Na vapor	3.41, 2.21, 9.10 μm	first near infrared	[76Gro]
Tl vapor	1.3 µm	SR suppressed by inhomogeneous broadening	[76Flu]
Cs vapor	2.931 μm	single pulses	[77Gib]
Rb atomic beam	$2.79 \dots 1.3 \ \mu m \ (7 \ lines)$	polarization characteristics	[78 Cru]
Cs vapor	3.1 μm	Doppler beats between different groups of atoms	[78Gro]
Eu vapor	557.7, 545.3, 1759.6 nm	first visible SR	[79Bre]
Eu vapor	605.7 nm (dominant)	20 lines	[79Cah]
Rb	2.79 μm	beats from different isotopes	[79Mar]
Cs vapor	2.931 μm	measurement of tipping angle	[79Vre]
Cs vapor	2.931 μm	triggered SR	[80Car]
Cs vapor	2.931 μm	delay time dependence on Fresnel number	[81Vre]
Cs vapor	3.01 µm	hyperfine structure quantum beats	$[81 \mathrm{Rys}]$
CH_3F	$496 \ \mu m$	single pulse, wide range of lengths, cross sections, pressures	[81 Ros]
$C_{12}H_{10}$ in diphenyl	373.9 nm	SR in an organic material	[83Zin]
O_2 centers in KCl	592.8, 629.1 nm	two-color SR in a solid	[84Flo]
Cr: ruby, Nd:YAG at 100 K	694, 1061 nm	phase changes in ringing	[84Var]
Ga vapor	5.75, 1.21 μm	SR and the first observation of subradiance	[85Pav]
Cr: ruby, Nd:YAG at 100 K	694, 1061 nm	triggered SR, transition to ASE	[86Var]
O_2 in KCl	629.1 nm	transition to ASE	[87Mal]
O_2 centers in KCl	629.1 nm	two-dimensional SR from a lamel- lar sample	[87Sch]
O_2 centers in KCl	629.04 nm	spatial coherence of SR	[88Sch]
CH_3F	496 µm		[91Sch]
CH_3CN	373 μm		[91Sch]
NH ₃	291 μm	two-photon noise initiated SR	[92Bak]
Pair of Ba ions	493 nm	microscopically resolved SR and subradiance	$[96 \mathrm{DeV}]$
Na BE-condensate	589 nm	superradiant Rayleigh scattering accompanied by emission of di- rected atomic beams	[99Ino]

Table 6.2.1. Selected superradiance experiments in the infrared and visible.

6.2.3.1 A microscopic observation of superradiance and subradiance

In 1996 DeVoe and Brewer [96DeV] observed directly the collective spontaneous emission of two trapped ¹³⁸Ba⁺ ions as a function of the ion–ion separation r. The ions were confined to fixed positions in a radio-frequency Paul trap. For two atoms (ions) the decay from the fully excited state $|e, e\rangle$ to the ground state $|g, g\rangle$ may go through the symmetric state $|s\rangle := (|e, g\rangle + |g, e\rangle)/\sqrt{2}$, as well as through the antisymmetric state $|a\rangle := (|e, g\rangle - |g, e\rangle)/\sqrt{2}$. If the distance between the atoms r is much less than the wavelength of the emission, $kr \ll 1$, then the decay rate through the symmetric channel is twice that of the single atom rate (superradiance), see (6.2.9), while the transition probability through the antisymmetric channel is approximately zero. The latter is a special case of subradiance. However, if the correct r dependence of the electromagnetic interaction between the two atoms is taken into account, then both the symmetric, and the antisymmetric.

metric channels have a nonvanishing decay rate. In the experiment, so far, only $kr \gg 1$ could be realized, and then the following approximate formula is valid for the collective decay rates: $\Gamma_{\pm}(r) = \gamma(1 \pm \beta \sin(kr)/kr)$, corresponding to the symmetric (+) and the antisymmetric (-) channels. For an ideal case with nondegenerate levels the value of the constant is $\beta = 1.5$. The experimental conditions of [96DeV] corresponded to $\beta = 0.33$. The collective decay rate depends also on the relative orientations of the atomic dipoles. The formula for Γ_{\pm} above is for the case that the dipoles are parallel to each other and perpendicular to the vector \mathbf{r} joining them. This orientation could be achieved by a polarized exciting laser beam with polarization vector $\boldsymbol{\epsilon}_1$ perpendicular to \mathbf{r} . For collinear dipoles the collective effect is suppressed. (For the general dependence on the distance and relative orientation see [95Bre, 96Ben]). In the experiment the antisymmetric state $|a\rangle$ was excited and the decay rate was measured. The relevant part of the curve together with the experimental results are shown in Fig. 6.2.1.



Fig. 6.2.1. Theoretical (solid curve) and experimental values of collective lifetimes $(1/\Gamma_{-})$ at r = 1380, 1470, and 1540 nm. The single-ion lifetime in the same trap is shown by the dashed horizontal line. Full circles with error bars: exciting laser polarization $\epsilon_1 \perp \mathbf{r}$, the dipoles are parallel. Crosses (error bars omitted): $\epsilon_1 || \mathbf{r}$, the dipoles are collinear, a collective effect is not expected. The point for which the lifetime is smaller than the single-atom value is a signature of superradiance, while the point above that value corresponds to subradiance [96DeV].

6.2.3.2 Superradiance experiments in pencil-shaped macroscopic samples

These experiments can be divided again into two classes with respect to the time scale of the relaxation processes. In gases the relaxation times are typically in the nanosecond range whereas in solids these processes are faster, by at least one or two orders of magnitude. Therefore to avoid phase relaxation the solid samples must be kept at low temperature.

In order to observe SR, one has to achieve a possibly complete inversion on a nondegenerate two-level transition, without competing transitions. This is usually accomplished via a third higher level, or by a cascade decay through several levels. In some of the experiments magnetic fields have been applied to remove degeneracy. The excitation goes either along the axis of the cylinder or with uniform lateral pumping. The typical setup and level scheme is shown in Fig. 6.2.2 [73Skr, 76Mac]. The geometry of the sample corresponds usually to a Fresnel number $\mathcal{F} = 1$. The resulting series of SR pulses of the experiment [73Skr] are shown in Fig. 6.2.3. The multiple peaks are attributed to Burnham–Chiao ringing [69Bur, 96Ben]. According to an interference experiment in ruby [84Var], the consecutive peaks are in opposite phases, as required by the one-dimensional propagation theory (see the explanation in the second paragraph after (6.2.18)).



Fig. 6.2.2. Level scheme for hydrogen fluoride, HF, and principle of the experimental arrangement of the first observation of SR [73Skr, 76Mac]. The windows of the sample cell are tilted to prevent multiple reflections.

Fig. 6.2.3. Oscilloscope trace of an SR pulse observed in HF gas at 84 μ m on the $J = 3 \rightarrow 2$ rotational transition, and theoretical fit [73Skr]. The parameters are: pump intensity $I = 1 \text{ kW/cm}^2$, gas pressure p =1.3 mtorr, cell length L = 1 m. The small peak on the oscilloscope trace at t = 0 is the 3 μ m pump pulse, highly attenuated [73Skr].

Fig. 6.2.4. Delay time $T_{\rm D}$ of the output pulse vs. $[\ln(\theta/2\pi)]^2$. The dashed line is used to correct for the delay of the injection pulse with respect to the pump pulse [77Gib].

An important step in resolving the question of quantum initiation of SR was the experiment of Vrehen and Schuurmans [79Vre]. They injected small area pulses, so-called tipping pulses, at the relevant wavelength into the sample immediately after it had been completely inverted by a short pump pulse. As long as the area of the tipping pulse θ was smaller than θ_0 (see the note after (6.2.19)), the delay time of the output pulse would not be affected. However, when $\theta > \theta_0$ the delay was reduced. Therefore, by measuring the delay time as a function of the area of the injected pulse, the magnitude of θ_0 could be found. This plot is shown in Fig. 6.2.4. As expected, T_D increased linearly for large injection pulses. For small injection pulses the initiation of SR was dominated by quantum fluctuations, and T_D was constant. At the cross-over $\theta = \theta_0$ by definition. The results confirm the theoretical value $\theta_0 \simeq 2/\sqrt{N}$.

Theoretical [81Mat, 83Wat, 89Ave, 91Ave] and experimental [81Vre] investigations confirm that transverse effects can significantly influence the pulse shape. The excitation generates a Gaussian profile in the inversion and this in turn leads to a radial dependence of the intensity peaks, delay and ringing. According to the analysis of [81Mat], detectors not resolving this transversal dependence average out the ringing, this is how single pulses have been observed in [77Gib]. In the experiment [73Skr] the detector viewed a small area in the near field of the beam, and the central portion of the output pulse exhibited ringing (Fig. 6.2.3). According to another experiment, [81Vre], the observed results for the delay times were in reasonable agreement with the one-dimensional approach for Fresnel numbers $\mathcal{F} = 0.8...1$. For larger values of \mathcal{F} , both the delay time and its fluctuations were reduced. According to the estimations of [81Vre]: $\langle T_{\rm D} \rangle \sim |\ln 2\pi N/\mathcal{F}|^2$, and $\Delta(T_{\rm D}) \simeq 2.3/[\mathcal{F}\ln(N/\mathcal{F})]$. The measured reduction is the consequence of the above mentioned averaging detection method.

6.2.3.3 Superradiant-type Rayleigh scattering from a Bose–Einstein condensate

An effect which is not strictly SR, but very similar to it has been observed [99Ino] in a Bose–Einstein (BE) condensate, which is an ideal cooperative system of atoms. The sample, which was $L = 200 \ \mu\text{m}$ in length and 20 μm in diameter, contained several million sodium atoms. It was exposed to laser pulses that intersected the condensate perpendicular to its long axis with pulse durations between 10 and 800 μ s. This excitation was slightly red detuned from the $3S_{1/2}$, $F = 1 \rightarrow 3P_{3/2}$, F = 0, 1, 2 transition. Pulses of intensity above 1 mW/cm² triggered collimated emission of light beams from both ends of the sample along its long axis. Typical images of the emitted pulses consisted of a few bright spots with angular widths equal to the diffraction limit for a source with a diameter of 14 μ m.

An entirely new feature of this experiment was that the atoms recoiled by the interaction with the exciting photons could be directly seen as highly directional beams flying out from the sample, when the confining magnetic trap was turned off. As required by momentum conservation, these beams propagated at an angle of 45° with the long axis of the condensate.

The observed effect was actually cooperative Rayleigh scattering with the characteristic properties of superradiance. Time-resolved measurements showed that by increasing the intensity of the excitation the evolution of the emitted pulses was shortened and became more highly peaked. The emission was also accompanied by a faster than normal exponential decay of the number of atoms remaining in the condensate at rest after excitation. For sufficiently intense excitation more peaks appeared in the emission, and several bunches of recoiled atoms were observed, as a result of sequential scattering. The radiated intensity is expected to be proportional to the *square* of the contrast of the matter wave interference pattern, although in the experiment [99Ino] the density of the condensate has not been varied.

This strong cooperative effect could be observed only when the polarization of the exciting laser was perpendicular to the long axis of the condensate, so that dipole radiation from the atoms could be coherently amplified along this axis. To the contrary, excitation with polarization parallel to the long axis resulted in ordinary weak isotropic Rayleigh scattering. The superradiant features also disappeared when the temperature was raised above the condensation value, mainly because the collective phase memory time T_2^* was then reduced by Doppler broadening with a factor of 30. Therefore the appearance of the superradiant emission indicated sensitively the appearance of the Bose–Einstein phase transition.

6.2.4 Outlook

The fast development of experimental techniques for trapping atoms and ions will allow to extend SR experiments in the microscopic domain. As a further development of the works described in [85Har] and [96DeV] one expects to observe SR from a few atoms (well-defined number) confined in a trap in a possibly symmetrical arrangement. This technique is promising also in modeling and probably realizing information processing with quantum-optical methods (quantum informatics), see also Sect. 5.1.10.3. At present the main obstacle in the way of making a quantum computer is decoherence, the same effect that usually destroys superradiance.

As for the extended systems, the wavelength can be further reduced. A strong stimulation of superradiance research is the fact that it can be an intensive source of X rays and γ rays from nuclear transitions, where no easily handled mirrors are present (Vol. VIII/1B, Chap. 7.1: "X-ray lasers"). Note also that the $\sim N^2$ dependence of the intensity has already been realized in Free Electron Lasers (FEL). In this sense a FEL can also be considered as a source of superradiation (Vol. VIII/1B, Chap. 6.1: "Free-electron lasers"). Possible other sources of SR could be the electron-hole plasma and/or excitons in semiconductors. In these cases the characteristic quadratic dependence of the radiated intensity on the number of the emitting centers must be proven. One also expects new experiments demonstrating SR from Bose–Einstein condensates. In connection with lasing without inversion, also a scheme of superradiance without inversion has been considered [98Mal].

Several theoretical works predict special effects in photon statistics if superradiance is initiated by squeezed vacuum instead of true vacuum, for references see [96Ben]. Even a "superradiant laser" has been proposed by Haake et al. which would produce intensive squeezed light [93Haa]. It is interesting, however, that experimental investigations of the photon statistics of any superradiant light are completely missing.

References for 6.2

27Dir	Dirac, P.A.M.: Proc. R. Soc. (London) A 114 (1927) 243.
30Wei	Weisskopf, V., Wigner, E.: Z. Phys. 63 (1930) 54.
54Blo 54Dic	Bloembergen, N., Pound, R.V.: Phys. Rev. 95 (1954) 8. Dicke, R.H.: Phys. Rev. 93 (1954) 99.
69Bur 69McC	Burnham, D.C., Chiao, R.Y.: Phys. Rev. 188 (1969) 667. McCall, S.L., Hahn, E.L.: Phys. Rev. 183 (1969) 457.
71Bon	Bonifacio, R., Schwendimann, P., Haake, F.: Phys. Rev. A 4 (1971) 302.
72Fri	Friedberg, R., Hartmann, S.R., Manassah, J.T.: Phys. Lett. A 40 (1972) 365.
73Skr	Skribanowitz, N., Herman, I.P., MacGillivray, J.C., Feld, M.S.: Phys. Rev. Lett. 30 (1973) 309.
74Aga	Agarwal, G.S.: Quantum statistical theories of spontaneous emission, Springer tracts in modern physics, Vol. 70, Quantum optics, Berlin: Springer-Verlag, 1974.
75All 75Bon	Allen, L., Eberly, J.H.: Optical resonance and two-level atoms, New York: Wiley, 1975. Bonifacio, R., Lugiato, L.A.: Phys. Rev. A 11 (1975) 1507.
76Flu 76Gro 76Mac	 Flusberg, A., Mossberg, T., Hartmann, S.R.: Phys. Lett. A 58 (1976) 373. Gross, M., Fabré, C., Pillet, P., Haroche, S.: Phys. Rev. Lett. 36 (1976) 1035. MacGillivray, J.C., Feld, M.S.: Phys. Rev. A 14 (1976) 1169.
77Gib	Gibbs, H.M., Vrehen, Q.H.F., Hikspoors, H.M.J.: Phys. Rev. Lett. 39 (1977) 547.
78Cru 78Gro	Crubellier, A., Liberman, S., Pillet, P.: Phys. Rev. Lett. 30 (1978) 1237. Gross, M., Raimond, J.M., Haroche, S.: Phys. Rev. Lett. 40 (1978) 1711.
79Bre 79Cah 79Haa 79Mar 79Pol 79Vre	 Brechignac, C., Cahuzac, Ph.: J. Phys. Lett. 40 (1979) L123. Cahuzac, Ph., Sontag, H., Toschek, P.E.: Opt. Commun. 31 (1979) 37. Haake, F., King, H., Schröder, G., Haus, J., Glauber, R.: Phys. Rev. A 20 (1979) 2047. Marek, J.: J. Phys. B. 12 (1979) L229. Polder, D., Schuurmans, M.F.H., Vrehen, Q.H.F.: Phys. Rev. A 19 (1979) 1192. Vrehen, Q.H.F., Schuurmans, M.F.H.: Phys. Rev. Lett. 42 (1979) 224.
80Car	Carlson, N.W., Jackson, D.J., Schawlow, A.L., Gross, M., Haroche, S.: Opt. Commun. 32 (1980) 350.
81Haa 81Mac 81Mat 81Ros 81Rys 81Sch 81Vre	 Haake, F., Haus, J., King, H., Schröder, G., Glauber, R.: Phys. Rev. A 23 (1981) 1322. MacGillivray, J.C., Feld, M.S.: Phys. Rev. A 23 (1981) 1134. Mattar, F.P., Gibbs, H.M., McCall, S.L., Feld, M.S.: Phys. Rev. Lett. 46 (1981) 1123. Rosenberger, A.T., DeTemple, T.A.: Phys. Rev. A 24 (1981) 868. Ryschka, M., Marek, J.: Phys. Lett. A 86 (1981) 98. Schuurmans, M.F.H., Vrehen, G.F.H., Polder, D.: Adv. At. Mol. Phys. 17 (1981) 167. Vrehen, Q.H.F., Weduve, J.J.: Phys. Rev. A 24 (1981) 2857.

80	References for 6.2
82Gro 82Vre	Gross, M., Haroche, S.: Phys. Rep. 93 (1982) 301. Vrehen, Q.H.F., Gibbs, H.M., in: Topics in current physics, Vol. 21, Dissipative systems in quantum optics, Bonifacio, R. (ed.), Berlin: Springer-Verlag, 1982, p. 111.
83Wat 83Zin	Watson, E.A., Gibbs, H.M., Mattar, F.P., Cornier, M., Claude, Y., McCall, S.L., Feld, M.S.: Phys. Rev. A 27 (1983) 1427. Zinoviev, P.V., Lopina, S.V., Naboikin, Yu.V., Silaeva, M.B., Samartsev, V.V., Sheibut, Yu.E.: Zh. Eksp. Teor. Fiz. 85 (1983) 1945; Sov. Phys. JETP (English Transl.) 58 (1983) 1129.
84Flo 84Var	 Florian, R., Schwan, L.O., Schmid, D.: Phys. Rev. A 29 (1984) 2709. Varnavsky, O.P., Kirkin, A.N., Leontovich, A.S., Malikov, R.F., Mozharovsky, A.M., Trifonov, E.D.: Zh. Eksp. Teor. Fiz. 86 (1984) 1227; Sov. Phys. JETP (English Transl.) 59 (1984) 716.
85Har 85Pav	Haroche, S., Raimond, J.M.: Adv. At. Mol. Opt. Phys. 20 (1985) 347. Pavolini, D., Crubellier, A., Pillet, P., Cabaret, L., Liberman, S.: Phys. Rev. Lett. 54 (1985) 1917.
86Leo 86Var	Leonardi, C., Persico, F., Vetri, G.: Riv. Nuovo Cimento 9 (1986) 1. Varnavsky, O.P., Golovlev, V.V., Kirkin, A.N., Malikov, R.F., Mozharovsky, A.M., Benedict, M.G., Trifonov, E.D.: Zh. Eksp. Teor. Fiz. 90 (1986) 1596; Sov. Phys. JETP (English Transl.) 63 (1986) 937.
87Mal 87Sch	Malcuit, M.S., Maki, J.J., Simkin, D.J., Boyd, R.W.: Phys. Rev. Lett. 59 (1987) 1189. Schiller, A., Schwan, L.O., Schmid, D.: J. Lumin. 38 (1987) 243.
88Sch	Schiller, A., Schwan, L.O., Schmid, D.: J. Lumin. 40–41 (1988) 541.
89Ave 89Bau 89Sch 89Zhe	 Avetisyan, Yu.A., Zaitsev, A.I., Malyshev, V.A., Trifonov, E.D.: Zh. Eksp. Teor. Fiz. 95 (1989) 1541; Sov. Phys. JETP (English Transl.) 68 (1989) 891. Bausch, R., Borgs, P., Kree, R., Haake, F.: Europhys. Lett. 10 (1989) 445. Schwan, L.O., Schwendiman, P., Sigmund, E.: Phys. Rev. A 40 (1989) 7093. Zheleznyakov, V.V., Kocharovsky, V.V., Kocharovsky, Vl.V.: Usp. Fiz. Nauk 159 (1989) 193; Sov. Phys. Usp. (English Transl.) 32 (1989) 835.
90Baz	Bazhanov, N.A., Bulyanitsa, D.S., Zaitsev, A.I., Kovalev, A.I., Malyshev, V.A., Tri- fonov, E.D.: Zh. Eksp. Teor. Fiz. 97 (1990) 1995; Sov. Phys. JETP 70 (1990) 1128.
91Ave 91Sch	Avetisyan, Yu.A., Zaitsev, A.I., Malyshev, V.A., Trifonov, E.D.: Opt. Spektrosk. 70 (1991) 1345; Opt. Spectrosc. (English Transl.) 70 (1991) 786. Scherrer, D.P., Kälin, A.W., Kesselring, R., Kneubühl, F.K.: Appl. Phys. B. 53 (1991) 250.
92Bak 92Jan	Bakos, J.S., Scherrer, D.P., Kneubühl, F.K.: Phys. Rev. A 46 (1992) 410. Janssen, D., Stahl, A.: Europhys. Lett. 18 (1992) 33.
93And 93Haa	Andreev, A.V., Emelyanov, V.I., Ilinskiî, Yu.A.: Cooperative effects in optics, Bristol: IOP, 1993. Haake, F., Kolobov, M., Fabré, C., Giacobino, E., Reynaud, S.: Phys. Rev. Lett. 71 (1993) 995.

- 94Mal Malyshev, V.A., Trifonov, E.D., Schwan, L.O.: Opt. Spektrosk. 76 (1994) 524; Sov. Phys. Opt. Spectrosc. (English Transl.) 76 (1994) 470.
- 94Tri Trifonov, E.D.: Opt. Spektrosk. **77** (1994) 61; Opt. Spectrosc. (English Transl.) **77** (1994) 51.
- 95Bre Brewer, R.G.: Phys. Rev. A 52 (1995) 2965.
- 96Ben Benedict, M.G., Ermolaev, A.M., Malyshev, V.A., Sokolov, V.I., Trifonov, E.D.: Superradiance, Bristol: IOP, 1996.
- 96DeV DeVoe, R.G., Brewer, R.G.: Phys. Rev. Lett. 76 (1996) 2049.
- 98Mal Malyshev, V.A., Ryzhov, I.V., Trifonov, E.D., Zaitsev, A.I.: Laser Phys. 8 (1998) 494.
- 99Ino Inouye, S., Chikkatur, A.P., Stamper-Kurn, D.M., Stenger, J., Pritchard, D.E., Ketterle, W.: Science 285 (1999) 571.

7.1 Modulators

B. Kuhlow

7.1.1 Introduction

Modulation of light is used for a number of applications including the impression of information onto optical beams, Q-switching of lasers, for generation of giant optical pulses, and optical beam deflection.

Besides direct modulation by changing the laser current there are two main methods in modulation: internal and external modulation. The internal modulation method uses a change of the quality factor, Q, in a laser cavity by inserting a modulator. Due to the very sensitive lasing condition on losses in the cavity a highly efficient switching operation can be obtained. Therefore, this method is used for laser Q-switching and for active mode locking to generate ultra-short optical pulses. Internal modulators are very similar to external modulators that are widely used. Therefore, the latter ones will be treated next in more detail. However, an important requirement for intracavity modulators is the use of low loss optical materials mostly combined with the shortest possible optical length.

Light modulation is defined as a change of the amplitude or intensity, phase, polarization, or frequency of an incident light wave. It can be caused by changing the refractive index, absorption coefficient, or direction of transmitted light in the modulator medium. Various physical effects are known for interaction of light with external forces. An easy and simple way to deflect or switch a light beam is to use mechanical methods, e.g. moving mirrors or shutters, but their modulation frequencies are rather low. Higher frequencies are attainable with electrical or optical addressable effects. Among these absorption changes in semiconductors and dyes may be efficient and fast but usable laser wavelengths are very limited. Of major importance, however, is the method of refractive index change in dielectrics and semiconductors, which is efficiently used in many kinds of optical modulators including the important electro-optic and acousto-optic modulators. Especially the electro-optic effect allows for much higher modulation frequencies than all other methods.

A large variety of modulation devices have been investigated [72Den, 84Yar, 97Lee]. But, generally, there is no single "best" modulator. The choice of device is highly dependent on the specific application. The simplest type of modulator is the temporal-only or point modulator, where the modulation signal at the output is identical within any cross-section perpendicular to the propagation of light beam. A more general modulation device performing both spatial and temporal modulation is usually called a spatial light modulator. A survey of different modulator types is given in [97Lee].

In what follows only bulk electro-optic and acousto-optic light modulators and deflectors as well as a few examples of electro-optic Kerr cells using liquids will be treated which all are of major importance in laser applications.

7.1.2 Light propagation in crystals

The basic idea behind electro-optic and acousto-optic devices is to alter the optical properties of a material with an applied voltage in a controlled way. In certain types of crystals it is possible to effect a change in the index of refraction that is proportional to an electric field or acoustic strain field. These are referred to, respectively, as the electro-optic and photoelastic effects. They allow widely used means of controlling the intensity or phase of light when propagating through the crystal.

For any anisotropic (optically inactive) crystal class there are in general two orthogonal linearly polarized waves (i.e. direction of D) with corresponding different phase velocities (i.e. index of refraction) for a given wave vector. The mutually orthogonal polarization directions and the indices of refraction of the two waves are found most easily in terms of the index ellipsoid (or indicatrix) that is a construct with geometric characteristics [84Yar].

The index ellipsoid assumes its simplest form in the principal coordinate system as

$$\frac{x^2}{n_x^2} + \frac{y^2}{n_y^2} + \frac{z^2}{n_z^2} = 1.$$
(7.1.1)

Here n_x , n_y , n_z are the principal values of refraction along the principal crystal axes x, y, z, which are the directions in the crystal along which D and E are parallel.

Three types of optical crystals are possible:

- 1. biaxial crystals $(n_x \neq n_y \neq n_z)$,
- 2. uniaxial crystals $(n_x = n_y = n_o; n_z = n_e),$
- 3. and isotropic crystals $(n_x = n_y = n_z = n_o)$.

They are established by the optical symmetry of a crystal and the corresponding dielectric tensor. The existence of an "ordinary" and an "extraordinary" ray with different indices of refraction is called birefringence. For a general direction of propagation, the section of the ellipsoid through the origin perpendicular to this direction is an ellipse where the major and minor axes represent the two orthogonal polarizations of D and the indices of refraction for this particular direction of propagation.

An electric field applied to certain crystals causes a slight change in the index of refraction, due to a redistribution of bond charges at the atomic level and possibly a slight deformation of the crystal lattice. A similar effect can be obtained by an acoustic strain field. In general, these alterations are not isotropic; that is, the changes vary with direction in the crystal. The net result is a change in the inverse dielectric constants (impermeability) tensor $1/n_{ij}^2$ of the index ellipsoid in its general form:

$$\Delta \left(\frac{1}{n^2}\right)_{ij} = r_{ijk}E_k + s_{ijkl}E_kE_l + p_{ijkl}S_{kl} , \quad i, j, k, l = 1, 2, 3 = x, y, z , \qquad (7.1.2)$$

where E_x , E_y , E_z are the components of the applied electrical field vector in principal coordinates and S_{kl} are the second-order strain tensor elements defined by the symmetric part of the deformation gradient $S_{kl} = (\partial u_k / \partial x_l + \partial u_l / \partial x_k)/2$, where $u_{l,k}$ is the displacement. A summation over repeated indices is assumed. The linear (or Pockels) electro-optic effect is represented by the third-rank tensor r_{ijk} . The quadratic (or Kerr) electro-optic effect and the photoelastic effect are represented by fourth-rank tensors s_{ijkl} and p_{ijkl} , respectively. Because of permutation symmetries the linear electro-optic tensor r_{ijk} reduces to a 6×3 matrix and the fourth-rank tensors s_{ijkl} and p_{ijkl} reduce to a 6×6 matrix each. This leads to contracted notations r_{ik} and s_{ij} , p_{ij} , where $i, j = 1, \ldots, 6$ and k = 1, 2, 3. The form of these tensors, but not the magnitude of tensor elements, can be derived from symmetry considerations, which dictate which of the coefficients are zero, as well as the relationships between the remaining coefficients [67Nye, 72Wem, 84Yar]. Crystals lacking a center of symmetry are non-centrosymmetric and exhibit a linear (Pockels) electro-optic effect. The induced changes in the impermeability tensor are linear in an applied electric field. On the other hand, all crystals exhibit a quadratic (Kerr) electro-optic effect as well as a photoelastic effect. Then, the changes in the impermeability tensor elements are quadratic in an applied electric field. When the linear effect is present, it generally dominates over all quadratic effects.

7.1.3 Linear electro-optic effect

The Pockels effect is named after the German physicist Friedrich Carl Alwin Pockels (1865–1913) who studied this phenomenon extensively in the year 1893. An electric field applied in a general direction to a non-centrosymmetric crystal produces a linear change in the constants $(1/n^2)_i$ due to the linear electro-optic effect according to

$$\Delta\left(\frac{1}{n^2}\right)_i = r_{ik}E_k , \quad i = 1, \dots, 6 \quad \text{and} \quad k = 1, 2, 3 = x, y, z , \qquad (7.1.3)$$

where r_{ik} are the elements of the linear electro-optic tensor in contracted notation and E_x , E_y , E_z are the components of the applied electrical field in principal coordinates. A summation over repeated indices k is assumed. Using the rules of matrix multiplication, we have, for example $\Delta \left(\frac{1}{n^2}\right)_6 = r_{61}E_1 + r_{62}E_2 + r_{63}E_3$.

The magnitude of $\Delta(1/n^2)$ is typically in the order of less than 10^{-5} . Therefore, these changes are small perturbations. The index ellipsoid in the presence of an electric field can be written as [84Yar]:

$$\left(\frac{1}{n_x^2} + r_{1k}E_k\right)x^2 + \left(\frac{1}{n_y^2} + r_{2k}E_k\right)y^2 + \left(\frac{1}{n_z^2} + r_{3k}E_k\right)z^2 + 2yzr_{4k}E_k + 2zxr_{5k}E_k + 2xyr_{6k}E_k = 1.$$
(7.1.4)

The presence of cross terms indicates that the ellipsoid is rotated and the lengths of the principal dielectric axes are changed. A new set of principal axes x', y', and z' can always be found by a coordinate rotation where the perturbed ellipsoid will be represented again in the standard form, see (7.1.1),

$$\frac{x'^2}{n_{x'}^2} + \frac{y'^2}{n_{y'}^2} + \frac{z'^2}{n_{z'}^2} = 1$$
(7.1.5)

with eigenvalues $1/n_{x'}^2$, $1/n_{y'}^2$, $1/n_{z'}^2$ depending on the strength and direction of the applied field and of course on the given matrix elements r_{ij} for the used crystal class. Extensive tables of linear electro-optic coefficients are listed in [79Coo].

7.1.3.1 Modulator devices

An electro-optic modulator is a device with operation based on an electrically induced change in index of refraction or change in natural birefringence. Two types of modulators can be classified: *longitudinal* and *transversal* ones.

In the *longitudinal* configuration the voltage is applied parallel to the direction of light propagation in the device. The electrodes must be transparent for the light or a small aperture at the

center at each end of the crystal has to be introduced. The applied electric field inside the crystal is E = V/L, where L is the length of the crystal. The induced phase shift is proportional to the voltage V but not to physical dimensions of the device.

In transverse configuration the voltage is applied perpendicular to the direction of light propagation. The electrodes do not obstruct the light as it passes through the crystal. The applied electric field is E = V/d, where d is the lateral separation of the electrodes. Thus, the voltage necessary to achieve a desired degree of modulation can be largely reduced by reduction of d. However, the transverse dimension d is limited by the increase of capacitance, which affects the modulation bandwidth or speed of the device.

Different kinds of modulation can be obtained using an electro-optic bulk modulator in combination with polarizers and passive birefringence elements [84Yar, 95Mal]. Usually it is desirable to modulate only one optical parameter at a time; simultaneous changes in the other parameters must then be avoided. First, it is assumed that the modulation field is uniform throughout the length of the crystal and the modulation frequency is very low ($\omega_m \ll 2\pi c/nL$).

Examples of modulator devices using the linear electro-optic effect in KDP, ADP, LiNbO₃, LiTaO₃, GaSe, and the quadratic electro-optic effect in BaTiO₃ are given in [84Yar, 89Gha]. Pockels cells typically can be operated at fairly low voltages (roughly 5 to 10 times less than an equivalent Kerr cell) and the response time is shorter than for Kerr cells. The response time of KDP is less than 10 ps, and such a modulator can work up to about 25 GHz. In LiNbO₃ electro-optic modulation up to 50 GHz has been demonstrated [92Dol]. High-speed modulators above 40 GHz for telecommunication applications in combination with semiconductor lasers today are mostly based on electro-absorption effect in some special semiconductor materials [02Tak]. The following presentation corresponds widely to examples given in [84Yar, 95Mal].

7.1.3.1.1 Phase modulation

A light wave can be phase-modulated, without change in polarization or intensity, using a suitable cut of an electro-optic crystal. One of two possibilities is required for a crystal and its orientation:

- 1. A crystal having principal axes that will not rotate with applied voltage, but change uniformly. An example is $LiNbO_3$ with applied field along its optical z-axis. This solution is suitable for randomly polarized laser beams.
- 2. A crystal having a characteristic plane perpendicular to the direction of propagation where the plane rotates under the influence of an applied field. Then the input wave must be polarized along one of the new principal axes x' or y' and will not alter this polarization during modulation. An example is KDP with electric field applied along its optical z-axis. Such a configuration is shown in Fig. 7.1.1.

According to Fig. 7.1.1 with a polarizer along x^\prime the optical wave at the output of the crystal exhibits a phase shift

$$\Delta \Phi = \frac{2\pi}{\lambda} \cdot \Delta n_{x'} \cdot L . \tag{7.1.6}$$



Fig. 7.1.1. Longitudinal phase modulator.

The electrically induced change in refractive index is $\Delta n_{x'} \approx -\frac{1}{2}n_{x'}^3 rE$, where r is the corresponding electro-optic coefficient.

For a *longitudinal* modulator with applied electrical field E = V/L the induced phase shift is (apart of sign)

$$\Delta \Phi = \frac{\pi}{\lambda} \cdot n_{x'}^3 r V , \qquad (7.1.7a)$$

which is independent of L and linearly proportional to V.

For a transverse modulator with applied electric field E = V/d the induced phase shift is

$$\Delta \Phi = \frac{\pi}{\lambda} \cdot n_{x'}^3 r V \cdot \frac{L}{d} , \qquad (7.1.7b)$$

which is a function of the aspect ratio L/d.

The half-wave voltage producing $\Delta \Phi = \pi$ for a longitudinal modulator is

$$V_{\pi} = \frac{\lambda}{n_{x'}^3 r} \tag{7.1.8a}$$

and for a transverse modulator

$$V_{\pi} = \frac{\lambda}{n_{x'}^3 r} \cdot \frac{d}{L} \,. \tag{7.1.8b}$$

If the applied modulation voltage and corresponding electric field is sinusoidal in time the phase at the output changes accordingly, as

$$\Phi = \frac{2\pi}{\lambda} \left(n_{x'} - \frac{1}{2} n_{x'}^3 r E_{\rm m} \cdot \sin \omega_{\rm m} t \right) \cdot L = \frac{2\pi}{\lambda} \cdot n_{x'} L - \delta \cdot \sin \omega_{\rm m} t .$$
(7.1.9)

Thus, the optical field is phase-modulated with phase-modulation index

$$\delta = \frac{\pi}{\lambda} \cdot n_{x'}^3 r E_{\rm m} L = \pi \cdot \frac{V_{\rm m}}{V_{\pi}} . \tag{7.1.10}$$

Neglecting a constant phase term, the output wave can be developed into a Bessel function series $E_0(t) = E_i \sum_n J_n(\delta) \cos(\omega + n\omega_m)t$ which consists of components at frequency ω and higher harmonic frequencies $(\omega + n\omega_m)$, $n = \pm 1, \pm 2, \ldots$ The distribution of energy into the sidebands is a function of modulation index δ .

7.1.3.1.2 Polarization modulation (dynamic retardation)

Polarization modulation involves the coherent addition of two orthogonal waves, resulting in a change of the input polarization state at the output. A suitable configuration is shown in Fig. 7.1.2. The crystal and applied voltage are configured to produce dynamically fast and slow axes in the crystal cross-section. The polarizer is positioned such that the input light is decomposed equally into the two orthogonal linear eigenpolarizations along these axes.



Fig. 7.1.2. Longitudinal polarization modulator with input polarizer oriented along the x principal axis at 45° with respect to the perturbed axes x'and y'.

If the light propagating along the z principal axis is polarized along the x- (or y-) axis the corresponding refractive indices of the fast and slow axes x' and y', respectively, are

$$n_{x'} \approx n_x - \frac{1}{2}n_x^3 r_x E$$
 and $n_{y'} \approx n_y + \frac{1}{2}n_y^3 r_y E$

where n_x and n_y are the unperturbed indices of refraction and r_x , r_y are the corresponding electrooptic coefficients for the used material and orientation of the applied field. The phase difference or retardation is

$$\Gamma = \frac{2\pi}{\lambda} \cdot \left(n_{y'} - n_{x'}\right) \cdot L = \frac{2\pi}{\lambda} \cdot \left(n_y - n_x\right) \cdot L + \frac{\pi}{\lambda} \cdot \left(n_y^3 r_y + n_x^3 r_x\right) \cdot EL = \Gamma_0 + \Gamma_i \,. \tag{7.1.11}$$

 Γ_0 is the natural retardation without applied field and Γ_i is the field-induced part of the retardation. In general, the output wave is elliptically polarized. Special points are $\Gamma = \pi/2$, where the electrical field vector is circularly polarized, and $\Gamma = \pi$, where the wave is again linearly polarized, but rotated by 90° to its input direction of polarization. The half-wave voltage in this more general case of a longitudinal modulator is $V_{\pi} = \lambda/(n_y^3 r_y + n_x^3 r_x)$, which is proportional to the wavelength λ and inversely proportional to the relevant electro-optic coefficients. To cancel out an occurring natural birefringence, the phase retardation Γ_0 can be made a multiple of 2π by slightly polishing the crystal to adjust its length, or by introducing a variable compensator, or more practical by applying a bias voltage.

7.1.3.1.3 Amplitude modulation

The intensity of a light wave can be modulated in several ways. Some possibilities are:

- 1. including a dynamic retarder configuration with either crossed or parallel polarizer at the output, or
- 2. using a phase modulator configuration in one branch of a Mach–Zehnder interferometer, or
- 3. choosing a dynamic retarder with push–pull electrodes.

An intensity modulator constructed using a dynamic retarder with crossed polarizers, as shown in Fig. 7.1.3, yields for the transmission $T = I_0/I_i$, the ratio of output to input intensity, the relation

$$T(V) = \sin^2\left(\frac{\Gamma}{2}\right) = \frac{1}{2} \cdot \left[1 - \cos\Gamma\right] = \frac{1}{2} \cdot \left[1 - \cos\left(\Gamma_0 + \frac{\pi V}{V_{\pi}}\right)\right]$$
(7.1.12)

For linear modulation around the 50% transmission point a fixed bias of $\Gamma_0 = \pi/2$ must be introduced, either by placing an additional phase retarder, a quarter-wave plate at the crystal output, or by applying a bias voltage of $V_{\pi}/2$. In case of natural birefringence the values have to be changed accordingly.



Fig. 7.1.3. Longitudinal intensity modulator using crossed polarizers and a quarter-wave plate as a bias to produce linear modulation.

For sinusoidal modulation of V the retardation at the output including bias is $\Gamma = \pi/2 + \Gamma_{\rm m} \sin \omega_{\rm m} t$, where the *amplitude modulation index* is

$$T_{\rm m} = \pi \cdot \frac{V_{\rm m}}{V_{\pi}} \,. \tag{7.1.13}$$

The transmission in case of $V_{\rm m} \ll 1$ becomes $T(V) = \frac{1}{2}[1 + \Gamma_{\rm m} \sin \omega_{\rm m} t]$, which is linear proportional to the modulation voltage. If the signal is large, then the output intensity becomes distorted, and again higher-order odd harmonics appear.

7.1.3.1.4 Design considerations

Modulators operating in transverse configuration have a half-wave voltage typically in the order of tens of volts compared to about 10,000 V for longitudinal devices [89Gha]. For phase modulation, a crystal orientation is required that would give the maximum change in index of refraction, whereas for amplitude modulation a maximum birefringence must be produced.

For $L \ll 2\pi c/\omega_{\rm m} n$ the transit time $\tau = nL/c$ of light in the crystal is of no relevance for the modulation frequency and the electro-optic crystal can be modeled as a lumped capacitor. As the modulation frequency increases beyond such limit the optical phase no longer follows adiabatically the time-varying refractive index. The result is a reduction in the modulation index parameters, δ for phase modulation and $\Gamma_{\rm m}$ for amplitude modulation, by a factor

$$\sigma = \frac{\sin\left(\frac{1}{2}\omega_{\rm m}\tau\right)}{\left(\frac{1}{2}\omega_{\rm m}\tau\right)} = \operatorname{sinc}\left(\frac{1}{2}\omega_{\rm m}\tau\right) \,. \tag{7.1.14}$$

If $\tau = 2\pi/\omega_{\rm m}$ then the transit time of light is equal to the modulation period and the net retardation is reduced to zero. If, somewhat arbitrarily, one takes the highest useful modulation frequency from $\tau = \pi/\omega_{\rm m}$ it follows $(\nu_{\rm m})_{\rm max} = c/(2Ln)$. E.g. using a KDP crystal (n = 1.5) with a length of L = 2 cm, it yields $(\nu_{\rm m})_{\rm max} = 5$ GHz.

7.1.3.2 Traveling-wave modulator

Applying the modulation voltage as a traveling wave, propagating collinearly with the optical wave, can largely extend the limitation of the transit time on the bandwidth of a modulator. Figure 7.1.4 illustrates a transverse traveling-wave modulator. (The modulation field direction is along the x-axis, whereas both light and traveling wave propagate along z.) The electrode is designed to be part of the driving transmission line in order to eliminate electrode charging time effects on the bandwidth. Therefore, the transit time problem is addressed by adjusting the phase velocity of the modulation signal to be equal to the phase velocity of the optical signal [84Yar, 89Gha, 90Alf].



Fig. 7.1.4. Transverse traveling-wave electro-optic modulator.

A mismatch in the phase velocities of the modulating signal and optical wave will produce a reduction in the modulation index δ or Γ_m by a factor

$$\sigma_{\rm tw} = \operatorname{sinc}\left(qL\right) \,, \tag{7.1.15}$$

where

$$q = \frac{\omega_{\rm m}}{2c} \cdot (n_{\rm m} - n) = \frac{\omega_{\rm m}}{2} \cdot \left(\frac{1}{v_{\rm m}} - \frac{1}{v_0}\right) \ .$$

 $v_{\rm m}$ and $n_{\rm m} = \sqrt{\varepsilon}$ are the phase velocity and index of refraction of the modulating signal, v_0 and n are the according terms for the light beam, respectively. In the case of amplitude modulation this equation holds only if there is no natural birefringence in the cross section of the crystal.

Whereas for low frequencies the modulation indices δ and $\Gamma_{\rm m}$ are linearly proportional to the crystal length L these become a sinusoidal function of L at higher frequencies. For a given mismatch q the maximum modulation index can be achieved for crystal lengths $L = \pi/2q$, $3\pi/2q$, etc. The always occurring mismatch between $n_{\rm m}$ and n produces a walk-off between the optical wave and the modulation wave. The maximum useful modulation frequency is taken to be $(\nu_{\rm m})_{\rm max} = c/[2L(n_{\rm m} - n)]$, showing an increase in the modulation frequency limit or useful crystal length by a factor of $(1 - \frac{n_{\rm m}}{n})^{-1}$ for a traveling-wave modulator.

7.1.3.3 Examples

7.1.3.3.1 Crystal classes

7.1.3.3.1.1 Crystal class $\overline{4}2m$

Widely used modulator crystals such as KH_2PO_4 (KDP), KD_2PO_4 (KD*P), $(NH_4)H_2PO_4$ (ADP), $(NH_4)D_2PO_4$ (AD*P) belong to this crystal class.

These crystals are naturally uniaxial with optical axis along z. The three non-vanishing electrooptic tensor elements are $r_{41} = r_{52}$, r_{63} .

Applying a field along the z-axis gives from (7.1.4)

$$\frac{\left(x^2 + y^2\right)}{n_{\rm o}^2} + \frac{z^2}{n_{\rm e}^2} + 2r_{63}E_z \cdot xy = 1 ,$$

where $n_{\rm o}$ and $n_{\rm e}$ are the ordinary and extraordinary refractive indices, respectively. The new index ellipsoid resulting from rotation of the principal plane (x, y) by 45° is

$$\left(\frac{1}{n_{\rm o}^2} + r_{63}E_z\right) \cdot x'^2 + \left(\frac{1}{n_{\rm o}^2} - r_{63}E_z\right) \cdot y'^2 + \frac{z^2}{n_{\rm e}^2} = 1 \; .$$

The lengths of the major axes of this ellipsoid depend on the applied field. Assuming $r_{63}E_z \ll n_o^{-2}$ gives the refractive indices of the two linearly polarized eigenmodes as $n_{x'} = n_o - \frac{1}{2}n_o^3 r_{63}E_z$ and $n_{y'} = n_o + \frac{1}{2}n_o^3 r_{63}E_z$. The birefringence of this plate is $n_{y'} - n_{x'} = n_o^3 r_{63}E_z$ giving a phase retardation of $\Gamma = \frac{2\pi}{\lambda} \cdot n_o^3 r_{63}V$ and a half-wave voltage of $V_{\pi} = \lambda/(2n_o^3 r_{63})$. E.g. the half-wave voltage for a z-cut KDP plate at $\lambda = 633$ nm is 9.3 kV.

7.1.3.3.1.2 Crystal class 3m

The uniaxial crystals LiNbO₃ or LiTaO₃ belong to this crystal class. The non-vanishing electrooptic tensor elements are $r_{51} = r_{42}$, $r_{22} = -r_{12} = -r_{61}$, $r_{13} = r_{23}$, r_{33} . z-cut: Applying a field along the optical z-axis of the crystal gives from (7.1.4)

$$\left(\frac{1}{n_{\rm o}^2} + r_{13}E_z\right) \cdot x^2 + \left(\frac{1}{n_{\rm o}^2} + r_{13}E_z\right) \cdot y^2 + \left(\frac{1}{n_{\rm e}^2} + r_{33}E_z\right) \cdot z^2 = 1.$$

The principal axes of the new index ellipsoid remain unchanged and the lengths of the new semiaxes are $n_x = n_o - \frac{1}{2}n_o^3 r_{13}E_z$, $n_y = n_o - \frac{1}{2}n_o^3 r_{13}E_z$, and $n_z = n_e - \frac{1}{2}n_e^3 r_{33}E_z$.

Under the influence of an applied electric field parallel to the z-axis (c-axis), the crystal remains uniaxially anisotropic. Light propagating along z will experience the same phase change regardless of its polarization state. Thus, an unpolarized laser can be modulated with such a device. In a longitudinal configuration the voltage-driven phase change is $\Delta \Phi = \frac{\pi}{\lambda} \cdot n_o^3 r_{13} V$ and the half-wave voltage for phase modulation is $V_{\pi} = \lambda/(n_o^3 r_{13})$. E.g. for LiNbO₃ at $\lambda = 633$ nm it results in $V_{\pi} =$ 5.5 kV. Since no phase retardation between any two orthogonal polarized waves is introduced, no amplitude modulation can be achieved. Other z-cut uniaxial electro-optic crystals behave similar, except crystals with $\bar{4}2m$ or $\bar{4}$ symmetry, which become biaxial under the influence of an electric field.

x-cut: However, if the field is oriented along the *z*-axis and a light beam is propagating along the *x*-axis (or equivalently along the *y*-axis), a birefringence of $n_z - n_y = (n_e - n_o) - \frac{1}{2}(n_e^3 r_{33} - n_o^3 r_{13}) E_z$ occurs. This can be used for a transverse modulator.

The phase retardation for light passing through the crystal is thus

$$\Gamma = \frac{2\pi}{\lambda} \cdot (n_{\rm e} - n_{\rm o})L - \frac{\pi}{\lambda} \cdot (n_{\rm e}^3 r_{33} - n_{\rm o}^3 r_{13})V \cdot \frac{L}{d} .$$

For light linearly polarized along z direction the electrically induced phase change is $\Delta \Phi = \frac{\pi}{\lambda} \cdot n_{\rm e}^3 r_{33} V \cdot \frac{L}{d}$. Because of the natural birefringence, an amplitude modulator using these crystals requires a phase compensator. The half-wave voltage in case of arbitrary polarization is

$$V_{\pi} = \frac{d}{L} \cdot \frac{\lambda}{(n_{\mathrm{e}}^3 r_{33} - n_{\mathrm{o}}^3 r_{13})}$$

Thus, the half-wave voltage in a transverse modulator can be largely reduced by reduction of d. It typically lies in the range of 100 V (see Table 7.1.2). Transverse modulators using LiNbO₃ or LiTaO₃ have been demonstrated up to 4 GHz modulation frequency [67Kam].

7.1.3.3.1.3 Crystal class $\overline{4}3m$

Examples of this group are InAs, CuCl, ZnS, GaAs, and CdTe, which are due to the cubic crystal class naturally isotropic. The last two kinds of crystals are used for modulation in the infrared, since they remain transparent beyond 10 μ m. The three non-vanishing electro-optic tensor elements are r_{41} , $r_{52} = r_{41}$, and $r_{63} = r_{41}$.

Longitudinal modulator, *z*-cut: In the presence of an electric field in z direction the index ellipsoid becomes using (7.1.4)

$$\frac{x^2}{n^2} + \frac{y^2}{n^2} + \frac{z^2}{n^2} + 2r_{41}xyE_z = 1 .$$

The principal dielectric axes x and y are rotated around the z-axis by 45°. The new index ellipsoid in principle coordinate system (x', y', z') becomes

$$\left(\frac{1}{n^2} + r_{41}E_z\right) \cdot x'^2 + \left(\frac{1}{n^2} - r_{41}E_z\right) \cdot y'^2 + \frac{z'^2}{n^2} = 1 ,$$

and the principal indices of refraction are $n_{x'} = n - \frac{1}{2}n^3r_{41}E_z$, $n_{y'} = n + \frac{1}{2}n^3r_{41}E_z$ and $n_{z'} = n$. This crystal cut is used in a longitudinal modulator.

In case of a Phase Modulator (PM), the light must be polarized either in x' or y' direction, giving an electrically induced phase change of $\Delta \Phi = \frac{\pi}{\lambda} \cdot n^3 r_{41} V$ resulting in a half-wave voltage of $V_{\pi} = \lambda/(n^3 r_{41})$ (PM).

For Amplitude Modulation (AM) the front polarizer can be aligned along the x-axis so that equal amplitudes of the x' and y' modes are exited. The phase retardation results in $\Gamma = \frac{2\pi}{\lambda} \cdot n^3 r_{41} V$ giving a half-wave voltage of $V_{\pi} = \lambda / [2(n^3 r_{41})]$ (AM).

Transverse modulation: For transverse modulation the field must be applied along a cube diagonal direction and the crystal must be cut appropriately. E.g. in case of a field oriented in $\langle 110 \rangle$ direction and having a magnitude of $E_x = E_y = \frac{1}{\sqrt{2}} \cdot E$ it results in $n_{x'} = n + \frac{1}{2}n^3r_{41}E$, $n_{y'} = n - \frac{1}{2}n^3r_{41}E$ and $n_{z'} = n$. The obtained phase retardation is $\Gamma = \frac{2\pi}{\lambda} \cdot n^3r_{41} \cdot \frac{L}{d} \cdot V$ and the half-wave voltage is given by $V_{\pi} = \frac{d}{L} \cdot \frac{\lambda}{2(n^3r_{41})}$.

Data of phase retardation and electro-optical properties of $\overline{4}3m$ crystals for a field along directions $\langle 001 \rangle$, $\langle 110 \rangle$, or $\langle 111 \rangle$ are found in [61Nam, 84Yar].

7.1.3.3.2 Selected electro-optic materials and modulator systems

In Table 7.1.1 selected electro-optic materials and their properties are presented.

KDP and its isomorphic single crystal KD*P are widely used for electro-optic Q-switches in Nd:YAG-, Nd:YLF-, Ti:sapphire-, and alexandrite lasers. The damage threshold of these crystals of about 10 GW/cm² (< 500 ps) is very high. The longitudinal half-wave voltage at $\lambda = 0.546 \mu m$ is 7.65 kV for KDP and 2.98 kV for KD*P, respectively.

 $LiNbO_3$ is also used for these applications, mostly in transverse modulators. The half-wave voltage is below 1 kV, but the damage threshold is only in the range of 250 MW/cm².

In Table 7.1.2 performances of some commercially available EO modulators with driver electronics are listed. EO modulator systems are available from a number of different suppliers. Among these are in alphabetic order and without claim of completeness: Agere-Systems, Conoptics, Cy-Optics, e2v technologies, Electro-Optical Products Corp., EOSPACE, Fujitsu, Gsaenger, IPAG-Innovative, JDS Uniphase, Lasermetrics, Lucent Technologies, New Focus, Nova Phase, Sciro, Sumitomo, and Quantum Technology. Many others may exist and can be found in the net, e.g. via www.globalspec.com. Devices exemplarily listed in Table 7.1.2 may also be offered by other suppliers with similar or even superior features. Especially, high-speed modulators for telecommunication system applications, as shown for one example in the Table's last row, can operate up to data rates of 40 Gbit/s and beyond at rather low drive voltages. 0.9 V drive voltage at 40 Gbit/s for a LiNbO₃ EO modulator has been demonstrated at Fujitsu Laboratories in 2002.

7.1.3.4 Electro-optic beam deflector

The electro-optic effect is also used to deflect light beams. A simple realization of such a deflector using e.g. a KDP crystal is shown in Fig. 7.1.5 [75Yar]. It consists of two KDP prisms with edges along the x', y', and z directions having their z-axes opposite to one another but are otherwise similarly oriented. The electric field is applied along z and the light propagates in the y' direction with its polarization along x'. The index of refraction "seen" by ray A, which propagates entirely in the upper prism, is given as $n_A = n_0 - \frac{n_o^3}{2} \cdot r_{63}E_z$, while in the lower prism with opposite field direction with respect to the z-axis, ray B "sees" $n_B = n_0 + \frac{n_o^3}{2} \cdot r_{63}E_z$. The deflection angle comes out as

Material	Sym- metry	Wave- length λ [µm]	EO-coefficient r_{ik} $[10^{-12} \text{ m/V}]$		Index of refraction n	Figure of merit $n^3 r$ $[10^{-12} \text{ m/V}]$	Dielectric constant ε_i
CdTe	$\bar{4}3m$	1.0 10.6	(T) $r_{41} = 4.5$ (T) $r_{41} = 6.8$		n = 2.84 n = 2.60	103 120	(S) 9.4
GaAs	$\bar{4}3m$	$1.15 \\ 10.6$	(T) $r_{41} = 1.43$ (T) $r_{41} = 1.51$		n = 3.43 n = 3.3	58 54	(T) 12.3
GaP	$\bar{4}3m$	$0.55 - 1.5 \\ 0.633$	(T) $r_{41} = -1.0$ (S) $r_{41} = -0.97$		n = 3.66 - 3.08 n = 3.32	35	(S) 10
$\beta - ZnS$	$\bar{4}3m$	0.633	(S) $r_{41} = -1.6$		n = 2.35		
ZnSe	$\bar{4}3m$	$0.633 \\ 10.6$	(S) $r_{41} = 2.0$ (T) $r_{41} = 2.2$		n = 2.60 n = 2.39	35 30	(S) 9.1
ZnTe	$\bar{4}3m$	$0.633 \\ 10.6$	(T) $r_{41} = 4.04$ (T) $r_{41} = 3.9$		n = 2.99 n = 2.70	108 77	(S) 10.1
$\begin{array}{l} \mathrm{KDP} \\ \mathrm{(KH_2PO_4)} \end{array}$	$\bar{4}2m$	0.546	(T) $r_{41} = 8.77$ (T) $r_{63} = 10.3$		$n_{\rm o} = 1.5115$ $n_{\rm e} = 1.4698$		(T) $\varepsilon_1 = \varepsilon_2 = 42$ (T) $\varepsilon_3 = 21$
		0.633	(T) $r_{41} = 8.0$ (T) $r_{63} = 11$		$n_{\rm o} = 1.5074$ $n_{\rm e} = 1.4669$		(S) $\varepsilon_1 = \varepsilon_2 = 44$ (S) $\varepsilon_3 = 21$
$\begin{array}{l} \mathrm{KD}^*\mathrm{P} \\ \mathrm{(KD_2\mathrm{PO}_4)} \end{array}$	$\bar{4}2m$	0.546	(T) $r_{41} = 8.8$ (T) $r_{63} = 26.4$		$n_{\rm e} = 1.4683$ $n_{\rm o} = 1.5079$		(S) $\varepsilon_1 = \varepsilon_2 = 58$ (T) $\varepsilon_3 = 50$
		0.633	(T) $r_{63} = 24.1$		$n_{\rm o} = 1.502$ $n_{\rm e} = 1.462$		(S) $\varepsilon_3 = 48$
ADP (NH_4H_2 -	$\bar{4}2m$	0.546	(T) $r_{41} = 23.76$ (T) $r_{63} = 8.56$		$n_{\rm o} = 1.5266$ $n_{\rm e} = 1.4808$		(T) $\varepsilon_1 = \varepsilon_2 = 56$ (T) $\varepsilon_3 = 15$
$PO_4)$		0.633	(T) $r_{41} = 23.41$ (T) $r_{63} = 8.8$		$n_{\rm o} = 1.5220$ $n_{\rm e} = 1.4773$		(S) $\varepsilon_1 = \varepsilon_2 = 58$ (S) $\varepsilon_3 = 14$
$LiNbO_3$	3m	0.633	(T) $r_{13} = 9.6$ (T) $r_{22} = 6.8$ (T) $r_{33} = 30.9$ (T) $r_{51} = 32.6$	(S) $r_{13} = 8.6$ (S) $r_{22} = 3.4$ (S) $r_{33} = 30.8$ (S) $r_{51} = 28$	$n_{\rm o} = 2.286$ $n_{\rm e} = 2.200$		$\begin{array}{l} (\mathrm{T}) \ \varepsilon_1 = \varepsilon_2 = 78 \\ (\mathrm{T}) \ \varepsilon_3 = 32 \\ (\mathrm{S}) \ \varepsilon_1 = \varepsilon_2 = 43 \\ (\mathrm{S}) \ \varepsilon_3 = 28 \end{array}$
LiTaO ₃	3m	0.633	(T) $r_{13} = 8.4$ (T) $r_{22} = -0.2$ (T) $r_{33} = 30.5$	(S) $r_{13} = 7.5$ (S) $r_{22} = 1$ (S) $r_{33} = 33$ (S) $r_{51} = 20$	$n_{\rm o} = 2.176$ $n_{\rm e} = 2.180$		$\begin{array}{l} (T) \ \varepsilon_1 = \varepsilon_2 = 51 \\ (T) \ \varepsilon_3 = 45 \\ (S) \ \varepsilon_1 = \varepsilon_2 = 41 \\ (S) \ \varepsilon_3 = 43 \end{array}$

Table 7.1.1. Selected Electro-Optic (EO) materials and their properties [79Coo, 84Yar, 86Kam].

(T): low frequency from dc through audio range; (S): high frequency.

$$\theta = \frac{L \cdot (n_B - n_A)}{d} = \frac{L}{d} \cdot n_o^3 r_{63} E_z .$$
(7.1.16)

To get the number of resolvable spots, N, it is assumed that the crystal is placed at the waist of a Gaussian beam with spot diameter \bar{d} . The far-field beam divergence is $\theta_{\text{beam}} = 2\lambda/(\pi n \bar{d})$. Taking $d = \bar{d}$ and $n \approx 1$ (in air) the maximum number of resolvable spots is

$$N = \frac{\theta}{\theta_{\text{beam}}} = \frac{\pi}{2\lambda} \cdot Ln_o^3 r_{63} E_z .$$
(7.1.17)

An electric field that induces a birefringent retardation (in the distance L) of $\Delta \Gamma = \pi$ will yield N = 1.

Supplier	Crystal	Spectral range $\Delta \lambda$ [nm]	Aperture diameter [mm]	Contrast at wavelength	Half-wave voltage V_{π} at wavelength	Operating frequency (-3 dB-limit)
Conoptics	ADP	300750	3.5	500:1 at 633 nm	263 V at 500 nm	1125 MHz
Conoptics	KD*P	3001100	3.1	700:1 at 1064 nm	482 V at 1064 nm	1200 MHz
Quantum Technology	$LiTaO_3$	8002500	1	100:1 at 633 nm	107 V at 633 nm	DC1 GHz
New Focus	MgO:LiNbO ₃	5001600	1	> 100:1 at 1 µm	4579 V at 1 μm	2.04.6 GHz
New Focus	MgO:LiNbO ₃	5001600	$1\times 2~\mathrm{mm}^2$	> 100:1 at 1 µm	79 V at 1 μm	6.8 or 9.2 GHz
Fujitsu	$\rm LiNbO_3$	15301570	fiber-coupled	> 100:1 at 1550 nm	5 V at 1550 nm	DC 30 GHz

Table 7.1.2. Performances of some commercially available EO modulators with driver electronics.



Fig. 7.1.5. Double-prism electro-optic beam deflector.

7.1.4 Kerr electro-optic effect modulators

The first ever investigated electro-optic effect is named after the Scottish physicist John Kerr (1824–1907) who discovered it in the year 1875. He found that an isotropic transparent substance becomes birefringent when placed in an electric field. The quadratic electro-optic effect is a higher-order effect and is normally neglected when the linear electro-optic effect is present. Unlike the linear effect, it exists in a medium with any symmetry. An electric field, in general, changes the dimension and orientation of the index ellipsoid. This change is dependent upon the direction of the applied electric field as well as the 6×6 matrix elements s_{ij} , see text below (7.1.2).

The equation of the index ellipsoid in the presence of an electric field can be written

$$\left(\frac{1}{n_x^2} + s_{1j}\xi_j^2\right)x^2 + \left(\frac{1}{n_y^2} + s_{2j}\xi_j^2\right)y^2 + \left(\frac{1}{n_z^2} + s_{3j}\xi_j^2\right)z^2 + 2yzs_{4j}\xi_j^2 + 2zxs_{5j}\xi_j^2 + 2xys_{6j}\xi_j^2 = 1,$$
(7.1.18)

where $\xi_1^2, \ldots, \xi_6^2 = E_x^2, E_y^2, E_z^2, 2E_yE_z, 2E_zE_x, 2E_xE_y$.

A table containing the form of the quadratic electro-optic coefficients is given in [67Nye, 84Yar]. Extensive tables of quadratic electro-optic coefficients are listed in [79Coo].

7.1.4.1 Kerr effect in isotropic media

An optically isotropic medium placed in a static electric field becomes birefringent. This effect is associated mostly with the alignment of the molecules in the presence of the field. The medium then gains optically uniaxial anisotropy where the electric field defines the optical axis. Choosing an electric field E along the z-axis of an isotropic crystal the index ellipsoid becomes

$$\left(\frac{1}{n^2} + s_{12}E_z^2\right)x^2 + \left(\frac{1}{n^2} + s_{12}E_z^2\right)y^2 + \left(\frac{1}{n^2} + s_{11}E_z^2\right)z^2 = 1, \qquad (7.1.19)$$

where $s_{13} = s_{23} = s_{12}$ and $s_{33} = s_{11}$ have been used. Further, for isotropic media $s_{44} = s_{55} = s_{66} = \frac{1}{2}(s_{11} - s_{12})$. This index ellipsoid can be written in a more compact form as

$$\frac{\left(x^2 + y^2\right)}{n_o^2} + \frac{z^2}{n_e^2} = 1 \tag{7.1.20}$$

using $n_o = n - \frac{1}{2}n^3 s_{12}E^2$ and $n_e = n - \frac{1}{2}n^3 s_{11}E^2$. The birefringence becomes $n_e - n_o = \frac{1}{2}n^3(s_{12} - s_{11})E^2 = -n^3 s_{44}E^2$. According to empirical results the Kerr birefringence is often written as $n_e - n_o = K\lambda E^2$, where K is the so-called Kerr constant and λ is the vacuum wavelength. Kerr constant and quadratic electro-optic coefficients in isotropic media are, therefore, related by $s_{44} = -K\lambda/n^3$. Table 7.1.3 lists Kerr constants of several liquids used in Kerr modulator devices. Table 7.1.4 lists some selected quadratic electro-optic coefficients.

Substance Wavelength Index of Kerr constant $\lambda \ [\mu m]$ refraction n $K \, [\mathrm{m/V^2}]$ 4.9×10^{-15} 0.546Benzene $\mathrm{C}_{6}\mathrm{H}_{6}$ 1.503 4.14×10^{-15} 0.633 1.496 3.88×10^{-14} Carbon disulfide CS_2 0.5461.633 3.18×10^{-14} 0.6331.619 2.83×10^{-14} 0.6941.612 1.84×10^{-14} 1.0001.596 1.11×10^{-14} 1.6001.582 -3.5×10^{-14} Chloroform $CHCl_3$ 0.589 5.1×10^{-14} Water H_2O 0.589 1.37×10^{-12} Nitrotoluene $C_5H_7NO_2$ 0.589 2.44×10^{-12} $C_6H_5NO_2$ Nitrobenzene 0.589

Table 7.1.3. Kerr constants for some selected liquids at 20°C [84Yar].

Table 7.1.4. Some selected quadratic electro-optic coefficients [79Coo].

Substance	Symmetry	Wavelength $\lambda \; [\mu m]$	Electro-optic coefficients $n_o^3 s_{ij} \ [10^{-18} \text{ m}^2/\text{V}^2]$	Index of refraction n
BaTiO ₃	m3m	0.500	$n_{o}^{3}(s_{11} - s_{12}) = 72,000$ $n_{o}^{3}s_{44} = 44,000$ near $T_{c} = 120^{\circ}C$	$n_{\rm o} = 2.42$
${\rm KTa}_{0.65}{\rm Nb}_{0.35}{\rm O}_3$	m3m	0.633	$n_{\rm o}^3(s_{11} - s_{12}) = 34,700$ at 20°C	$n_{\rm o} = 2.29$
$\mathrm{Pb}_{0.93}\mathrm{La}_{0.07}(\mathrm{Zr}_{0.65}\mathrm{Ti}_{0.35})0_3$	∞m	0.550	$n_{\rm o}^3(s_{33} - s_{13}) = 26,000$ near $T_{\rm c} = 63^{\circ}{\rm C}$	$n_{\rm o} = 2.450$

A Kerr modulator consists of a glass cell containing two transverse electrodes that is filled with a polar liquid and placed between crossed linear polarizers whose transmission axes are at $\pm 45^{\circ}$ to the applied electric field. The great value of such a device lies in the fact that it can respond effectively to frequencies up to 10 GHz. Kerr cells, usually containing nitrobenzene or carbon disulfide, have for many years been used as Q-switches in pulsed laser systems. A cell of length L and electrode distance d gives a retardation of

$$\Gamma = 2\pi \cdot KL \cdot \frac{V^2}{d^2} \,. \tag{7.1.21}$$

A typical nitrobenzene cell where d = 1 cm and L is several cm will require a voltage in the order of 3×10^4 V to respond as a half-wave plate. A drawback is that nitrobenzene is both poisonous and explosive. Therefore, transparent solids like the crystals KTN (KTa_{0.65}Nb_{0.35}O₃), barium titanate (BaTiO₃), or PLZT (Pb_{0.93}La_{0.07}(Zr_{0.65}Ti_{0.35})O₃) e.g. are of interest for such modulators.

7.1.5 Acousto-optic modulators

An acoustic wave propagating in an optically transparent medium produces a periodic modulation of the index of refraction via the elasto-optic effect. This provides a moving phase grating which may diffract portions of an incident light into one or more directions. This phenomenon, known as the acousto-optic diffraction, has led to a variety of optical devices that perform spatial, temporal, and spectral modulation of light. Acousto-optic devices are used in laser applications for electronic control of the intensity or position of a laser beam.

Acoustic waves in solids can appear as longitudinal or transverse (shear) waves, whereas in liquids and gases only longitudinal waves are possible. The acousto-optic interaction is very similar to that of electro-optic modulation, except that in the acousto-optic interaction an RF field is required in any case. A number of different useful materials, such as SiO₂, TeO₂, LiNbO₃, PbMoO₄, have been exploited for fabrication of acousto-optic modulator devices.

7.1.5.1 The photoelastic effect

The photoelastic effect in a material causes coupling of mechanical strain to the optical index of refraction. This effect occurs in all states of matter and is commonly described by the change in the optical impermeability tensor. Using contracted indices the relation is [86Got]

$$\Delta\left(\frac{1}{n^2}\right)_i = p_{ij}S_j , \quad i,j = 1, 2, \dots, 6$$

where p_{ij} are the strain-optic coefficients and S_j are the strain components.

The equation of the index ellipsoid in the presence of strain can then be written as

$$\left(\frac{1}{n_x^2} + p_{1j}S_j\right)x^2 + \left(\frac{1}{n_y^2} + p_{2j}S_j\right)y^2 + \left(\frac{1}{n_z^2} + p_{3j}S_j\right)z^2 + 2yzp_{4j}S_j + 2zxp_{5j}S_j + 2xyp_{6j}S_j = 1 ,$$
(7.1.22)

where n_x , n_y , n_z are the principal indices of refraction and p_{ij} are defined in the principal coordinate system. Again summation over repeated indices is assumed.

As an example a longitudinal sound wave is considered propagating along the z direction of an isotropic medium (e.g. water). For a sinusoidal particle displacement $\boldsymbol{u}(z,t) = A \, \bar{\boldsymbol{z}} \cos(\Omega t - Kz)$ with amplitude A, sound frequency Ω , and wave number $K = 2\pi/\Lambda$ the associated strain field is

$$S_3 = KA \cdot \sin\left(\Omega t - Kz\right) = S \cdot \sin\left(\Omega t - Kz\right) . \tag{7.1.23}$$

For an isotropic medium, the relevant elasto-optic coefficients are [84Yar]: $p_{13} = p_{23} = p_{12}$, $p_{33} = p_{11}$, and $p_{43} = p_{53} = p_{63} = 0$. The new index ellipsoid results in

$$\left(\frac{1}{n^2} + p_{12}S \cdot \sin\left(\Omega t - Kz\right)\right) x^2 + \left(\frac{1}{n^2} + p_{12}S \cdot \sin\left(\Omega t - Kz\right)\right) y^2 + \left(\frac{1}{n^2} + p_{11}S \cdot \sin\left(\Omega t - Kz\right)\right) z^2 = 1.$$
(7.1.24)

Here no mixed terms are involved and the principal axes remain unchanged. The new principal indices of refraction are

$$n_x = n - \frac{1}{2}n^3 p_{12}S \cdot \sin\left(\Omega t - Kz\right) ,$$

$$n_y = n - \frac{1}{2}n^3 p_{12}S \cdot \sin\left(\Omega t - Kz\right) ,$$

$$n_z = n - \frac{1}{2}n^3 p_{11}S \cdot \sin\left(\Omega t - Kz\right) .$$

The medium now carries a volume-index (phase) grating with grating constant $K = 2\pi/\Lambda$ that travels at speed $v = \Omega/K$.

Another example may be a sound wave along the $\langle 001 \rangle$ direction (z-axis) in a cubic crystal of symmetry m3m (e.g. germanium). Considering a shear wave polarized in $\langle 010 \rangle$ direction (y-axis) with particle displacement $\boldsymbol{u}(z,t) = A \, \bar{\boldsymbol{y}} \cos(\Omega t - Kz)$, the strain field associated with this shear comes out to be

$$S_4 = KA \cdot \sin\left(\Omega t - Kz\right) = S \cdot \sin\left(\Omega t - Kz\right) . \tag{7.1.25}$$

According to the relevant elasto-optic coefficients for the point group symmetry m3m $p_{14} = p_{24} = p_{34} = p_{54} = p_{64} = 0$, $p_{44} \neq 0$ the index ellipsoid thus becomes [84Yar]

$$\frac{x^2 + y^2 + z^2}{n^2} + 2yz \cdot p_{44}S \cdot \sin\left(\Omega t - Kz\right) = 1.$$
(7.1.26)

The new principal axes are obtained by rotating the coordinates around the x-axis by 45° , and the new principal indices of refraction are given by

$$\begin{split} n_{x'} &= n \;, \\ n_{y'} &= n - \frac{1}{2} n^3 p_{44} S \, \sin(\Omega t - K z) \;, \\ n_{z'} &= n + \frac{1}{2} n^3 p_{44} S \, \sin(\Omega t - K z) \;. \end{split}$$

As before, a moving optical volume-index grating is excited from the shear acoustic wave via the strain-optic effect.

The forms of the elasto-optic coefficients are identical to those for the quadratic electro-optic effect [67Nye, 84Yar]. Tables of elasto-optic coefficients p of various materials are found e.g. in [72Pin, 79Hel, 86Got].

7.1.5.2 Interaction regimes

The quantity of the factor $Q = (2\pi\lambda L)/(n\Lambda^2)$ determines the interaction of light with sound waves. Herein λ is the vacuum wavelength of light, n is the refractive index of the medium, L is the distance the light travels through the acoustic wave, and Λ is the acoustic wavelength [67Kle].

7.1.5.2.1 Raman–Nath regime

 $Q \ll 1$: The approximation holds for $L \ll (n\Lambda^2)/(2\pi\lambda)$. The laser beam may be incident roughly normal to the acoustic beam. Then several diffraction orders appear distinguished by the numbers $-m, \ldots, -2, -1, 0, 1, 2, \ldots, m$ at diffraction angels according to $\sin \theta_m = m\lambda/n\Lambda$ with corresponding frequencies $\omega - m\Omega, \ldots, \omega - 2\Omega, \omega - \Omega, \omega, \omega + \Omega, \omega + 2\Omega, \ldots, \omega + m\Omega$, see Fig. 7.1.6.



Fig. 7.1.6. Raman–Nath diffraction of light into multiple orders.

The intensities are given by Bessel functions J_m . The diffraction efficiency of the *m*th-order Raman–Nath diffraction is thus given by

$$\eta = J_m^2 \left(\beta\right) = J_m^2 \left(kL\Delta n\right) , \qquad (7.1.27)$$

where Δn is the change in refractive index caused by the acoustic wave. The diffraction efficiency of order ± 1 is maximum when the modulation index is $\beta = 1.85$. The zeroth order is completely quenched when $\beta = 2.4$, because of $J_0(2.4) = 0$. For small values of β it results in [86Got]

$$\frac{I_1}{I_0} = \beta^2 = \frac{\pi^2}{2} \cdot \left(\frac{L}{\lambda}\right)^2 M_2 I_a , \quad \text{where} \quad M_2 = \frac{n^6 p^2}{\rho v^3} .$$
(7.1.28)

Here ρ is the mass density and $I_{\rm a}$ is the acoustic intensity. If the RF sound carrier is modulated with an information-bearing signal, the modulation of the diffracted light will be reasonably linear provided β is less than 2.4. The main disadvantage of operating in the Raman–Nath regime is the small interaction length and the excessive acoustic power required.

Therefore, most acousto-optic devices operate in the Bragg regime; the common exception being acousto-optic mode lockers.

7.1.5.2.2 Bragg regime

 $Q \gg 1$: The approximation holds for $L \gg (n\Lambda^2)/(2\pi\lambda)$. The physical properties can be treated, for instance, by coupled-mode theory. It shows up that at one particular angle of incidence $\theta_{\rm B}$ of a laser beam only one diffraction order is produced by constructive interference in the interaction region, possible other orders are annihilated by destructive interference. This can be understood as a resonant reflection of light from the acoustic wave. A convenient description can be given for light and sound waves as colliding photons and phonons governed by the laws of conservation of energy and momentum. Momentum conservation requires

$$k_{\rm d} = k_{\rm i} + K$$
, (7.1.29)

where

 $k_{\rm i} = 2\pi n/\lambda$: wave vector of incident light beam, $k_{\rm d} = 2\pi n/\lambda$: wave vector of diffracted light beam, $K = 2\pi/\Lambda = 2\pi f/v$: wave vector of acoustic wave.

Conservation of energy takes the form $\omega_{\rm d} = \omega_{\rm i} \pm \Omega$.

Different materials and a variety of configurations can be used for modulation. These can be described by terms such as longitudinal- and shear-mode, isotropic and anisotropic. While these all rely on the basic principles of momentum and energy conservation, different modes of operation have very different performances.

7.1.5.2.2.1 Isotropic interaction

In an isotropic interaction incident and diffracted light beams see the same refractive index in the crystal. There is no change in polarization associated with the interaction. These interactions usually occur in liquids, in homogenous crystals, or in birefringent crystals suitably cut, see Fig. 7.1.7.



Fig. 7.1.7. Bragg diffraction in an isotropic medium.

If θ is the angle between the incident or diffracted beam and the acoustic wavefront, the momentum conservation yields for Bragg condition

$$K = 2k \cdot \sin \theta_{\rm B} \quad \text{or} \quad 2\Lambda \cdot \sin \theta_{\rm B} = \frac{\lambda}{n} .$$
 (7.1.30)

The separation angle between first and zeroth order is twice the angle of incidence and, therefore, twice the Bragg angle. In practice, the Bragg angle is often small and can be written as

$$\theta_{\rm B} = \frac{\lambda}{2n\Lambda} = \frac{\lambda f}{2nv} \,. \tag{7.1.31}$$

7.1.5.2.2.2 Anisotropic interaction

The refractive index in an optically anisotropic medium such as a birefringent crystal depends on the direction as well as the polarization of the light beam. In general, the refractive indices of the incident and diffracted light beams are different.



Fig. 7.1.8. Bragg diffraction in an anisotropic medium.

From the wave vector diagram shown in Fig. 7.1.8 one obtains [95Cha]

$$\sin \theta_{\rm i} = \frac{\lambda}{2n_{\rm i}\Lambda} \cdot \left[1 + \frac{\Lambda^2}{\lambda^2} \left(n_{\rm i}^2 - n_{\rm d}^2 \right) \right] ,$$

$$\sin \theta_{\rm d} = \frac{\lambda}{2n_{\rm d}\Lambda} \cdot \left[1 - \frac{\Lambda^2}{\lambda^2} \left(n_{\rm i}^2 - n_{\rm d}^2 \right) \right] .$$
(7.1.32)

Notice that the first term on the right-hand side of these equations is the same as for the isotropic case, while the remaining terms denote the modification due to the effect of anisotropy. This significantly changes the angle/frequency characteristics of acousto-optic diffraction.

As an example, Bragg diffraction in a positive uniaxial crystal (e.g. LiNbO₃) is considered. An incident extraordinary wave polarized parallel to the *c*-axis with refractive index n_e is diffracted into an ordinary wave polarized perpendicular to the *c*-axis with refractive index n_o . Then, θ_i and θ_d are both functions of λ/Λ . Bragg diffraction is only possible when $|n_o - n_e| \leq \lambda/\Lambda \leq |n_o + n_e|$. In the small angle approximation, the separation angle between first and zeroth order is nearly the same as in the case of isotropic diffraction.

For acousto-optic modulators and acousto-optic deflectors, optical birefringence in principle is not necessary. However, it is a requirement for materials used in acousto-optic tunable filters.

7.1.5.2.2.3 Efficiency

Using coupled-mode theory the fraction of power of an incident light beam transferred into the diffracted beam after traveling a distance L is for validity of the Bragg condition

$$\frac{I_{\rm d}}{I_{\rm i}} = \sin^2\left(\kappa L\right) \,,\tag{7.1.33}$$

where κ is a coupling constant. A slight deviation from the Bragg condition caused by a wave vector mismatch Δk results in

$$\frac{I_{\rm d}}{I_{\rm i}} = (\kappa L)^2 \cdot {\rm sinc}^2 \left[(\kappa L)^2 + \left(\frac{L\Delta k}{2}\right)^2 \right]^{\frac{1}{2}} , \qquad (7.1.34)$$

 $\begin{array}{c} {\rm Landolt\text{-}B\" ornstein} \\ {\rm New \ Series \ VIII/1A2} \end{array}$
where $\operatorname{sin}(x) = \operatorname{sin}(x)/x$. The mismatch Δk can be caused either by misalignment of the laser beam or by distortion of the acoustic wavefront due to the finite size of the transducer [84Yar].

The coupling constant κ can be expressed in terms of the strain components and the elastooptic coefficients of the material. The efficiency for Bragg diffraction according to (7.1.33) results in

$$\eta = \frac{I_{\rm d}}{I_{\rm i}} = \sin^2 \left(\frac{\pi n^3}{2\lambda_0 \cos \theta_{\rm B}} \cdot pSL \right) , \qquad (7.1.35)$$

where pS are the matrix elements $p_{ij}S_j$. The strain acoustic amplitude S is related to the acoustic intensity I_a by $S = [2I_a/(\rho v^3)]^{\frac{1}{2}}$, where ρ is the mass density and v is the velocity of sound in the crystal.

A homogeneously excited transducer of height H, width L, and acoustic power $P_{\rm a}$ yields $I_{\rm a} = P_{\rm a}/(HL)$. Combining these relations it gives

$$\eta = \frac{I_{\rm d}}{I_{\rm i}} = \sin^2 \left\{ \frac{\pi}{\sqrt{2}\,\lambda_0 \cos\theta_{\rm B}} \cdot \left[M_2 P_{\rm a} \cdot \left(\frac{L}{H}\right) \right]^{\frac{1}{2}} \right\} \,, \tag{7.1.36}$$

where $M_2 = n^6 p^2 / (\rho v^3)$ is a figure of merit for judging the usefulness of a material with respect to efficiency.

For small acoustic power levels, the diffraction efficiency is thus linearly proportional to the acoustic power

$$\eta = \frac{\pi^2}{2\lambda_0^2 \cos^2 \theta_{\rm B}} \cdot \left[M_2 P_{\rm a} \cdot \left(\frac{L}{H} \right) \right] , \qquad (7.1.37)$$

which is the basis of an acousto-optic modulator. This approximation is valid when the peak efficiency is below 70 %. The acoustic power required for 100 % modulation (i.e. total conversion of the incident light) is given by

$$P_{\rm a} = \frac{\lambda_0^2 \cos^2 \theta_{\rm B}}{2M_2} \cdot \left(\frac{H}{L}\right) \,. \tag{7.1.38}$$

Therefore, according to (7.1.37) a small aspect ratio (H/L) is desirable for an efficient operation of a modulator.

7.1.5.2.2.4 Bandwidth

A modulated acoustic wave is defined by its center frequency f_0 and modulation bandwidth Δf . The attainable modulation bandwidth Δf of an acousto-optic modulator results from the angular spread of the beams of light and sound. For finite beams of light and sound in a modulator a relation of Bragg angle change $\Delta \theta$ and acoustic frequency change Δf comes out as

$$\Delta f = \frac{2nv \cdot \cos\theta}{\lambda} \cdot \Delta\theta \,. \tag{7.1.39}$$

The angle of incidence covers a range of $\Delta \theta = \delta \theta_{o} + \delta \theta_{a}$, where $\delta \theta_{o}$ and $\delta \theta_{a}$ are the angular spread of light wave vector and sound wave vector, respectively. The diffracted light beam for each fixed angle of incident light has an angular spread into sidebands of $2\delta \theta_{a}$, see Fig. 7.1.9. Each direction corresponds to a different frequency shift. In order to recover best the intensity modulation of the diffracted light beam, mixing the spectrally shifted components collinearly in a square-law detector is necessary, for optimum condition $\delta \theta_{o} \approx \delta \theta_{a} = \frac{1}{2}\Delta \theta$.



Fig. 7.1.9. Diffraction geometry of a Bragg acoustooptic modulator.

In most practical cases the incident laser beam is a focused Gaussian beam with beam waist diameter \bar{d} . The corresponding optical beam divergence is $\Delta \theta = 4\lambda/(\pi n \bar{d})$. Then, taking $\delta \theta = \frac{1}{2}\Delta \theta$ it yields

$$\left(\Delta f\right)_{\rm m} = \frac{1}{2}\Delta f = \frac{2}{\pi} \cdot \frac{v}{\bar{d}} \cdot \cos\theta \approx \frac{1}{\tau} , \qquad (7.1.40)$$

where $\tau = \bar{d}/v$ is the acoustic transit time across the optical aperture [95Cha]. Thus, the modulation bandwidth is roughly equal to the reciprocal of the acoustic transit time across the optical beam. The maximum fractional bandwidth of an acousto-optic modulator is usually determined by the condition that the diffracted beam does not interfere with the undiffracted beam, i.e. $\Delta \theta < \theta_{\rm B}$. Then, one obtains from (7.1.39) and the relation $\theta_{\rm B} = \lambda f/2nv$ (7.1.31), and when assuming $\Delta \theta \approx \Delta \theta_{\rm B}$ for the maximum modulation bandwidth

$$\frac{(\Delta f)_{\rm m}}{f} \approx \frac{\Delta f}{2f} \le \frac{1}{2} \,. \tag{7.1.41}$$

Thus, the maximum modulation bandwidth is approximately one-half of the acoustic frequency. Therefore, large modulation bandwidths are only available with high-frequency Bragg diffraction.

A useful figure of merit [66Gor], including the modulator bandwidth Δf and the center acoustic frequency f_0 and being independent of the modulator width L, is:

$$2\eta f_0 \Delta f = M_1 \cdot \frac{2\pi^2}{\lambda_0^3 \cos \theta_{\rm B}} \cdot \left(\frac{P_{\rm a}}{H}\right) \quad \text{with} \quad M_1 = \frac{n^7 p^2}{\rho v} . \tag{7.1.42}$$

Another quantity [67Dix], which is independent of the acoustic and optical beam dimensions, is

$$\eta f_0 = M_3 \cdot \frac{\pi^2}{2\lambda_0^3 \cos^2 \theta_{\rm B}} \cdot P_{\rm a} \quad \text{with} \quad M_3 = \frac{n^7 p^2}{\rho v^2} \,.$$
 (7.1.43)

Values of M_1 , M_2 , and M_3 for a number of materials are listed in Table 7.1.5.

Another important material parameter is the acoustic attenuation that limits the center frequency, bandwidth, and useful aperture of acousto-optic devices. Following the theory the dominant contribution to acoustic attenuation in crystals is caused by relaxation of the thermal phonon distribution toward equilibrium. A widely used result of this theory is the relation derived by Woodruff and Ehrenreich [61Woo], which shows that the acoustic attenuation increases quadratically with frequency. Commonly the acoustic attenuation is given in units of dB per microsecond of acoustic propagation time and GHz^2 , i.e. $\text{dB}/(\mu s \cdot \text{GHz}^2)$, see Table 7.1.5.

Table 7.1.5 lists relevant properties of selected acousto-optic materials. The listed figures of merit are normalized relative to that of fused silica which has the following absolute values: $M_1 = 7.83 \times 10^{-8} \text{ m}^2 \text{sKg}^{-1}, M_2 = 1.51 \times 10^{-15} \text{ s}^3 \text{Kg}^{-1}, M_3 = 1.3 \times 10^{-13} \text{ m}^2 \text{s}^2 \text{Kg}^{-1}$.

Ref.	p.	110]
------	----	------

7.1 Modulators

Table 7.1.5.	Selected acousto-o	ptic materia	ls [72Kor, 86G	lot, 95Cha].						
Material	Optical transmission T	Density ρ [g/cm ³]	Acoustic mode	Acoustic velocity v	Acoustic attenuation α_a	Optical polarization	Refractive index n	Figures of (relative tc	merit) fused quart	(z
	[1114]			[srt /mm]	[(zup . sml)/gm]			M_{1}	M_2	M_3
Fused silica	0.24.5	2.2	L	5.96	7.2	-	1.46	1.0	1.0	1.0
$LiNbO_3$	$0.4. \ldots 4.5$	4.64	L[100]	6.57	1.0	$35^{\circ} Y$ Rot.	2.2	8.5	4.6	7.7
			$ m S(100)35^\circ$	~ 3.6	~ 1.0	[100]	2.2	2.3	4.2	3.8
TiO_2	$0.45\ldots 6.0$	4.23	L[110]	7.93	1.0	-1	2.58	18.6	6.0	14
$PbMoO_4$	$0.42\dots 5.5$	6.95	L[001]	3.63	5.5	-	2.39	14.6	23.9	24
BGO	$0.45\dots 7.5$	9.22	L[110]	3.42	1.6	Arb.	2.55	3.8	6.7	6.7
${\rm TeO_2}$	$0.35\dots 5.0$	6.0	S[110]	0.62	17.9	Cir.	2.26	13.1	795	127
$_{\rm GaP}$	$0.6. \dots 10.0$	4.13	L[110]	6.32	8.0	=	3.31	75.3	29.5	71
			S[110]	4.13	2.0	-	3.31		16.6	26
$\mathrm{Hg_2Cl_2}$	0.3028.0	7.18	S[110]	0.347	8.0	=	2.27	4.3	703	73
GaAs	1.011.0	5.34	L[110]	5.15	15.5	=	3.37	118	69	137
GeAsSe-glass	$1.0\dots 14.0$	4.4	L	2.52	1.7	-	2.7	54.4	164	129
Ge	2.022.0	5.33	L[111]	5.5	16.5	=	4.0	1117	482	1214
${ m Tl}_{3}{ m AsS}_{4}$	$0.6. \dots 12.0$	6.2	L[001]	2.15	5.0	=	2.83	1.52	523	416
${ m Tl}_{3}{ m AsSe}_{3}$	$1.26\dots 13.0$	7.83	L[100]	2.05	14.0	=	3.34	209	2259	1772
L: longitudinal	mode; S: shear m	ode.								

7.1.5.3 Acousto-optic intensity modulator

In an analog acousto-optic modulator the acoustic wave is Amplitude-Modulated (AM) and the incident laser beam is focused onto the sound grating as shown in Fig. 7.1.9. As already described for bandwidth considerations the ratio of divergences of light beam and acoustic beam should be near 1. The three acoustic waves, namely carrier, the upper, and the lower sidebands will generate three correspondingly diffracted light waves traveling in separate directions. The modulated light intensity is determined by overlapping collinear heterodyning of the diffracted optical carrier beam and the two sidebands. The focused-beam-type modulator has certain disadvantages. The diffraction spread associated with the narrow optical beam tends to lower the diffraction efficiency. More importantly, the focusing of the incident beam results in a high peak intensity that can cause damage for even relatively low laser power levels. Then it is necessary to open up the optical aperture. Due to the basic issues of acoustic transient time, the temporal bandwidth of the modulator will be severely degraded.

A variety of acousto-optic modulators are commercially available suited for external or intracavity applications. Compared to the competing electro-optic modulators, the acousto-optic modulators have some advantages that include low driving power, high extinction ratio, and insensitivity to temperature change. When bandwidths below the GHz range are required, acousto-optic modulators are mostly preferred to their electro-optic counterparts. Low-cost types, usually made of special glasses, are useful for bandwidth up to about 10 MHz. The use of superior materials such as PbMoO₄ and TeO₂ has raised the modulation bandwidth up to about 50 to 100 MHz. And a GaP modulator can operate down to about 2 ns rise time, which corresponds to a maximum modulation bandwidth of about 500 MHz. Table 7.1.6 lists a few selected acousto-optic modulators with some of their typical parameters.

Material	Wavelength $\lambda \; [\mu m]$	Center frequency f_0 [MHz]	RF bandwidth Δf [MHz]	Rise time $t_{\rm r}$ [ns]	Diffraction efficiency η [%]
PbMoO ₄	0.633	80	40	25	80
${\rm TeO_2}$	0.633	110	50	20	75
${\rm TeO}_2$	0.633	200	100	7	65
GaP	0.83	500	250	4	50
GaP	0.83	1000	500	2	30

Table 7.1.6. Typical performances of a few selected acousto-optic modulators [95Cha].

GeAsSe glass is particularly suited to wideband applications at 1.06 μ m wavelength due to its exceptionally low acoustic loss (1.7 dB/(μ s · GHz²)). At 10.6 μ m, the most popular material is single-crystal Ge. A Ge modulator shows a typical rise time of 30 ns and an efficiency of 5%/Watt. Acousto-optic materials with high optical quality, such as fused silica, are exclusively used inside a laser cavity. These applications include *Q*-switching, mode locking, and cavity dumping.

7.1.5.4 Acousto-optic deflector

Basically, acousto-optic deflectors operate in the same way as acousto-optic Bragg modulators; the main difference is that the frequency rather than the amplitude of the sound wave is modulated – Frequency-Modulated (FM) wave. Figure 7.1.10 shows the operation principle of a Bragg deflector.



Fig. 7.1.10. Bragg diffraction acousto-optic deflector (Bragg cell).

Basic equation for this application is

$$\Delta \theta = \frac{\Delta f \cdot \lambda}{nv \cdot \cos \theta_{\rm B}} \,. \tag{7.1.44}$$

In a deflector, the most important performance parameters are resolution and speed. Therefore, the divergence of the optical beam should be small, see Fig. 7.1.10. Resolution, or the maximum number of resolvable spots, is defined as the ratio of the range of deflection angle divided by the angular spread of the diffracted beam. Taking the angular beam divergence in case of a Gaussian beam as

$$\delta\theta_0 = \frac{2\lambda}{\pi n\bar{d}}$$

it results in

$$N = \frac{\Delta\theta}{\delta\theta_0} = \frac{\pi \bar{d}}{2v \cdot \cos\theta_{\rm B}} \cdot \Delta f \approx \tau \cdot \Delta f , \qquad (7.1.45)$$

where $\tau = \bar{d}/v$ is the acoustic transit time across the optical aperture. A common figure of merit is the ratio of the total number of resolvable spots to the access time, determined by the feasible bandwidth according to

$$\frac{N}{\tau} = \Delta f \ . \tag{7.1.46}$$

Thus, a high speed–capacity product is only achievable when the bandwidth is large, which requires, as in the case of a modulator, a high modulation frequency f of the device. Basically, there is a trade-off between the speed and the resolution of AO deflectors. The maximum resolution is limited to a few thousands. Table 7.1.7 shows examples of some acousto-optic deflectors (Bragg cells).

Besides bandwidth, the acoustic attenuation across the aperture is a basic parameter, which in most solids is proportional to the square of the acoustic frequency, i.e. f^2 . For an allowable average attenuation across the band, $\bar{\alpha}$ [dB], the maximum deflector resolution can be estimated as

$$N_{\max} = \frac{\bar{\alpha}}{\alpha_{\rm a} \cdot f_0} , \qquad (7.1.47)$$

where $\alpha_{\rm a} \left[dB/(\mu s \cdot GHz^2) \right]$ is the acoustic attenuation coefficient and f_0 is the center frequency.

Material	Center frequency f_0 [MHz]	Bandwidth Δf [MHz]	Access time τ [us]	Resolution $N = \tau \Delta f$	Efficiency η [% / W]
$\overline{T_0 O_{\tau_0}(\mathbf{S})}$	00	50	40	2000	110
$1eO_2(3)$	90	50	40	2000	110
TeO_2 (S)	160	100	10	1000	95
GaP(S)	1000	500	2.0	1000	30
GaP (S)	2000	1000	1.0	1000	12
$LiNbO_3$ (S)	2500	1000	1.0	1000	10
GaP(L)	2500	1000	0.25	250	44
GaP(L)	3000	2000	0.15	300	10
GaP(S)	3000	2000	0.25	500	8
$LiNbO_3$ (S)	3000	2000	0.30	600	6

Table 7.1.7. Performance of acousto-optic Bragg cells at $\lambda = 830$ nm [95Cha].

L: longitudinal mode; S: shear mode.

7.1.6 Glossary

D	displacement vector
\boldsymbol{E}	electric field vector
c	velocity of light in vacuum
d	width of the electro-optic crystal
\bar{d}	optical beam width
f	acoustic frequency
$(\Delta f)_{\rm m}$	modulation bandwidth
Η	acoustic beam height
Ι	intensity
\boldsymbol{k}	optic wave vector
k	optic wave number
K	acoustic wave vector
K	acoustic wave number
L	length of the electro-optic crystal, acoustic beam width
M	figure of merit
N	number of resolvable spots
n	optical refractive index
n_x, n_y, n_z	principal indices of refraction
$\Delta\left(\frac{1}{n^2}\right)$	change in an impermeability tensor element
K	Kerr constant
$P_{\rm a}$	acoustic power
p	strain-optic coefficient
r	linear electro-optic coefficient
s	quadratic electro-optic coefficient
S	strain
v	sound velocity
V	applied voltage
V_{π}	half-wave voltage
(x, y, z)	principal dielectric coordinate system
$(x^\prime,y^\prime,z^\prime)$	perturbed principal dielectric coordinate system

α_{a}	acoustic attenuation
δ	phase modulation index
ϵ	dielectric constant
η	diffraction efficiency
κ	coupling constant
λ	vacuum wavelength of light
ϕ	deflection angle
ρ	mass density
σ	modulation index reduction factor
au	transit time
$ u_{ m m}$	modulation frequency
ω	light radian frequency
$\omega_{ m m}$	modulation radian frequency
Φ	phase of the optical field
$\Delta \Phi$	phase shift of light
Г	phase retardation
$\Gamma_{\rm m}$	amplitude modulation index
Λ	acoustic wavelength
$ heta_{ m B}$	Bragg angle
$\delta\theta_{\rm o},\delta\theta_{\rm a}$	beam divergence: optical, acoustic
Ω	acoustic radian frequency

References for 7.1

61Nam 61Woo	Namba, C.S.: J. Opt. Soc. Am. 51 (1961) 76. Woodruff, T.O., Ehrenreich, H.: Phys. Rev. 123 (1961) 1553.
66Gor	Gordon, E.I.: Appl. Opt. 5 (1966) 1629.
67Dix 67Kam 67Kle 67Nye	Dixon, R.W.: J. Appl. Phys. 38 (1967) 3634. Kaminow, I.P., Sharpless, W.M.: Appl. Opt. 6 (1967) 351. Klein, W.R., Cook, B.D.: IEEE Sonics Ultrason. 14 (1967) 123. Nye, J.F.: Physical Properties of Crystals, Oxford: Clarendon Press, 1967.
72Den 72Kor 72Pin	Denton, R.T.: Laser handbook, Arecchi, F.T., Schulz-DuBois, E.O. (eds.), Amsterdam: North Holland Publ. Co., 1972, p. 703–724. Korpel, A.: Acousto-Optics, Applied Solid State Science, Vol. 3, Wolfe, R. (ed.), New York: Academic Press, 1972. Pinnow, D.A.: Laser handbook, Arecchi, F.T., Schulz-DuBois, E.O. (eds.), Amsterdam: North Holland Publ. Co., 1972, p. 995–1008.
72Wem	Wemple, S.H.: Laser handbook, Arecchi, F.T., Schulz-DuBois, E.O. (eds.), Amsterdam: North Holland Publ. Co., 1972, p. 975–994.
75Yar	Yariv, A.: Quantum Electronics, New York: J. Wiley & Sons, 1975.
79Coo	Cook, W.R., Jaffe, H.: Landolt Börnstein, Vol. III/11, Hellwege, KH. (ed.), Heidelberg: Springer-Verlag, 1979, p. 287.
79Hel	Hellwege, KH.: Landolt Börnstein, Vol. III/11, Heidelberg: Springer-Verlag, 1979.
84Yar	Yariv, A., Yeh, P.: Optical waves in crystals, New York: J. Wiley & Sons, 1984.
86Got 86Kam	Gottlieb, M.: Handbook of laser science and technology, Vol. IV, part 2, Boca Raton, Fl.: CRC Press, 1986, p. 319–341. Kaminow, I.P.: Handbook of laser science and technology, Vol. IV, part 2, Boca Raton, Fl.: CRC Press, 1986, p. 253–278.
89Gha	Ghatak, A., Thyagarajan, K.: Optical electronics, New York: Cambridge Univ. Press, 1989.
90Alf	Alferness, R.C.: Guided wave optoelectronics, Chap. 4, Tamir, T. (ed.), New York: Springer-Verlag, 1990.
92Dol	Dolfi, D.W., Ranganath, T.R.: Electron. Lett. 28 (1992) 1197.
95Cha 95Mal	Chang, I.C.: Handbook of optics II, New York: McGraw-Hill, Inc., 1995, p. 12.1–12.54. Maldonado, T.A.: Handbook of optics II, New York: McGraw-Hill, Inc., 1995, p. 13.1–13.35.
97Lee	Lee, N.L.: Handbook of photonics, Gupta, M.C. (ed.), New York: CRC Press, 1997, p. 393–434.
02Tak	Takeuchi, H., Saitho, T., Ito, H.: Technical Digest OFC2002, WV1.

7.2 Thin-film technology

D. RISTAU

7.2.1 Introduction

Thin-film technology plays a key role in the development and application of modern lasers. Optical components with advanced thin-film systems are employed in each functional section of a laser, starting at the optical pumping scheme and the generation of laser radiation, over the manipulation and steering of the beam until shaping and focusing for its final application. As a consequence, the quality of optical coatings is often recognized as a limiting factor for the reliability and economic efficiency of many laser applications. Therefore, optical coatings and their specifications should be carefully considered during the design and implementation of laser systems. In this chapter, a brief review will be given on the state of the art of optical thin-film technology and on the characterization of optical coatings. Selected quality parameters of optical laser components will be presented and discussed with respect to typical applications in laser technology and modern optics.

7.2.2 Basic principle of optical thin-film systems

The fundamental structure of a dielectric layer system is illustrated in Fig. 7.2.1a. Thin transparent layers with a thickness in the range of the wavelength, at which the system is applied, are deposited on the surface of the optical component (refractive index $n_{\rm T}$). In order to adapt the optical function



Fig. 7.2.1a. Basic structure of a thin-film system: Transparent layers of at least two different materials $(n_{\rm H}, n_{\rm L})$ are deposited on a substrate $(n_{\rm T})$.

Fig. 7.2.1b. Path of the incident beam in a single layer. The reflection and transmission coefficients at the interfaces x are denoted by r_x and t_x , respectively.

 $t_1t_1'r_1'r_2^2$

 $t_1 t_1' r_1'^2 r_2^3$

of the coated surface to the required specifications, at least two layer materials with a high refractive index $n_{\rm H}$ and a low index $n_{\rm L}$ have to be selected. Besides the contrast in the indices of refraction, the choice of the thin-film materials is dependent on the application wavelength and the expected environmental influences on the coating system. The optical function of the layer structure is based on the interference of partial waves reflected and transmitted at the numerous interfaces between the layers. Thus, the spectral behavior of the coating can be described by tracing the incoming wave through the layer system and accumulating all contributions of the induced partial waves. This calculation method is quite comprehensible for a single-layer system (see Fig. 7.2.1b), where the interfering partial waves can be calculated with the Fresnel equations taking into account the phase shift acquired by the waves passing through the bulk of the layer. But, considering the enormously increasing number of partial beams, the resulting equations become extremely complicated for multilayer structures. Nowadays, the matrix formalism, which can be deduced from the boundary conditions of the electric and magnetic field at the interfaces [64Bor], is usually employed for the calculation of thin-film systems. In this approach, each layer is represented by a matrix M, which contains all specific parameters of the film [88The]. The matrix M_m of a thin film at the position m within the layer system relates the electric (E) and magnetic (H) field strength at the interface between the (m-1)th and mth layers:

$$\begin{pmatrix} E_{m-1} \\ H_{m-1} \end{pmatrix} = \mathsf{M}_m \begin{pmatrix} E_m \\ H_m \end{pmatrix} = \begin{pmatrix} \cos\varphi_m & \frac{1}{n_m} \cdot \mathbf{i} \cdot \sin\varphi_m \\ n_m \cdot \mathbf{i} \cdot \sin\varphi_m & \cos\varphi_m \end{pmatrix} \begin{pmatrix} E_m \\ H_m \end{pmatrix} .$$
(7.2.1)

The phase shift φ_m in (7.2.1) is dependent on the index of refraction n_m , the thickness d_m , and the angle of propagation θ_m within the *m*th layer:

$$\varphi_m = \frac{2\pi n_m d_m \cos \theta_m}{\lambda} \,. \tag{7.2.2}$$

The wavelength of the incoming electric wave is denoted by λ . As the outstanding advantage of the matrix method, the calculation of a layer structure can be simply accomplished by a multiplication of the matrices M_m for the constituent single layers:

$$\mathsf{M} = \mathsf{M}_1 \times \mathsf{M}_2 \times \mathsf{M}_3 \times \dots \times \mathsf{M}_k , \qquad (7.2.3)$$

where M denotes the transfer matrix of the entire stack with a number of k layers, and the multiplication is performed according to the rule for 2×2 -matrices.

The transmittance of the multilayer stack is given by the ratio of the transmitted-field to the incoming-field amplitude. In the case of non-absorbing layers, i.e. the imaginary parts of the refractive indices are zero, the transmittance can be directly expressed as a function of the matrix elements M_{ij} (with n_0 : refractive index of the surrounding medium, n_T : refractive index of the substrate):

$$T = 1 - R = 4 \left(2 + \frac{n_0 M_{11}^2}{n_{\rm T}} + \frac{n_{\rm T} M_{22}^2}{n_0} + n_0 n_{\rm T} M_{12}^2 + \frac{M_{21}^2}{n_0 n_{\rm T}} \right)^{-1} \,. \tag{7.2.4}$$

The matrix method can be easily transferred in computer codes [80Lid] and has been developed to a standard tool in optical thin-film technology. Often, the unit QWOT (Quarter-Wave Optical Thickness) is used in these software tools to express the optical thickness D_m of the layers at the design wavelength λ_d :

$$D_m = \frac{4n_m d_m}{\lambda_d} \quad [\text{QWOT}] . \tag{7.2.5}$$

At integer QWOT-values, corresponding to multiples of $\lambda_d/4$, the matrices M_m reduce to simple expressions for normal incidence. Also, these thickness values represent an extreme value in the spectral behavior, because the according phase shift φ_m is a multiple of π . As a consequence,



Fig. 7.2.2. Reflectance spectra of QWOT-coating stacks of Ta_2O_5/SiO_2 with different numbers (k = 1, 5, 13, 21) of single layers for a design wavelength of 1.064 µm.

layers with integer QWOT-values can be accurately deposited on the basis of monitoring the reflection or transmission at the wavelength λ_d during the production process. If such an optical thickness is approached, the extreme values in the transfer behavior of the growing layer can be precisely detected, and the deposition can be stopped. Thus, the unit QWOT is mainly of practical importance, and most of the layer system designs used in thin-film production are based on QWOT-layers of two deposition materials. The designs are often expressed using the notations H for a QWOT layer of the high-index material and L for the low-index material, respectively. For example, a periodic stack composed of an odd number k of QWOT layers is denoted by the following expression:

$$\underbrace{H|L|H|L|\cdots|L|H|}_{k \text{ layers}} \text{ substrate }.$$
(7.2.6)

Such a HL-stack is the basic design for dielectric high-reflecting mirrors and output couplers. For normal incidence, the reflectance at the design wavelength can be directly determined:

$$R_{k} = \left(\frac{n_{\rm H}^{k+1} - n_{0}n_{\rm T}n_{\rm L}^{k-1}}{n_{\rm H}^{k+1} + n_{0}n_{\rm T}n_{\rm L}^{k-1}}\right)^{2} \approx 1 - 4n_{0}n_{\rm T}\frac{n_{\rm L}^{k-1}}{n_{\rm H}^{k+1}}.$$
(7.2.7)

The approximation in (7.2.7) is meaningful for the condition $(n_{\rm H}/n_{\rm L})^k \gg 1$, which can be fulfilled by most combinations of deposition materials used in practice. Obviously, the reflectivity of the QWOT-stack increases with the refractive-index contrast between the deposition materials and the number k of layers. Typical spectra for a QWOT-stack of Ta₂O₅/SiO₂ are illustrated in Fig. 7.2.2 for different numbers of layers. Around the reflectance band, the spectra oscillate with decreasing amplitude originating at the reflectivity of the uncoated substrate as minimum value. If the angle of incidence is tuned from 0° to higher angles, the spectral characteristic is shifted towards shorter wavelengths, and the reflectivity of the mirror decreases. In most cases, a dielectric mirror with optimized reflectivity for an angle of 0° will be nearly transparent if it is applied at an angle of 45° at the same wavelength.

Besides mirror coatings, antireflective coatings are fundamental thin-film systems in optical technology. Depending on the wavelength range, the substrate, and the intended residual reflectivity, antireflective coatings may be composed of non-QWOT-layers of more than two materials. A typical example for a demanding application is ophthalmology, where antireflective coatings on eye-glasses have to cover the whole VIS-range and must be resistant against mechanical and chemical influences. For most applications in laser technology, only one or a few wavelengths have to be considered resulting in less complicated designs. A total suppression of the residual reflection for a single wavelength is already achievable with two non-QWOT layers for most optical materials, see Fig. 7.2.3. This so-called V-coating can be routinely produced for all laser wavelengths and possesses good optical properties as well as high laser-induced-damage thresholds.



Fig. 7.2.3. Reflectance spectrum of a two-layer antireflective coating for the wavelength 550 nm. Design: $0.37H(Ta_2O_5, n_H = 2.03)/1.30L(SiO_2, n_L = 1.45).$

7.2.3 Production of optical coatings

The production process for high-quality thin films can be divided into stages which are nearly independent of the optics type and the thin-film deposition process applied. First, a thin-film design which defines the sequence of single layers and their materials is elaborated. This design has to meet the requirements of the user including the spectral transfer function, the stability, and the optical quality of the coatings. In the next step, the optical coating system is deposited on the optical components with a process selected in view of the individual specifications of the application. Finally, the coated product is inspected and characterized with respect to the critical operating parameters within the quality management system of the manufacturer.

The design of even highly complicated optical coating systems is no longer an obstacle. Nowadays, commercial computer programs are available which automatically calculate the desired design on the basis of advanced calculation algorithms and optimization techniques including special needle methods and genetic codes [88Bau]. For most software products, the spectral transfer function needed is loaded, and the computer produces suggestions for coating designs. Of course, the power of the software environments must be combined with the expert knowledge of a scientist possessing experience in the properties of different coating materials (see Fig. 7.2.4) and their combination.

Presently, the production of a coating system according to a specific design is the primary difficulty in thin-film technology. Most industrial coating areas are still dominated by thermal processes like e-beam evaporation, which has been introduced in the 1960ies, or even by boat evaporation techniques, which date back to the 1940ies. In these thermal processes, the coating material is evaporated and condenses as film on the optical components under vacuum conditions.



Fig. 7.2.4. Spectral transmission bands of materials often used for thin films or substrates.



Fig. 7.2.5. Microstructure of a single layer of MgF_2 deposited by conventional boat evaporation on a silicon substrate [99Ris].

Since the typical kinetic energy of the admolecules is confined to a few tenths of an eV, they attain a very limited mobility on the surface of the growing layer and can reach only bonding positions near to their location of incidence. This low-energy growing mechanism results in a porous microstructure with severe disadvantages for the mechanical and optical film quality, see Fig. 7.2.5. For example, water and other contaminants from the atmospheric environment are strongly adsorbed by this porous microstructure leading to high optical absorption values and an elevated thermal shift of the layer systems. In addition, the porous microstructure suffers from a reduced mechanical stability, and it is the origin of high total scatter values of the coating system. In order to increase the mobility of the admolecules and to improve the quality of thermally deposited films, the substrates often have to be heated to temperatures in the range from 200 to 350°C. Thus conventional coating processes cannot cover the broad spectrum of temperaturesensitive optical components, as plastic optics and many crystal materials used in laser technology. For the demanding specification of many modern technology fields, the quality of conventionally deposited optical coatings is not sufficient. As a consequence, several new concepts have been developed to overcome the specific difficulties of thermal evaporation processes during the last two decades [92Gue].

Considering the condensation process of the layers, two basic approaches can be discussed to improve the optical and microstructural quality: In the first concept, the Ion-Assisted Deposition (IAD), the mobility of the admolecules is enhanced by an additional bombardment of the growing layer with ions of inert or reactive gases. In this process, auxiliary energy is transferred by collisions with the ions resulting in a longer mean free path of the admolecules, which accordingly can reach locations with nearly optimal bonding conditions in the growing layer. Hence, the ion-assisted deposited thin-film structure possesses a higher packing density with improved mechanical and environmental stability. Furthermore, especially for oxide materials, the stoichiometry and consequently, the optical quality of the layer can be optimized by employing oxygen for the generation of the assisting ions. Also, the coating of temperature-sensitive substrates is possible with IAD-processes, because heating of the substrates is no longer necessary. As an additional technical and economical advantage, IAD-processes can be easily integrated in existing deposition plants for conventional deposition by installing an appropriate ion source.

As the second alternative for an improvement of the coating quality, a direct increase of the admolecule energy by sputtering processes can be considered. In these deposition concepts, the coating material is sputtered from a target by high-energetic ions which are produced within a discharge maintained in the deposition plant or by a separate ion gun (Ion-Beam Sputtering: IBS,





Fig. 7.2.6. Principle of an ion-beam sputtering process. Sputtering of the target material is achieved by inert ions, which are produced in a Kaufman ion source.

Fig. 7.2.7. Microstructure of a high-reflecting stack for the wavelength 633 nm produced by ion-beam sputtering. (Detail of a surface of the break through the system).

[78Wei], see Fig. 7.2.6). In contrast to thermal evaporation, the kinetic energy of the sputtered molecules ranges on a significantly higher level between 5 to 20 eV. Therefore, the microstructure of layer systems deposited by sputter processes is compact and extremely stable, see Fig. 7.2.7. As a result of the separation of the process steps, ion generation, sputtering of the coating material, and finally the condensation of the layer structure, these coatings exhibit also extremely low defect densities and contamination levels. Actually, coatings produced by IBS-processes achieve nearly ideal properties with lowest total optical losses below 1 ppm for laser mirrors in the visible and near-infrared spectral region. Therefore, IBS processes are presently employed for a variety of special coating challenges including filter systems for telecommunication, low-loss mirrors for optical measurement techniques, high-power laser applications, and coating of optoelectronic circuits. With respect to the economic efficiency, IBS-processes still suffer from the small effective deposition area and the low deposition rates, which range between a factor 5 to 10 below typical rates of evaporation processes. Several industrial and institutional research groups are presently investigating techniques to improve the economy of IBS-processes and to simultaneously advance or maintain the extremely high quality of the coating products.

7.2.4 Quality parameters of optical laser components

Some of the quality parameters frequently considered in laser technology and optics are compiled in Table 7.2.1 in conjunction with the corresponding measurement techniques. Besides the spectral transfer functions $R(\lambda)$ and $T(\lambda)$, which are determined by spectrophotometry, laser-induced damage thresholds and optical losses must be carefully thought about for most laser applications. The measurement of these laser-specific parameters requires adapted measurement facilities described within International Standards which are frequently updated with respect to the latest developments in laser technology [06ISO2]. The measurement procedures for absorptance (ISO 11551 [97ISO3]) and total scatter (ISO 13696 [02ISO]) have been tested in several international interlabo-

Specification	Parameter	Unit	Standard	Measurement principle
Laser-Induced Damage Threshold (LIDT)	cw-LIDT 1 on 1-LIDT S on 1-LIDT certification	W/cm J/cm^2 J/cm^2 J/cm^2	ISO 11254-1: ISO 11254-1: ISO 11254-2: DIS 11254-3:	cw-laser irradiation irradiation with single pulses repetitive irradiation with pulses irradiation sequence
optical losses	absorptance total scattering	ppm ppm	ISO 11551: ISO 13696:	laser calorimetry integration of scattered radiation
transfer function	reflectance transmittance	% %	FDIS 13697: ISO 15368:	precise laser ratiometric method spectrophotometry
surface quality	form tolerances scratch/digs roughness	λ/N	ISO 10110:	17 parts containing different types of imperfections
stability	abrasion environmental stability		ISO 9211: ISO 9022:	different test methods more than 20 parts containing a va- riety of conditioning methods

 Table 7.2.1. Selected quality parameters of optical laser components and the corresponding characterization techniques.

ratory tests. Also, the practicability of the standards for the measurement of laser-induced damage thresholds (ISO 11254, parts 1 and 2, [01ISO2]) has been confirmed by detailed investigations at a variety of different operation conditions of the test lasers [05Las]. Additional parameters for the surface quality and environmental stability of optical laser components are described in many national and international standards, which have been elaborated by the optics industry and users during the last five decades [97ISO1, 97ISO2, 06ISO1]. In laser applications, especially the surface quality of the optics is a crucial factor, because the optical losses and the power handling capability of the coatings are directly dependent on the surface roughness and defects. The environmental stability should be always considered for industrial laser applications, extreme climates, or high-power laser systems. For example, the spectral characteristic of a coating may be shifted under the influence of high temperatures or extreme humidity.

7.2.5 Measurement of critical parameters of laser components

Compared to application in optics, laser technology imposes much higher demands on optical coatings with respect to their optical losses and damage thresholds. For projecting a laser system, these specifications have to be discussed also in the context of their specific measurement methods. Therefore, a short outline on the determination of absorption, scattering, and laser-induced damage threshold will be given in the following sections.

7.2.5.1 Calorimetric measurement of absorption

Absorption in optical laser components leads to a conversion of a fraction of the impinging laser power into heat, which dissipates in the bulk of the component and induces distortion effects. For



Fig. 7.2.8. Exponential method for the evaluation of temperature recorded during a calorimetric absorptance measurement according to ISO 11551. The irradiation time is indicated by the rectangular graphs at the bottom of the diagram.



Fig. 7.2.9. Example for a laser calorimetric measurement with high sensitivity on an uncoated fused silica substrate (thickness: 1 mm, measured absorptance: 2.8 ppm) at 1.064 μ m. The temporal behavior of the laser power is also indicated.

example in laser material processing, these thermal distortion effects can induce a shift of the focal plane on the work piece, which in turn, deteriorates the process result. The determination of absorptance according to ISO 11551 is based on the laser calorimetric method [76Gib, 88DeB, 98Wil], which provides the means for absolute measurements by uncomplicated calibration techniques. For the calorimetric measurement, a temperature sensor is attached to the specimen which is located in a thermally isolating chamber. After thermal equilibrium between the sample and the environment is reached, the sample is irradiated by a laser beam with known power P starting at the time t_1 for a heating period with a duration $t_{\rm B}$. During this heating period, the temperature of the specimen increases according to the absorbed laser power. At a defined time t_2 , the laser is switched off, and the temperature of the specimen decreases as a consequence of heat dissipation to the environment. For the evaluation of the calorimetric measurement, the recorded temperature curves of the heating and cooling cycles are considered. In most cases, the temperature behavior is modeled directly on the basis of a solution for the heat conduction equation with the boundary conditions according to the sample geometry. For a sample with infinite thermal conductivity and small temperature increases, the corresponding solutions of the heat equation can be expressed by exponential functions:

heating curve:
$$T(t) = T(t_1) + \frac{\alpha P}{\gamma C_{\text{eff}}} (1 - \exp(-\gamma(t - t_1)))$$
,
cooling curve: $T(t) = T(t_1) + \frac{\alpha P}{\gamma C_{\text{eff}}} (\exp(-\gamma(t - t_2)) - \exp(-\gamma(t - t_1)))$

where α denominates the absorptance, and γ represents the coefficient for heat losses induced by radiation and convection, respectively. The effective heat capacity C_{eff} is combined of the contributions from the sample and the holder in conjunction with additional heat contact effects to other arrangements in the calorimetric chamber. An example for an evaluation of a calorimetric measurement according to this exponential method is depicted in Fig. 7.2.8. The advantage of this data reduction technique, which is recommended for lower laser powers and exposure times longer than 60 s, is the involvement of all measured data points in the procedure. For an absolute determination of the absorptance value, the laser power P must be measured with a calibrated power monitor.

The laser calorimetric method had been qualified in various global round-robin tests and by fundamental investigations in the sensitivity of the method, see Fig. 7.2.9. Nowadays, absorptance values according to ISO 11551 are given in product catalogs of most optical companies.

7.2.5.2 Measurement of total scattering

In laser technology, the amount of radiation scattered by the optical components is of major concern. For example, scattering may deteriorate the quality of laser material processing or it may limit the precision in laser metrology applications. In the UV-spectral range, losses by optical scattering may directly impair the economic efficiency, because the production of UV-laser photons is expensive compared to most other prominent wavelengths. Also, laser safety aspects have to be considered in high-power laser applications, where even low optical scattering may lead to dangerous laser power levels in the environment of the laser and beam-steering system. For the measurement of Total Scattering (TS) by laser components, the International Standard ISO 13696 is applied by most companies and research institutes. In this standard procedure, the radiation scattered by the specimen is collected and integrated by an Ulbricht-sphere [1900Ulb] or a Coblentz-hemisphere [13Cob]. The power of the scattered radiation is measured by a detector and related to the corresponding power of a 100% diffuse reflecting standard. Depending on the halfspace of integration, total backward and total forward scattering are distinguished.

A typical measurement facility, which is equipped with an Ulbricht-sphere for the visible and near-infrared spectral range, is depicted in Fig. 7.2.10. This apparatus consists of separate chambers for the beam preparation and the scatter measurement device, which can be flushed with gases in order to reduce the contribution of Rayleigh scattering to the zero signal of the system below 1 ppm. The integrating performance of the Ulbricht sphere is sensitively dependent on the diffuse reflectivity of the coating on the inner walls. For wavelengths below 200 nm, appropriate materials with the required diffuse reflectance above 99% are not available. Therefore, Coblentz-hemispheres coated with adapted high-reflecting metal and protection layers are preferred for TS-measurement in the deep and vacuum UV-spectral range. In comparison to the integration principle of the Ulbricht-sphere, the collection effect of the Coblentz-sphere can be considered as an optical imaging of the scattered radiation onto the detector, which is positioned at the conjugate point with respect to the sample.

The total scatter value of a coated component is directly dependent on the surface quality of the substrate and the reflectance of the coating. Therefore, substrates for laser applications should



Fig. 7.2.10. Example for a measurement facility for total scattering according to ISO 13696 with an Ulbricht-sphere for the visible and near-infrared spectral range.





have a surface roughness well below 1 nm rms and surface imperfections better than $(5/1) \times 0.010$ according to ISO 10110.

An introduction into the problems of surface roughness and scattering is given in [99Ben]. As an example for the influence of reliable loss measurements on the progress of optical thin-film technology, a learning curve is presented in Fig. 7.2.11 for the optical losses of high-reflecting mirrors produced by ion-beam sputtering. An improvement in optical losses by more than three orders magnitude to values around 1 ppm could be achieved since the introduction of IBS-technology.

7.2.5.3 Laser-induced damage thresholds

The Laser-Induced Damage Threshold (LIDT) of an optical component is one of the most important quality parameters in the development and application of high-power laser systems. Even today, the power handling capability of optical components is a limiting factor in many laser applications including nuclear fusion technology, material processing, or laser medicine. In many laboratories, LIDT-values are determined on the basis of the International Standard ISO 11254, which describes a protocol for the objection of selected sites on the specimen to a focused high-power laser beam with defined output energy or power. After the test, each irradiation site is optically inspected (see Fig. 7.2.12), and the state of damage is documented in conjunction with the corresponding laser power. For the evaluation of these raw data, the so-called survival curve [83Sei] is deduced, which represents the damage probability as a function of the laser power. In this technique, the damage threshold is defined by the highest power value with a damage probability of zero. Besides 1 on 1- and cw-tests outlined in ISO 11254-1, which are of interest for fundamental research of optical coatings, irradiation sequences with repetitive laser pulses (S on 1-tests) are also considered in ISO 11254-2. These S on 1-tests represent the typical operation condition of a component in practical applications.

Even though the fundamental principle of LIDT-measurements is uncomplicated, the experimental expense is considerable. For comparable measurements, laser systems operating in single transversal and longitudinal modes are recommended. Also, a precise beam diagnostic package has to be installed in order to characterize the beam parameters at the specimen or a conjugate location. Besides online damage detection systems, an inspection using a differential interference contrast microscope is prescribed for a reliable identification of the damage state of the irradiated sites.

Extreme care must be taken, if damage thresholds of optical components are compared for different operation conditions. Scaling laws of LIDT-value have been studied for the main param-





Fig. 7.2.12. Damage site on chirped mirror for fslasers (TiO_2/SiO_2). The damage occured at an energy of 0.1 J/cm² during an S on 1-test.

Fig. 7.2.13. Learning curve for high-power laser mirrors at 248 nm produced with different coating materials by thermal evaporation [98Ris].

eters wavelength, pulse duration, and spot diameter [81Wal] and are only applicable, if the actual damage mechanism is clearly identified. Even though 1 on 1-damage tests are not representative for the operation conditions of optical components in typical applications, they are still listed in catalogs of optics manufacturers and considered for a comparison of the power handling capabilities of competing products.

Reliable LIDT-measurements are often the key factor for the optimization of optical coating systems. A typical example is illustrated in Fig. 7.2.13 for high-reflecting mirrors applied in KrF-excimer lasers. Meanwhile the first decade of intense research in high-power UV-coatings was dominated by the combination of oxides and fluorides in one stack, the improved layer systems consist of only one material class. Best results of more than 20 J/cm² are presently achieved for HR-stacks of the material combination Al_2O_3/SiO_2 , which are deposited by a conventional, low-contamination PVD-process.

7.2.5.4 Quality parameters of laser components: present state

In Table 7.2.2 quality parameters of optical laser components for prominent laser systems are compiled. Besides optical absorptance and total scattering, typical LIDT-values are reported for operation conditions in practical applications. Particularly, the LIDT-value for cw-operation is given in units of linear power density, because most cw-damage mechanisms are thermal effects, which can be scaled with linear power density [97Put]. Moreover, a scaling of cw-LIDT-values in units of W/cm² would lead to an overestimation of the power handling capability of the component and could lead to fatal failures in its application. This aspect is of essential importance in the field of CO_2 -laser material processing, where optical components often consist of hazardous materials, which can cause severe health problems and contamination of the environment when emitted during catastrophic damage. The table also reflects the specific advantages of the IBSprocess in comparison to conventional evaporation: Extremely low losses and high LIDT-values can be achieved for HR-coatings in the visible and near-infrared spectral region. However, thermal processes still cover the entire spectral range from the vacuum ultraviolet to the far infrared and enable an efficient production of coatings with highest damage thresholds. Coatings for the

Wavelength	Type	Absorption ISO 11551	Total scattering ISO 13696	Laser-induced damage threshold, ISO 11254
157 nm	$\mathrm{HR/th}$		$1\dots 4\%$	
193 nm	m AR/th $ m HR/th$	0.72.5% 0.42.0%	0.20.5% 0.22.5%	$12 \text{ J/cm}^2 (10n1, 20 \text{ ns})$ $24 \text{ J/cm}^2 (10n1, 20 \text{ ns})$
248 nm	AR/th HR/th HR/IBS	$< 500~\rm{ppm}$	$\begin{array}{l} < \ 0.025 \ \% \\ < \ 0.2 \ \% \\ < \ 0.1 \ \% \end{array}$	$\begin{array}{l} 10 \ {\rm J/cm^2} \ (1{\rm on1}, \ 30 \ {\rm ns}) \\ > 20 \ {\rm J/cm^2} \ (1{\rm on1}, \ 30 \ {\rm ns}) \\ > 3 \ {\rm J/cm^2} \ (1{\rm on1}, \ 30 \ {\rm ns}) \end{array}$
633 nm	$_{ m HR/th}$ HR/IBS	< 30 ppm < 5 ppm	< 30 ppm < 5 ppm	-
1.064 µm	AR/th HR/th HR/IBS	< 20 ppm < 50 ppm < 1 ppm	< 100 ppm < 100 ppm < 1 ppm	$> 60 \text{ J/cm}^2 (12 \text{ ns}, 0.25 \text{ mm}) > 100 \text{ J/cm}^2 (12 \text{ ns}, 0.25 \text{ mm}) > 80 \text{ J/cm}^2 (12 \text{ ns}, 0.25 \text{ mm})$
10.6 µm	AR/th HR/th	< 0.16 % < 0.10 %	_	> 20 J/cm ² (100 ns, 1.4 mm) > 2 kJ/cm ² (1.2 ms, 250 μ m) > 3 kW/mm (cw, 100 μ m) > 25 J/cm ² (100 ns, 1.4 mm) > 2 kJ/cm ² (1.2 ms, 250 μ m)
	Wavelength 157 nm 193 nm 248 nm 633 nm 1.064 µm 10.6 µm	WavelengthType157 nmHR/th193 nmAR/th HR/th248 nmAR/th HR/th glas633 nmHR/th HR/IBS1.064 µmAR/th HR/IBS10.6 µmAR/th HR/th glas10.6 µmHR/th HR/th	Wavelength Type Absorption ISO 11551 157 nm HR/th $ISO 11551$ 193 nm AR/th HR/th $0.72.5\%$ 248 nm AR/th HR/th HR/tBS $0.72.0\%$ 633 nm AR/th HR/IBS < 500 ppm 1.064 μ m AR/th HR/tBS < 30 ppm < 5 ppm 10.6 μ m AR/th HR/th < 20 ppm < 1 ppm 10.6 μ m AR/th HR/th $< 0.16 \%$	WavelengthTypeAbsorption ISO 11551Total scattering ISO 13696157 nmHR/th 14% 193 nmAR/th HR/th $0.72.5\%$ $0.42.0\%$ $0.20.5\%$ $0.22.5\%$ 248 nmAR/th HR/th HR/IBS < 500 ppm $< 0.025 \%$ $< 0.1 \%$ 633 nmHR/th HR/IBS < 30 ppm $< 5 ppm$ < 30 ppm $< 5 ppm$ 1.064 μ mAR/th HR/IBS < 20 ppm $< 100 ppm< 1 ppm< 100 ppm< 100 ppm< 1 ppm10.6 \mumAR/thHR/th< 0.16 \%-$

Table 7.2.2. Selected quality parameters of optical coating systems for laser applications (types: HR:high-reflecting mirror, AR: antireflective coating, th: thermal evaporation, IBS: ion beam sputtering).

ArF- and F_2 -excimer laser are presently optimized by several working groups for semiconductor lithography and material processing. Therefore, the corresponding data are preliminary and may be significantly improved in the near future.

7.2.6 Examples for advanced laser components

The rapid development of laser technology imposes ever increasing demands on optical components and their production methods. Besides extreme requirements on the optical quality and spectral properties of optical coatings, often a flexible production with short delivery time for small numbers of components is inquired of the optical companies. The presently established development process for high-quality components, which involves a variety of iteration steps, is not adapted to these requirements. An ideal process concept would be linear, starting with the coating design, loading the design in the manufacturing system, and ending with coatings according exactly to the design specifications. Present research activities in thin film technology are focused on such a "Rapid Manufacturing" of optical coating systems.

Adapted on-line monitoring, process-tracing algorithms, and error detection during the deposition process are probably the key to rapid-manufacturing concepts in thin-film technology [02Dob]. Therefore, several approaches on the basis of different high-energetic deposition processes in combination with advanced on-line monitoring techniques are presently pursued by many research teams. As deposition processes with high potentiality for integrated manufacturing of optical coatings, IBS and magnetron sputtering techniques are considered. Besides an extremely high process reliability and stability, these concepts allow for the production of coatings with excellent optical quality. In most systems, broad-band spectrophotometers with high accuracy and spectral resolution are employed for a direct spectral evaluation of the actual coating part during the deposition.



Fig. 7.2.14. Spectra of a band-pass filter produced in a rapid-manufacturing cycle by ion-beam sputtering. The spectrum of the realized filter is compared to the spectrum of the design calculated for this application [01Gro].



Fig. 7.2.15. Thin-film polarizer for high-power laser applications. The spectrum of the realized polarizer is compared to the spectrum of the design which was optimized by a thin-film design program [01Gro].

For this purpose, calibrated spectra are periodically read into a computer system, which calculates the current state of the film structure on the basis of specially developed process tracing models. By comparing the present state of the deposition to the target design of the layer system, the switching between the single layers can be controlled, and errors can be detected. For some applications in laser technology, rapid-manufacturing processes are already employed in a laboratory scale. As an example, the spectra for a band-pass filter are depicted in Fig. 7.2.14. This band pass was designed for a special application in solid-state laser technology, where a high-reflecting region between 840 nm to 1060 nm was required, meanwhile the filter should transmit light in the wavelength range from 665 nm to 790 nm without losses and coloration effects. The filter system was produced by IBS in a single rapid-manufacturing cycle and fulfills the requirements within an error margin of less than 0.25%. In Fig. 7.2.15 a high-power thin-film polarizer deposited by IBS is illustrated. This layer system consists of 27 non-QWOT layers of TiO₂/SiO₂ with a total thickness of 3.2 μ m and enables polarization losses below 2.5%. Considering the present state of the art, an implementation of rapid-manufacturing techniques in the industrial production environment can be expected for this decade.

7.2.7 Summary and future trends

The production of optical thin films is an enabling technology for the development and application of future-oriented laser systems. Meanwhile the theoretical background for the design is mastered and implemented in commercial software environments, major deficiencies still exist for a reliable production and characterization of optical coatings. Intensive research and the development of new deposition processes as well as manufacturing strategies are necessary to meet the ever increasing demands of laser technology and modern optics. Besides the optimization of advanced deposition concepts with superior stability, rapid-manufacturing strategies are considered as a major tool to overcome the present limitations in optical thin-film technology.

Future trends in thin-film technology are governed by outstanding laser applications, which act as pace-makers for progresses in high-technology fields. A major driving force can be seen in semiconductor lithography, where on the way up to an ever increasing packing density the wavelength of the stepper systems is continuously reduced. According to Moore's law, the capacity of microchips is doubled every 18 months, meanwhile the production costs are halved. Presently, major activities of semiconductor lithography are concentrated on the wavelength 193 nm in conjunction with immersion techniques and on the implementation of next-generation lithography based on EUV sources operating in the wavelength range between 10 nm to 14 nm. Especially for the stepper systems operated with excimer lasers in the DUV spectral range, fluoride coatings are often employed for production of the optical components. The dominating coating process for these large aperture optics, which have a well defined surface figure and demand for reduced deposition temperature, is still conventional thermal deposition resulting in coatings with low mechanical and environmental stability. Therefore, IBS processes for the deposition of fluorides with improved stability, which have been investigated for many years [92Kol] within the framework of excimer laser development, have gained new interest [97Dij, 00Que, 04Yos]. For a realization of this so-called next-generation lithography in the EUV-range, mirrors for a wavelength of 13 nm are presently optimized within extensive international research initiatives. Besides reflectivity values of around 70%, a high thickness homogeneity has to be achieved over diameters of more than 300 mm. The present state in EUV thin-film technology can be described by a maximum reflectivity of approximately 69% for a Mo/Si Bragg reflector deposited by rf-magnetron sputtering. For the production of the masks for 13 nm lithography, IBS-concepts with broad ion-beam sources are presently studied in detail.

In the course of the development of new laser sources and frequency conversion systems with broad spectral emission ranges, an increasing demand for optical coatings with transfer properties specified over extended wavelength ranges can be observed. In this context the concept of "rugate filters", which are based on a continuous variation of the refractive index in the depth of the layer, is often discussed. As a major benefit, rugate filters can be designed to suppress higher-order stop bands in the spectra and to create spectral bands with extremely small variations in transmittance or reflectance. Furthermore, recent investigations indicate an advantage in the power-handling capability of these coatings in comparison to conventional optical coatings, which are composed of discrete layers with preset indices of refraction. For the realization of coatings with a defined variation of the refractive index, different approaches involving simultaneous evaporation from two independent deposition sources or ion-assisted deposition with varying composition of the reactive gas were investigated in thermal processes. Especially ion-beam sputter processes offer additional options for example by means of a target with two different material zones which is moved in the beam area allowing for the deposition of the pure materials and mixtures with compositions depending on the actual position of the target with respect to the ion beam [04Lee, 05Lap]. Recent results on these IBS-techniques document a clear prospect for the precise production of rugate filters and for the related progresses in optical thin-film technology.

Currently, ultra-short-pulse lasers are considered as an innovative tool for material processing, laser medicine, and biology, as well as the analysis and control of chemical reactions. For the development of these fs-lasers with ever increasing output power and improved beam parameters, special high-power broad-band coating systems are required. Besides an optical transfer function defined over an extremely extended spectral range, these coatings have to fulfill also demands with respect to their group delay dispersion for the compensation of dispersion effects in the laser systems. These coatings can be only achieved on the basis of complex designs, which are sensitive to thickness errors well below 0.1%. Improved production and controlling techniques are presently developed in order to adapt these coating types to the future demands of fs-technology.

In addition to the aspects outlined before, a variety of other challenges are imposed on thin-film technology by new developments in the laser field. Besides new crystal materials and micro-optics, coatings combining optical functions with other properties have to be mentioned. New approaches including rugate filter designs and advanced production techniques have to be considered to keep pace with the rapid progresses in laser technology.

References for 7.2

- 1900Ulb Ulbricht, R.: Die Bestimmung der mittleren räumlichen Lichtintensität durch nur eine Messung; Elektrotech. Z. **29** (1900) 595.
- 13Cob Coblentz, W.W.: The diffuse reflecting power of various substances; Natl. Bur. Stand. (U.S.) Bull. 9 (1913) 283–325.
- 64Bor Born, M., Wolf, E.: Principles of optics, Oxford: Pergamon Press, 1964.
- Gibbs, M., Wood, R.: Laser induced damage of mirror and window materials at 10.6 μm.
 Laser induced damage in optical material; NBS Spec. Publ. U. S. 462 (1976) 181–188.
- 78Wei Wei, D.T., Louderback, A.W.: US 4,142,958: Method for fabricating multi-layer optical films, 04, 1978.
- 80Lid Lidell, H.M.: Computer-aided techniques for the design of multilayer filters, Bristol: Adam Hilger Ltd, 1980.
- 81Wal Walker, T.W., Guenther, A.H., Nielsen, P.: Pulsed laser induced damage to thin film optical coatings, Part I: Experimental, Part II: Theory; IEEE J. Quantum Electron. 17 (1981) 2041–2065.
- 83Sei Seitel, S.C., Porteus, J.O.: Towards improved accuracy in limited scale laser damage testing via the "onset method"; NBS Spec. Publ. U. S. **6988** (1983) 502.
- 88Bau Baumeister, P.: Computer software for optical coatings; Photonics Spectra **22**(9) (1988) 143–148.
- 88DeB DeBell, G., Mott, L., von Gunten, M.: Thin film coatings for laser cavity optics. Proc. laser optics for intracavity and extracavity applications; Proc. SPIE (Int. Soc. Opt. Eng.) 895 (1988) 254–270.
- 88The Thelen, A.: Design of multilayer interference filters; New York, St. Louis, San Francisco, Auckland, Bogota, Hamburg, London, Madrid, Milan, Mexico, Montreal, New Delhi, Panama, Paris, Sao Paulo, Singapore, Sydney, Tokio, Toronto: McGraw-Hill Book Company, ISBN 0-07-063786-5, 1988.
- 92Gue Guenther, K.H.: Optical thin films deposited by energetic particle processes. Proc. SPIE (Int. Soc. Opt. Eng.) **1782** (1992) 344–355.
- 92Kol Kolbe, J., Schink, H.: Optical losses of dielectric VUV-mirrors deposited by conventional evaporation, IAD, and IBS. Guenther, K.H.: Thin Films for Optical Systems; Proc. SPIE (Int. Soc. Opt. Eng.) **1782** (1992) 435–446.
- 97Dij Dijon, J. et al.: High damage threshold fluoride UV mirrors made by ion beam sputtering. Proceedings of the 29th Annual Symposium on Optical Materials for High Power Lasers; Proc. SPIE (Int. Soc. Opt. Eng.) **3244** (1997) 406–416.
- 97ISO1 ISO 9022: Optics and optical instruments environmental test methods, Parts 1–21. International Organization for Standardization 1996, 1997.
- 97ISO2 ISO 10110: Preparation of drawings for optical elements and systems. Parts 1–12. International Organization for Standardization 1996, 1997.
- 97ISO3 ISO 11551: Test method for absorptance of optical laser components. International Standard, 1997 (revised version published in 2003), International Organization for Standardization.

126	References for 7.2
97Put	Puttik, V.K., Holm, R., Ristau, D., Natzschka, U., et al.: Continuous wave CO ₂ laser induced damage thresholds in optical components. Proceedings of the 29th symposium on laser induced damage in optical materials; Proc. SPIE (Int. Soc. Opt. Eng.) 3244 (1997) 188–198.
98Ris	Ristau, D.: Optics characterisation in thin film technology. Proceedings of the 4th In- ternational Workshop of Laser Beam and Optics Characterization, VDI-Verlag, 1998, p. 546–571
98Wil	Willamowski, U., Ristau, D., Welsch, E.: Measuring the absolute absorptance of optical laser components; Appl. Opt. 37 (36) (1998) 8362–8370.
99Ben	Bennett, J.M., Mattsson, L.: Introduction to surface roughness and scattering. Second adjuster Weshington D.C. Optical Society of America, 1000, ISBN 1-55752-600-5
99Ris	Ristau, D., Arens, W., Bosch, S., Duparré, A., Masetti, E., Jacob, D., Kiriakidis, G., Peiró, F., Quesnel, E., Tikhonravov, A.: UV-optical and microstructural properties of MgF ₂ -coatings deposited by IBS and PVD processes. Proceedings of the conference on advances in optical interference coatings; Proc. SPIE (Int. Soc. Opt. Eng.) 3738 (1999) 436–445.
00Que	Quensel, E. et al.: Optical and microstructural properties of MgF_2 UV coatings grown by ion beam sputtering. J. Vac. Sci. Technol. A $16(6)$ (2000) 2869–2876.
01Gro	Groß, T., Starke, K., Lappschies, M., Ristau, D.: Rapid prototyping of optical thin film filters. Proceedings of the 2nd international conference THE-coatings in manufacturing and engineering, Erice, Italy, Sept. 23–27, 2001, p. O2-1-O2-7.
01ISO2	ISO 11254: Test methods for laser induced damage threshold of optical surfaces. Part 1: 1-on-1 test, 2000; Part 2: S-on-1 test, 2001.
02Dob	Dobrowolski, J.A., Browning, St., Jacobson, M., Nadal, M.: 2001 OSA topical meeting on optical interference coatings (OIC2001): Manufacturing problem; Appl. Opt. 41 (16) (2002) 3039–3052.
02ISO	ISO 13696: Test method for radiation scattered by optical components; International Standard 2002, International Organization for Standardization.
04Lee	Lee, CC. et al.: Rugate filters made with composite thin film by ion beam sputtering. Proceedings of the Conference on Optical Interference Coatings, July 1st, 2004, OSA Technical Digest, p. THD10.
04Yos	Yoshida, T., Nishimoto, K., Etoh, K.: Fluoride antireflection multilayers with high trans- mittance for ArF excimer laser by ion beam sputtering method. Jpn. J. Appl. Phys. 43 (2B) (2004) L258–L260.
05Lap	Lappschies, M., Görtz, B., Ristau, D.: Optical monitoring of rugate filters. Proceedings of Advances in Optical Thin Films; Proc. SPIE (Int. Soc. Opt. Eng.) 5963 (2005) 1Z1–1Z9.
05Las	Laser-induced damage in optical materials. Proceedings of the Boulder damage sympo- sium. NIST Spec. Publ. and SPIE-Publications, 1969 to 2005. The publications of the first thirty years are available as CD-version: Laser induced damage in optical mate- rials. Collected papers. Published by the International Society of Optical Engineering, Washington, 98227-0010, 2005.
06ISO1	ISO 9211: Optical coatings, Parts 1–4, International Organization for Standardization 1994–2006.
06ISO2	Working groups of ISO/TC 172: Optics and photonics. International Organization for Standardization, website: http://www.iso.org/iso/en/stdsdevelopment/tc/.

7.3 Beam shaping

M. Scholl

7.3.1 Introduction

Beam shaping [97Aag, 99Sin, 00Dic] as a technical term is not clearly defined. In most relevant cases the task of beam shaping is as follows: generate by some optical system a spatial Power Density Distribution (PDD) with certain desired features from a given input beam – at least to some approximation. The requirements on the desired PDD result from the applications the beam is used for, e.g. adapting the intensity distributions for laser materials processing or splitting beams for multiplexing in optical communication. An important example is the generation of highly homogenous rectangular beam profiles for illumination of masks in lithography, for annealing applications or for digital light processors in video projection devices. In all these cases the transformation of the PDDs must be done with minimum loss of beam power to ensure the efficiency of the subsequent applications. This restriction usually rules out mask projection techniques which typically apply absorbing masks. Consequently, the characteristic feature of beam shaping is the use of pure phase masks to generate the desired output beam.¹

Sometimes the term beam shaping is used for the compensation of the ellipticity of astigmatic beams, too [99Sin, 00Dic]. Mostly this is done with ABCD-transformations realized with the help of quadratic phase elements, typically cylinder lenses. In a sense ABCD-transformations do not affect the internal structure of the PDD. In the spirit of the examples quoted above the term beam shaping applies only for non-quadratic phase transformations of the PDD which cannot be performed by ABCD-systems. In this contribution beam shaping will be used in this restricted sense in connection with non-quadratic phase elements only.

Furthermore, in situations where the typical time constant of the subsequent process allows fast sequential illumination of different parts of the PDD beam scanning may often be the simplest and most economical way to shape the required PDD. This should be checked first. Beam scanning, however, is a non-stationary technique which will not be considered here.

The most common stationary beam-shaping techniques which are used to generate a shaped PDD are beam transformation [96Hen, 97Aag, 00Dic, 03Ber], which shows the basic technique, beam integration [88Dic, 96Hen, 98Unn], combined beam integration with beamlet shaping [96Hen, 99Pah, 03Ber] and beam splitting [97Aag, 99Sin].

A typical arrangement for all these methods is shown in Fig. 7.3.1, which shows the basic idea of beam shaping: At the input plane (x_0, y_0) located at z = 0 the wavefront of the incoming beam is deformed by a phase element. The phase element forms a new wavefront in the plane (x', y') at $z' = 0 + \varepsilon$ in such a way that the desired PDD $I_1(x_1, y_1)$ builds up in the plane (x_1, y_1) at position z_1 behind some paraxial optical system. The phase element can be realized either by a mirror or by transmissive optics with an appropriate surface structure. The differences in the various beam-shaping techniques can be traced back to the properties of the structure function d of the phase element.

¹ However, using phase elements only it is difficult to generate abrupt intensity changes and highly structured profiles. For applications where high resolution is of prominent importance like in lithography lossy mask projection techniques are an indispensable tool.



Fig. 7.3.1. Typical arrangement used in beamshaping applications. The different shapes of the input and output PDDs $I_0(x_0, y_0)$, $I_1(x_1, y_1)$ are sketched.

For beam transformation a one-to-one mapping between points of the input and the output plane is assumed. A deterministic and invertible redistribution of the power in the PDD is inferred by the structure function. In beam transformation it is possible to perform an even more complete beam shaping by adjusting also the phase of the output beam with an appropriate second phase element. However, the input beam has to be known very precisely to determine the structure function. The technique is sensitive to deviations from the specified situation and to misalignment.

For the shaping of homogenous tophat profiles these disadvantages can be avoided by beam integration. The basic idea of beam integration by multifaceted mirrors or lenses is cutting the beam profile in segments called beamlets and superposing them in the output plane. The superposition acts as averaging of the beamlet profiles and a homogenous profile results. Due to the superposition of the beamlets a point in the output plane now has several preimages. In contrast to beam transformation there is no one-to-one correspondence between points in the input and output plane. Additional phase corrections in the output plane are not possible. Because of the segmentation the structure function d is periodic up to quadratic contributions.

Output PDDs different from tophats may be obtained by combining beam integration and beam transformation. This technique supposes that for each facet the input beamlet may approximately be treated as a plane-wave segment. The facet surfaces are structured appropriately so that the output profile is formed from this plane-wave segment by the facet. The superposition again averages possible differences and deviations in the output profiles.

Beam integration with beamlet shaping is much less sensible and may be applied if knowledge of the input beam is poor.

As coherence influences the result of beam shaping the design of the optics must take into account the coherence parameters [96Hen, 98Unn]. For beam integrators this crudely means that the facet width approximately should be of the order of the transverse coherence length of the beam.

Beam splitting differs from the previously quoted techniques in the structure of the output PDD. Usually an array of peaks with disjoint support has to be shaped. For a coherent beam this can be done with periodic structure functions. In this respect the situation is similar as in beam integration.

Manufacturing of beam-shaping elements requires methods to realize free form surfaces for mirrors or transmissive elements. Fast tool servo diamond turning machines are an appropriate tool for the manufacturing of some of the elements required for beam shaping. In many cases lithographic techniques for the production of diffractive optical elements with a proper surface structure offer an adequate manufacturing method.

For these most common techniques for the shaping of stationary beams the relevant design procedures are introduced in the following. The effects of coherence of the beam will be discussed briefly. Examples of techniques for the realization of phase elements are sketched. Detailed discussions are deferred to the literature [97Aag, 99Sin, 00Dic].

7.3.2 Beam-shaping techniques and design procedures

The beam-shaping techniques will be described here using the wave-optical scalar paraxial approximation [85Goo1, 86Sie, 87Bor, 05Hod].

The input beam at z = 0 in Fig. 7.3.1 can be described properly either by its complex field strength $E_0(\boldsymbol{x}_0) = \sqrt{I_0(\boldsymbol{x}_0)} \cdot \exp\{i\varphi_0(\boldsymbol{x}_0)\}\)$ for coherent beams or by its mutual intensity $J_0(\boldsymbol{x}_{01}; \boldsymbol{x}_{02}) = \langle E_0(\boldsymbol{x}_{01}) \cdot E_0^*(\boldsymbol{x}_{02}) \rangle$ for partially coherent beams [85Goo2, 87Bor]. The phase element in Fig. 7.3.1 is described in the complex amplitude transmission approximation [85Goo1, 86Sie, 87Bor, 97Aag, 05Hod]. Thus the field strength of the incoming beam is multiplied by

$$T = \exp\left\{-i\,k \cdot d(\boldsymbol{x})\right\} \tag{7.3.1}$$

with $k = 2\pi/\lambda$. The complex amplitude transmission approximation is admissible if the minimum feature size of d is approximately 10λ .

The optical structure function $d(\mathbf{x})$ characterizes the surface structure of the phase element. If the phase element is realized by a mirror the surface profile is given by Z = -2d, and for a transmissive element by Z = d/(n-1).

The optical system that transfers the beam from z' to the output plane at z_1 in Fig. 7.3.1 is assumed to be rotationally symmetric. It may be described by a 2×2 ABCD-matrix

$$\mathsf{M} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \quad \det(\mathsf{M}) = AD - BC = 1 \tag{7.3.2}$$

transforming geometrical-optical rays according to

$$\begin{pmatrix} \boldsymbol{x}_1 \\ \boldsymbol{u}_1 \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \cdot \begin{pmatrix} \boldsymbol{x}' \\ \boldsymbol{u}' \end{pmatrix}, \quad \text{or written formally} \quad \boldsymbol{\chi}_1 = \mathsf{M} \cdot \boldsymbol{\chi}', \qquad (7.3.3)$$

with

$$\boldsymbol{x} = \begin{pmatrix} x \\ y \end{pmatrix}, \quad \boldsymbol{u} = \begin{pmatrix} u_x \\ u_y \end{pmatrix}.$$
 (7.3.4)

The shorthand notation χ for the complete set of ray coordinates was introduced for later reference.

In (7.3.3), (7.3.4) (u_x, u_y) are the ray slopes in the corresponding directions. Obviously, the ABCD-system must not be an imaging system (B = 0) but should be as far as possible from imaging. Often the optical system is free propagation over a distance or a Fourier-transformation arrangement.

Without loss of generality the phase φ_0 of the incoming beam may be set to zero for further considerations. A non-vanishing φ_0 can be taken into account by a redefinition of d in (7.3.1). Furthermore, splitting the optical transformation in a phase transformation and a subsequent ABCD-transformation allows shifting quadratic terms between the two systems appropriately. This freedom can be used to adapt the range of d to an amount which is suitable for the manufacturing of the phase element.

In the following design principles are outlined which allow to determine the structure function d. Following the typical arrangement of Fig. 7.3.1 the connection between input and output beam must be derived from a transfer analysis as the first step of the design problem. From this connection an equation for the structure function d has to be set up and a solution has to be found.

7.3.2.1 Beam transformation

7.3.2.1.1 Coherent beams

For coherent beams the transfer analysis is performed in terms of the field strengths.

The input field strength $E_0(\mathbf{x}_0) = \sqrt{I_0(\mathbf{x}_0)}$ is transformed to the field strength $E'(\mathbf{x}')$ at the intermediate position z = z' by multiplication with the phase T of (7.3.1):

$$E'(\mathbf{x}') = \exp\left\{-\mathrm{i}\,k \cdot d(\mathbf{x}')\right\} \cdot E_0(\mathbf{x}') \,. \tag{7.3.5}$$

The subsequent transfer of E' to $E_1 = \rho_1 \exp{\{i \varphi_1\}}$ at $z = z_1$ is performed via the Collins integral [86Sie, 05Hod]. There the ABCD-matrix enters:

$$E_{1}(\boldsymbol{x}_{1}) = \int \mathrm{d}^{2}\boldsymbol{x}' \ K_{\mathsf{ABCD}}(\boldsymbol{x}_{1};\boldsymbol{x}') \cdot E'(\boldsymbol{x}') ,$$

$$K_{\mathsf{ABCD}}(\boldsymbol{x},\boldsymbol{x}') = \frac{-\mathrm{i}\,k}{2\,\pi\,B} \exp\left\{\mathrm{i}\,\frac{k}{2B} \left(A\boldsymbol{x}'^{2} - 2\boldsymbol{x}\cdot\boldsymbol{x}' + D\boldsymbol{x}^{2}\right)\right\} .$$
(7.3.6)

Finally, the complete transfer reads

$$E_1(\boldsymbol{x}_1) = \int d^2 \boldsymbol{x}_0 \ K_{\mathsf{ABCD}}(\boldsymbol{x}_1, \boldsymbol{x}_0) \cdot \exp\left\{-\mathrm{i} \, k \cdot d(\boldsymbol{x}_0)\right\} \cdot \sqrt{I_0(\boldsymbol{x}_0)}$$
$$= \sqrt{I_1(\boldsymbol{x}_1)} \cdot \exp\left\{\mathrm{i} \, \varphi_1(\boldsymbol{x}_1)\right\} \ . \tag{7.3.7}$$

The task in the design process is to solve this equation for the structure function d under the constraints that I_0 and I_1 are given. Note that the phase φ_1 of the output field is completely unspecified. This phase freedom must be used appropriately to allow for a solution for d [97Aag]. Mathematically it is not known if a strict solution exists at all. However, approximate solutions may be found via the Gerchberg–Saxton iteration procedure [72Ger, 96Hen, 97Aag] specified in Fig. 7.3.2. As the diagram is symmetric either the left upper or the right lower position may be chosen to start the iteration.

Position $z = z_0 = z'$		Position $z = z_1$
	forward Collins-integral	
$E_{(k)}' = \sqrt{I_0} \cdot \exp\left\{-\mathrm{i}k \cdot d_{(k)}\right\}$		$\tilde{E}_{1\;(k)} = \rho_{(k)} \cdot \exp\left\{\mathbf{i}\varphi_{1\;(k)}\right\}$
Apply contraint: Substitute $\rho'_{(k+1)} \rightarrow \sqrt{I_0}$ Keep $d_{(k+1)}$		Apply contraint: Substitute $\rho_{(k)} \rightarrow \sqrt{I_1}$ Keep $\varphi_{1\ (k)}$
$E'_{(k+1)} = \rho'_{(k+1)} \cdot \exp\left\{-i k \cdot d_{(k+1)}\right\}$	•	$\tilde{E}_{1(k)} = \sqrt{I_1} \cdot \exp\left\{ \mathbf{i} \varphi_{1(k)} \right\}$
	backward Collins-integral	

Fig. 7.3.2. Gerchberg–Saxton iteration procedure to solve (7.3.7) for d.

 $\begin{array}{c} {\rm Landolt-B\" ornstein} \\ {\rm New \ Series \ VIII/1A2} \end{array}$



Fig. 7.3.3. Solution of the design problem for the shaping of a homogenously illuminated cross (a) from a Gaussian input beam at $\lambda = 10.6 \ \mu\text{m}$. The iteration result for the surface structure $z^* = -2d$ of a mirror shaping the beam and deflecting it by 45° is shown in (b). (c) displays the simulation of the beam-shaping process [03Ber].

The quality of the approximation must be observed during the iteration. It can be specified by the transformation efficiency η :

$$\eta_1 = \frac{\int d^2 \boldsymbol{x}_1 \,\rho_{(k)}(\boldsymbol{x}_1) \cdot \sqrt{I_1(\boldsymbol{x}_1)}}{\sqrt{\int d^2 \boldsymbol{x}_1 \,\left(\rho_{(k)}(\boldsymbol{x}_1)\right)^2 \cdot \int d^2 \boldsymbol{x}_1 \,\left(I_1(\boldsymbol{x}_1)\right)}}$$
(7.3.8)

which measures the similarity of the shaped PDD with the required I_1 .² The transformation efficiency is used as a stopping criterion for the iteration in Fig. 7.3.2. An approximate solution for d is found if the iteration does no longer produce significant changes in η_1 .

Note that in beam transformation arbitrary phases in the output distribution can be arranged by a second phase element compensating the final phase φ_1 , as mentioned in the introduction.

An example for the solution of the design problem of beam transformation is shown in Fig. 7.3.3. Note that in the simulation of the beam-shaping result fluctuations like speckles appear. These may be due to vortex points with a helical structure of the phase comparable to a donut mode. With the help of a demodulation procedure these effects can be reduced to a minimum [03Ber]. If it is necessary to suppress the fluctuations even further contributions from higher spatial frequencies can be directed outside the area of interest within which the output PDD is specified. This amplitude freedom [97Aag] of the output PDD can result in power loss, however.

² Alternatively, an analogous expression η' can be computed at position z' from ρ' and $\int I_0$ due to the symmetry of Fig. 7.3.2.

7.3.2.1.2 Partially coherent beams and geometric optic approximation

For the discussion of beam transformation in geometric-optical approximation it is useful to consider partially coherent beams. The discussion can be performed most conveniently not directly with the mutual intensity $J(\boldsymbol{x}_1; \boldsymbol{x}_2) = \langle E(\boldsymbol{x}_1) \cdot E^*(\boldsymbol{x}_2) \rangle$ introduced above but in terms of the Wigner Distribution Function (WDF) $f(\boldsymbol{x}; \boldsymbol{u})$ [79Bas, 92Web]. The WDF is derived from J by

$$f(\boldsymbol{x};\boldsymbol{u}) = \frac{k^2}{(2\pi)^2} \int d^2 \boldsymbol{s} \exp\left\{-i\,k\,\boldsymbol{s}\cdot\boldsymbol{u}\right\} \cdot \left\langle E\left(\boldsymbol{x}+\frac{\boldsymbol{s}}{2}\right) \cdot E^*\left(\boldsymbol{x}-\frac{\boldsymbol{s}}{2}\right)\right\rangle \,, \tag{7.3.9a}$$

$$x = \frac{x_1 + x_2}{2}, \quad s = x_2 - x_1$$
 (7.3.9b)

with the vector notation of (7.3.4). By (7.3.9b) the coordinates \boldsymbol{x} and \boldsymbol{s} in (7.3.9a) are sum and difference of the two coordinates $\boldsymbol{x_1}, \boldsymbol{x_2}$ at which the field strengths E in (7.3.9a) are taken. The arguments $\boldsymbol{x}, \boldsymbol{u}$ of the WDF may be interpreted as ray coordinates, see (7.3.4). In the geometric-optical limit the expression $f(\boldsymbol{x}; \boldsymbol{u}) d^2 \boldsymbol{x} d^2 \boldsymbol{u} = dP$ gives the power fraction dP which is transported from the area element $d^2 \boldsymbol{x}$ at the point (x, y) into the space-angle element $d^2 \boldsymbol{u}$ in the direction (u_x, u_y) by the corresponding rays. In the geometric-optical approach f is the paraxial approximation of the radiance function of photometry.

The intensity I is found from the WDF (7.3.9a) as

$$I(\boldsymbol{x}) = \int d^2 \boldsymbol{u} f(\boldsymbol{x}; \boldsymbol{u}) .$$
(7.3.10)

Following Fig. 7.3.1 the beam-transfer analysis starts from an input WDF f_0 . Transit through the phase element is described by a convolution of f_0 with a transmission function τ . Explicitly this reads

$$f'(\mathbf{x}', \mathbf{u}') = f_0(\mathbf{x}', \mathbf{u}') \otimes_u \tau(\mathbf{x}', \mathbf{u}'), \qquad (7.3.11)$$

$$\tau(\boldsymbol{x}, \boldsymbol{u}) = \frac{k^2}{(2\pi)^2} \int d^2 \boldsymbol{s} \exp\left\{-i \, \boldsymbol{k} \, \boldsymbol{s} \cdot \boldsymbol{u}\right\} \, T\left(\boldsymbol{x} + \frac{\boldsymbol{s}}{2}\right) \cdot \, T^*\left(\boldsymbol{x} - \frac{\boldsymbol{s}}{2}\right) \,, \tag{7.3.12}$$

where the subscript u in (7.3.11) indicates that convolution is performed with respect to the variable u only.

The form (7.3.1) has to be used in the product $T \cdot T^*$ in (7.3.12). For sufficiently smooth structure functions d an expansion with respect to the variable s can be performed and only the lowest non-vanishing order needs to be kept. In this approximation the transmission function τ reads

$$\tau(\boldsymbol{x}, \boldsymbol{u}) = \delta\left(\boldsymbol{u} + \nabla d(\boldsymbol{x})\right) \tag{7.3.13}$$

with δ the Dirac delta function. Thus the convolution in (7.3.11) now merely acts as redefinition of the ray slope u:

$$\boldsymbol{u}_0 \to \boldsymbol{u}' = \boldsymbol{u}_0 - \nabla d(\boldsymbol{x}_0) \,. \tag{7.3.14}$$

This transformation property could have been expected in the geometric-optical approximation.

Transfer through the subsequent ABCD-system in Fig. 7.3.1 can be described without approximation using the ABCD-transformation properties of the WDF $f_1(\chi_1) = f'(\chi')$ or $f_1(\chi_1) = f'(M^{-1}\chi_1)$ and the transformation of the ray coordinates (7.3.3) [79Bas, 92Web, 05Hod]. Consequently, in this approach the transfer through the optical system of Fig. 7.3.1 is described completely in geometric-optical terms.

Combining the transformations and using (7.3.10) an integral expression for the output PDD in terms of the input WDF and the structure function d is obtained. The remaining integral can be done in a saddle-point or rotating-phase approximation [87Bor, 96Hen]. Together with the assumption of a vanishing phase φ_0 of the input beam this amounts to the use of $f_0(\boldsymbol{x}; \boldsymbol{u}) = I_0(\boldsymbol{x}) \cdot \delta(\boldsymbol{u})$ as input WDF. This shows that the divergence angle of the incoming beam must be smaller than the typical deflection angles $|\nabla d|$.

The transfer analysis finally yields the following equation for the structure function d

$$I_0(\boldsymbol{x}_0) = \det\left(\frac{\partial \boldsymbol{x}_1}{\partial \boldsymbol{x}_0}\right) \cdot I_1(\boldsymbol{x}_1) , \quad \boldsymbol{x}_1 = A \, \boldsymbol{x}_0 - B \, \nabla d(\boldsymbol{x}_0) \tag{7.3.15a}$$

or explicitly

$$\frac{I_0(\boldsymbol{x})}{I_1(A\boldsymbol{x} - B\,\nabla d(\boldsymbol{x}))} = \det(A\,\delta_{ij} - B\,\partial_i\partial_j d(\boldsymbol{x}))$$
$$= A^2 - AB\cdot(\partial_x^2 + \partial_y^2)\,d + B^2\cdot\left(\partial_x^2 d\,\partial_y^2 d - (\partial_x\partial_y d)^2\right)\,. \tag{7.3.15b}$$

This is a complicated nonlinear second-order partial differential equation for d. It may be solved by an iteration procedure [03Ber]. The equation may be interpreted in terms of the differential geometry of the surface described by d [03Ber].

The structure of the equations (7.3.15a) already results from the relation $I(\boldsymbol{x}_0)d^2\boldsymbol{x}_0 = I(\boldsymbol{x}_1)d^2\boldsymbol{x}_1$ together with the coordinate transformation in (7.3.15a). This asserts that there is a one-to-one mapping $\boldsymbol{x}_1(\boldsymbol{x}_0)$ which is assumed in beam transformation.

For problems which separate in the coordinates x, y, i.e. $I(x, y) = \tilde{I}(x) \cdot \hat{I}(y)$ and $d(x, y) = d_x(x) + d_y(y)$, equation (7.3.15a) splits in sets of separately integrable first-order ordinary differential equations

$$\tilde{I}_0(x_0) \,\mathrm{d}x_0 = \tilde{I}_1(x_1) \,\mathrm{d}x_1 \,, \tag{7.3.16}$$

$$x_1 = x_1(x_0) = A \cdot x_0 - B \cdot \frac{\mathrm{d}}{\mathrm{d}x_0} d_x(x_0)$$
(7.3.17)

and a similar set for the variable y. Equation (7.3.16) determines the coordinate mapping $x_1(x_0)$, equation (7.3.17) can be integrated to find the structure function d. The solutions of the equations for x, y solve the design problem in the case of separation. With the help of (7.3.16), (7.3.17) even analytic solutions can be found, e.g. the beam-shaping transformation of a Gaussian to a homogenous rectangular tophat distribution [97Aag].

7.3.2.1.3 Comparison of the geometric-optical solution with the solution from the coherent techniques

Figure 7.3.4 shows for a one-dimensional example the comparison of the geometric-optical solution with the solution from the coherent techniques shown above [96Hen].

As in the example of Fig. 7.3.4 the geometric-optical solution of the design problem often is a satisfactory approximation by itself. Alternatively, start functions for the iteration of Fig. 7.3.2 can be found from the geometric-optical techniques shown here.

7.3.2.2 Beam integration

Beam-integration techniques have been developed to solve the most prominent problem of beam shaping, i.e. to generate a highly homogenous rectangular tophat profile from an arbitrary input beam. The basic idea has been outlined in the introduction. Beam integration can be realized by many different systems like caleidoscopes, beam pipes or integrators. Only multifaceted beam integrators are discussed here as only these systems may be described in the arrangement of Fig. 7.3.1. The basic building element is a lens or a mirror with plane facets on a spherical enveloping



Fig. 7.3.4. One-dimensional example of beam shaping by beam transformation [96Hen]. The desired output PDD (a) is formed from a plane-wave input. In (b) the mirror surface function obtained by the geometric-optical approach I and by the Gerchberg–Saxton iteration II (curve shifted for convenience) are shown. The beam-shaping simulation for surface I is seen in (c), for surface II in (d).

surface. For discussion it is simpler to think of these elements as combination of a lenslet array and a focusing lens [96Hen, 98Unn, 99Pah]. Figure 7.3.5 shows examples of integrators composed of these elements.

Imaging integrators as in Fig. 7.3.5a are not strictly in the class described by Fig. 7.3.1. However, due to the imaging principle not only highly homogenous profiles but also very steeply ascending and falling edges can be obtained. Tophat profiles are approximated very well [98Unn].

The arrangement of non-imaging multifaceted beam integrators as in Fig. 7.3.5b strictly follows Fig. 7.3.1. The advantage that only one lenslet array is needed is bought at the expense of reduced edge steepness in the shaped PDD. However, for non-imaging integrators, Fig. 7.3.5b, it is possible to generalize the concept of beam integration to include beam transformation [96Hen] so that this integrator will be discussed.

A discussion of the non-imaging integrator in the framework of the geometric-optical approach of the previous section is of no use as it simply reproduces the basic expectation outlined in the introduction.

The transfer analysis through the system in the framework of the coherent approach is restricted to one transverse beam dimension for simplicity. The discussion will follow the arrangement of Fig. 7.3.5b [96Hen]. All length scales are normalized to the half facet width a, e.g. the transverse coordinate is $\xi = x/a$.



Fig. 7.3.5. Multifaceted beam integrators. (a) Imaging integrator or flies-eye integrator. The illumination array is imaged and superposed in the focal plane by the the imaging array. (b) Non-imaging integrator.

The transmission function of the lenslet array is the convolution of an array G of δ -functions with the transmission function $T_{\rm F}$ of the single facet respectively lenslet:

$$T(\xi) = G(\xi) \otimes T_{\rm F}(\xi) ,$$

$$G(\xi) = \sum_{n=-N}^{N} \delta(\xi - 2n) , \quad T_{\rm F}(\xi) = \operatorname{rect}(\xi) \cdot \exp\{-ik \cdot S_{\rm F}(\xi)\} .$$
(7.3.18)

Here the rect-function describes the boundaries of the facet. The facet or lenslet structure function $S_{\rm F}$ encodes its interior phase-transmission properties. A standard multifaceted integrator lens needs spheric lenslets $S_{\rm F} \propto \xi^2$. The transmission function of the single lenslet $T_{\rm F}$ becomes

$$T_{\rm F}(\xi) = \operatorname{rect}(\xi) \cdot \exp\left\{\mathrm{i}\,\pi\,N_{\rm F}\,\xi^2\right\} \tag{7.3.19}$$

with $N_{\rm F} = a^2/\lambda f$ the Fresnel number of the arrangement in Fig. 7.3.5b. So the lenslet curvature in (7.3.19) compensates the curvature of the subsequent focusing lens and thus mimics a lens with plane facets. The beam behind the lenslet array may be written as

$$E'(\xi) = E_0(\xi) \cdot (G(\xi) \otimes T_F(\xi)) .$$
(7.3.20)

The remaining step in the transfer analysis is the propagation through the focusing lens to the focal plane in Fig. 7.3.5b. Using (7.3.6) this amounts to a Fourier transformation \mathcal{F} and up to an inessential phase factor the output field E_1 reads

$$E_1(\xi) = \mathcal{F}[E_0 \cdot (G \otimes T_F)] = \mathcal{F}[E_0] \otimes (\mathcal{F}[G] \cdot \mathcal{F}[T_F]) \quad , \tag{7.3.21}$$

where the Fourier transform \mathcal{F} of some function $h(\xi)$ here is defined as

$$\mathcal{F}[h(\xi)] = \int \mathrm{d}\,\xi'\,h(\xi')\cdot\exp\left\{-\mathrm{i}\,2\,\pi\,N_{\mathrm{F}}\,\xi\cdot\xi'\right\}\,.$$

For the Fourier transform of the δ -array one finds

$$\mathcal{F}[G] = \sum_{n=-N}^{N} \exp\{-i 4\pi N_{\rm F} n \cdot \xi\}$$
(7.3.22)

expressing the interference of plane waves from the beamlets. The sum can be evaluated to an array of δ -functions. According to the bracket in (7.3.21) the peak array is modulated by the diffraction pattern of the facet transfer function. Due to the convolution each peak gets the form and width of the Fourier transform of the input field, cf. Fig. 7.3.6. The beam shaping solely shows in the modulation of the peak array.



Fig. 7.3.6. Output intensity of a beam integrator with plane facets with the structure according to (7.3.24). The Fresnel number is $N_{\rm F} = 24$, input beam is a donut mode [96Hen].

A necessary condition [96Hen] for the peak array to exist is that the distance $\Delta \xi$ of the peaks exceeds the peak diameter $2w_{\rm f}$, i.e. $w_{\rm f}/(\Delta \xi/2) < 1$, where $w_{\rm f}$ is the beam radius of $\mathcal{F}[E_0]$. Expressing $w_{\rm f}$, $\Delta \xi$ in terms of the parameters of the optical system and the input beam one finds

$$\frac{w_0}{a} \ge \frac{4}{\pi} \cdot M^2 \tag{7.3.23}$$

with w_0 the beam waist radius of the input beam and M^2 the beam propagation factor [92Web, 05Hod]. w_0/a in (7.3.23) roughly estimates the minimum number of facets that have to be illuminated by the input beam for the interference pattern to appear and the beam integration to be successful. This is also necessary for the assumption that within a facet the wavefront may be approximated by a plane wave.

The output power density of the beam integrator in Fig. 7.3.5b reads for coherent illumination [96Hen]

$$I_{\rm coh}(\xi) = \frac{1}{16} N_{\rm F}^2 \sum_{n=-N}^{N} \tilde{I}_0(\xi - n \cdot \Delta \xi) \cdot \tilde{I}_{\rm F}(n \cdot \Delta \xi)$$
(7.3.24)

with

$$\widetilde{I}_{0}(\xi) = \left|\mathcal{F}[E_{0}(\xi)]\right|^{2}, \quad \widetilde{I}_{\mathrm{F}}(\xi) = \left|\mathcal{F}[T_{\mathrm{F}}(\xi)]\right|^{2}, \quad \Delta \xi = \frac{1}{2N_{\mathrm{F}}}.$$
(7.3.25)

The first term in the sum in (7.3.24) describes an array of peaks each of the same form as the input beam would obtain when propagated through the system without the lenslet array. The peak distance is given by $\Delta \xi$. This array of peaks is modulated by the intensity pattern of the Fourier transform of the facet transfer function $T_{\rm F}$. For plane facets this is the near-field diffraction pattern of a slit or rectangular aperture in two dimensions. Figure 7.3.6 shows an example.

Although the derivation has been done for a coherent laser beam it can be shown that this structure also holds for partially coherent multimode beams [96Hen].

Note that the Fourier transform inverts the length scales of the input in (7.3.24): facets, i.e. the small scales of the input determine the envelope, that is the large scale of the output. The larger scale of the input beam diameter determines the small scale of the peaks. The large scale structure of the output PDD does not depend on the input PDD but only on the facets. As desired beam shaping becomes independent of the input beam.

In many applications with smoothing properties like materials processing the peaked structure of the output is unessential and the beam-shaping result is satisfactory. For plane-facet integrators the only free parameter which enters the transfer analysis is the Fresnel number $N_{\rm F} = a^2/\lambda f$. Thus the design problem is reduced to choosing a proper value for $N_{\rm F}$ and to adjusting a, f correspondingly. The choice of $N_{\rm F}$ may for example be dictated by the maximum peak distance $\Delta \xi$ (7.3.25) which is allowed by the materials processing application [96Hen].

7.3.2.3 Beam integration with beamlet shaping

The advantage of beam integration is that the beam shaping is independent of the input PDD. The advantage of beam transformation is found in the flexibility of the form of the output PDD. The aim of beam integration with beamlet shaping is to combine these advantages.

In the previous sections it has already been shown in which sense this can be realized. Following (7.3.24) the desired PDD is generated as the function $I_{\rm F}$ which modulates the peak array of the interfering beamlets. Thus each facet by itself must form a desired output PDD I_1 in the focal plane in Fig. 7.3.5b. The transmission function of the facet now reads

$$T_{\rm F}(\xi) = \exp\left\{-ik \cdot S_{\rm F}(\xi)\right\} = \exp\left\{i\pi N_{\rm F}\xi^2 - ik\,d_{\rm F}(\xi)\right\} .$$
(7.3.26)

The non-parabolic contribution $d_{\rm F}$ in the surface function $S_{\rm F}$ must be determined so that the second equation in (7.3.25) is satisfied with $I_1 = I_{\rm F}$. With the techniques of Sect. 7.3.2.1 an appropriate $d_{\rm F}$ can be found such that the desired PDD I_1 is formed from a plane-wave segment as input beam. In the computations the parameters A = 0, B = f, C = -1/f, D = 1 have to be used, cf. Fig. 7.3.5b. The beam transformation in Fig. 7.3.4 is an example for shaping of a single beamlet. Figure 7.3.7 shows the result of beam integration with beamlet shaping with superposed beamlets according to (7.3.24).

For two-dimensional beam shaping often the surface structure function has discontinuities at the facet boundaries. This is unsuitable for manufacturing. Discontinuities can be improved or even avoided, if appropriately curved facet boundaries are admitted [03Ber]. In this case the facets often are grouped in non-orthogonal lattices.

Beam integration with beamlet shaping can produce very satisfactory beam-shaping results even if deviations of the input beam from the specified situation occur. However, the crucial assumption in the design procedure is that the input beam for the beam shaping of each facet is a plane-wave segment. Note that different from the situation in pure beam transformation the phase of the input beam must not be assumed to vanish. It cannot be absorbed in a redefinition of the overall transmission function of the facetted phase element in Fig. 7.3.5b. It may be assumed, however, that the input beam phase is flat on a large scale. Thus there are no quadratic contributions, these could be absorbed in a redefinition of the focusing lens. The deviations caused by local phase front



Fig. 7.3.7. Computed output PDD for onedimensional line integration with beamlet shaping as shown in Fig. 7.3.4. The dense interference structure at wavelength $\lambda = 1.06 \ \mu m$ makes the PDD quasicontinuous.

deformations can be modeled by the assumption that the input plane waves incident on the facets vary slightly in the incidence angle. Due to the Fourier transform in (7.3.21), (7.3.24), (7.3.25) this shows as slight displacements of the PDDs from the various facets in the output PDD. Thus the desired profile and edges are flattened. The size of facets and the input beam have to be adapted properly to avoid this. As phase deformations may occur as result of partial coherence, too, the subject is discussed further in this framework.

7.3.2.4 Beam shaping and coherence

Like the stationary phase deformations discussed in the previous section statistical fluctuations degrade the quality of beam shaping, too. The induced partial coherence of the input beam may even spoil beam shaping.

As pointed out at the beginning of Sect. 7.3.2 a partially coherent laser beam may be described by its mutual intensity J. Often the mutual intensity can be expressed in terms of a quasihomogenous Schell model beam [77Car, 87Bor]:

$$J(\boldsymbol{x}_1; \boldsymbol{x}_2) = I(\boldsymbol{x}) \cdot \boldsymbol{\mu}(\boldsymbol{s}) \tag{7.3.27}$$

with \boldsymbol{x} and \boldsymbol{s} as defined in (7.3.9b). Here μ is the degree of coherence. The degree of coherence in many cases is described well by a Gaussian function with a width parameter σ which is the transverse coherence length. The model applies well not only for fiber-guided Nd:YAG laser beams [96Hen] or for excimer laser beams [98Unn, 03Ber] but is valid also for the description of beams with statistical phase fluctuations [96Hen]. If a quasihomogenous model beam is used as input beam for beam integration the resulting output intensity reads

$$I_{\rm pc}(\xi) = I_{\rm coh}(\xi) \otimes \mathcal{F}\left[\mu(\xi)\right] \tag{7.3.28}$$

with $I_{\rm coh}$ from (7.3.24). Remember that all lengths are normalized to half the facet width. Due to the convolution with the Fourier transform $\mathcal{F}[\mu]$ of the degree of coherence the peak array in the coherent intensity pattern is washed out. A criterion for this is in analogy to the derivation of (7.3.24) that the normalized width 2ε of $\mathcal{F}[\mu]$ should exceed the peak distance in the array, i.e.

$$\rho = \frac{\varepsilon}{\Delta\xi/2} > 1 . \tag{7.3.29}$$

 ρ is a smoothing parameter for the interference pattern. Expressing ε and $\Delta \xi$ in terms of the parameters of the optical system and the input beam the design rule that the ratio σ/a of coherence length and facet width should be of order unity is obtained.

Equation (7.3.29) only sets a lower limit on ρ . Partial coherence, however, not only washes out the peak array in the output PDD but also flattens edges and details of the coherent PDD. This is undesirable. An optimum design has to be found by optimizing the beam-shaping quality with respect to the parameter ρ . In the standard situation ε in (7.3.29) is fixed by the input beam, so $\Delta \xi$ has to be varied. This can be done by adding additional quadratic phase terms in the transmission function $T_{\rm F}$ (7.3.26) of the facet. A beam-shaping quality R may be defined in analogy to the beam-shaping efficiency (7.3.8). R has to be computed in dependence of ρ and the optimum value has to be identified for system design. Figure 7.3.8 shows the design procedure for the profile of Fig. 7.3.4 and Fig. 7.3.7 [96Hen].

Vice versa it is possible to alter the coherence properties of a beam to optimize beam shaping. This can be done with a phase-diffusing element in the input beam. Again the smoothing parameter ρ is used in the optimization procedure. Details are deferred to the literature [96Hen].

As has been seen partially coherent beams need a very careful design in adapting the system parameters like facet sizes and curvatures to the beam parameters. However, if the design rules are observed advanced beam-shaping applications like shown in Fig. 7.3.9 can be achieved.


Fig. 7.3.8. (a) Optimization of the output beam profile with respect to the smoothing parameter ρ . The intensity diagram corresponds to the maximum of the beam-shaping quality (R^6 for convenience). (b) Output beam profile for non-optimized value. (c) Measured output PDD of a line integrator manufactured according to optimization in (a). $\lambda = 1.06 \ \mu m$.

7.3.2.5 Beam splitting

In beam splitting the task is to generate from the input beam an array of beams. Often the output beams in the array are to have approximately equal intensities, but modulated arrays are required in some applications, too. In nearly all cases the input beam is coherent, often a Gaussian.

Usually this problem is thought of in terms of phase gratings with a nontrivial phase function within the slits. Up to different terminology this problem has already been discussed in Sects. 7.3.2.2 and 7.3.2.3. The peak array in the output PDD (7.3.24) exactly solves the task. The freedom of adding additional quadratic phase contributions to the facets can be used to adjust the peak distance in the array. As in Sect. 7.3.2.3 nontrivial phase functions in the facets can be used to adjust the array modulating function.

It must be mentioned that various other techniques for beam splitting or array illumination exist [99Sin]. However, they cannot be arranged as in Fig. 7.3.1 and will not be described here.



Fig. 7.3.9. Two-dimensional beam integration with beamlet shaping for an eximer laser [03Ber]. (a) Desired output PDD, an Archimedic spiral curve. (b) Facet surface profile designed with the geometric-optical approach and optimized with respect to the coherence properties of the input beam. (c) Measurement of the shaped PDD.

7.3.2.6 Manufacturing of beam-shaping elements

The phase elements of beam-shaping systems can be realized as reflecting, refracting or as diffracting elements with appropriately shaped surfaces. Although GRIN techniques are possible in principle a sufficiently precise adjustment of the local index of refraction is difficult.

Depending on the realization the manufacturing methods are chosen. The main requirement is that the manufacturing techniques have to be able to realize the free form surface given by the structure function d(x, y) in its full range with all its feature sizes with appropriate precision. This is usually $\lambda/10$ of the used wavelength or even better.

For all manufacturing techniques the dynamical range of the surface profile unfortunately must not be too large. Mostly the range is restricted to less than $10...50 \ \mu\text{m}$. It is therefore necessary in the design process to use all degrees of freedom to avoid high dynamical ranges. Most important the freedom to shift quadratic phase contributions to the optical system can be used.

Especially for the manufacturing of diffractive phase elements lithographic techniques can be used with success [99Pah, 99Sin]. For diffractive optical elements the surface profile given by the structure function is realized modulo the wavelength of the beam to be shaped. This corresponds



Fig. 7.3.10. Beam-shaping mirrors manufactured by diamond turning [03Ber]. Mirror (a) realizes the beam transformation to shape the cross profile discussed in Sect. 7.3.2.1. Mirror (b) realizes beam integration with beamlet shaping for the spiral shown in Fig. 7.3.9.

to a restriction of the phase to the range $0...2\pi$. Often this leads to feature sizes which are comparable to the wavelength. These are accessible to lithographic techniques. To realize the resulting surface structure both mask lithography and scanning lithography with electron or laser beams is used to expose a photo resist. In subsequent etching steps the surface structure can be generated either as a binary phase element or as blazed phase element. In binary phase elements [96Goe] the phase range $0...2\pi$ is divided in a number of levels and the profile given by the structure function is approximated by these levels. For low-level numbers binary phase elements suffer from low efficiency ($\eta = 41\%$ for 2 levels, $\eta = 95\%$ for 8 levels [99Sin]), high-level numbers make manufacturing expensive.

An alternative method is provided by gray-scale lithography [99Sin]. The photo resist is exposed to continuously varying intensity either by electron or laser beam scanning or by the use of gray scale mask. In the subsequent etching the gray scales are transferred as surface profiles.

A very prominent manufacturing technique for reflective and also for refractive phase elements is fast tool servo diamond turning [93Pyr]. In this technique the diamond tip of the turning tool is moved back and forth by piezoelectrical actuators during the revolution of the work piece. According to the manufacturing process the surface profile $Z(r, \varphi)$ is split as

$$Z(r,\varphi) = Z_{\rm rot}(r) + Z^*(r,\varphi) \tag{7.3.30}$$

in a rotationally symmetric part and a non-rotationally symmetric part. Remember that the surface profile Z and the structure function d are connected by Z = -d for mirrors and Z = d/(n-1) for transmissive elements. The dynamical range of Z^* is restricted as mentioned above.

Furthermore, the mechanics of the process does not allow profiling arbitrary steps in surface. Therefore this technique is mainly suited for smooth surface profiles not realized diffractively. Figure 7.3.10 shows examples of beam-shaping mirrors for applications in materials processing with CO_2 and excimer lasers.

However, the turning process induces a grating structure on the surface which may become disturbing for short wavelengths. Nevertheless for optics manufacturing feature sizes $< 5 \ \mu m$ and surface roughness of less than 1 nm has been demonstrated in selected examples.

7.3.3 Conclusion

Design procedures for the most common beam-shaping techniques have been outlined and a brief sketch of the manufacturing of beam-shaping elements has been given. The scalar paraxial approximation has been used throughout the analysis. Especially for the design of diffractive beam-shaping elements and for microoptical elements feature sizes in the order of the wavelength may occur. In the analysis of these elements the full vectorial diffraction theory has to be applied. This goes far beyond the scope of this contribution and may be found in the literature [97Kui]. Furthermore, a lot of other manufacturing techniques have been suggested. The discussion of these techniques is deferred to the literature [99Sin].

References for 7.3

- 72Ger Gerchberg, R.W., Saxton, W.O.: A practical algorithm for the determination of phase from image and diffraction plane pictures; Optik **35**(2) (1972) 237–246.
- 77Car Carter, W.H., Wolf, E.: Coherence and radiometry with quasihomogenous planar sources; J. Opt. Soc. Am. **47**(6) (1977) 785–796.
- Bastiaans, M.: Wigner distribution function and its application to first order optics;J. Opt. Soc. Am. 69 (1979) 1710.
- 85Goo1 Goodman, J.: Introduction to Fourier optics, New York, USA: John Wiley & Sons, 1985.
 85Goo2 Goodman, J.: Statistical optics, New York, USA: John Wiley & Sons, 1985.
- 86Sie Siegman, A.E.: Lasers, Mill Valley, CA: University Science Books, 1986.
- 87Bor Born, M., Wolf, E.: Principles of Optics, Oxford, UK: Pergamon Press, 1987.
- 88Dic Dickey, F.M., O'Neil, B.D.: Multifaceted laser beam integrators: general formulation and design concepts; Opt. Eng. 27 (1988) 999.
- 92Web Weber, H.: Propagation of higher-order intensity moments in quadratic-index media; Opt. Quantum Electron. **24** (1992) 1027.
- 93Pyr Pyra, M.: Nicht-rotationssymmetrische Laserspiegel Möglichkeiten und Grenzen der Herstellung, Dissertation, RWTH Aachen, Germany, 1993.
- 96Goe Goebel, B., Wang, L.L., Tschudi, T.: Multilayer technology for diffractive optical elements; Appl. Opt. **35** (1996) 4490–4493.
- 96Hen Henning, T.: Strahlformung für die Oberflächenbehandlung mit Hochleistungslasern, Dissertation, RWTH Aachen, Aachen, Germany: Shaker Verlag, 1996.
- 97Aag Aagedal, H., Wyrowski, F., Schmid, M.: Paraxial beam splitting and shaping, in: Diffractive optics for industrial and commercial applications, Turunen, J., Wyrowski, F. (eds.), Berlin, Germany: Akademie Verlag, 1997, 165–183.
- 97Kui Kuittinen, M., Turunen, J., Vahimaa, P.: Subwavelength-structured elements, in: Diffractive optics for industrial and commercial applications, Turunen, J., Wyrowski, F. (eds.), Berlin, Germany: Akademie Verlag, 1997, 303–323.
- 97Leg Leger, J.R.: Laser beam shaping, in: Microoptics: elements, systems and applications, Herzig, H.P. (ed.), London, UK: Taylor & Francis, 1997, 223–257.
- 98Unn Unnebrink, L.: Strahlformung partiell kohärenter Laserstrahlung: Optimierung eines Beleuchtungssystems für Maskenprojektionsverfahren mit UV-Lasern, Diplomarbeit, RWTH Aachen, Germany, 1998.
- 99Pah Pahlke, M.: Auslegung hocheffizienter, computergenerierter Beugungsstrukturen zur Strahlformung, Dissertation, Universität Stuttgart, Germany, 1999.
- 99Sin Sinzinger, S., Jahns, J.: Microoptics, Weinheim, Germany: Wiley-VCH Verlag, 1999.
- 00Dic Dickey, F.M., Holswade, S.C.: Laser beam shaping: Theory and techniques, New York, USA: Marcel Dekker Inc., 2000.
- 03Ber Bernges, J.: Auslegung und Anwendung nichtrotationssymmetrischer Optiken für die Laser-Materialbearbeitung, Dissertation, RWTH Aachen, Germany, 2003.
- 05Hod Hodgson, N., Weber, H.: Optical resonators fundamentals, advanced concepts, and applications, New York, USA: Springer, 2005.

143

8.1 Optical resonators

N. Hodgson

8.1.1 Introduction

The basic understanding of the properties of lasers and their radiation requires knowledge of the physics of optical resonators. The laser beam characteristics as well as efficiency and sensitivity against misalignment are determined mainly by the optical resonator. Without the resonator, light emitted by the active medium (solid-state material, gas, or liquid dye) differs insignificantly from the radiation emitted by thermal light sources. Characteristic properties of laser radiation, such as high brightness, high temporal and spatial coherence, and narrow bandwidth, are a result of the optical resonator feeding back the photons into the active medium.

The primary objective of laser resonator design is to create a specified beam quality and, at the same time, maximize output power and efficiency. Unfortunately, this design goal usually is accompanied by technical constraints such as limited space, low sensitivity to misalignment and pump power variations, small bandwidth, as well as by the practical challenge of low cost and reasonable manufacturability. Thus, it is critical to understand the physical properties of optical resonators in order to achieve a laser design that meets technical and commercial requirements.

Resonators currently used in commercial lasers can be divided into four categories: linear stable resonators, linear unstable resonators, ring resonators, and waveguide resonators.

8.1.1.1 Linear stable resonators

In linear stable resonators, light travels between an output coupling mirror and a high-reflecting end mirror. The resonator can be folded by additional high-reflecting mirrors. The main characteristic of stable resonators is the Gaussian beam being the fundamental mode, whereby the Gaussian beam diameter is determined by the geometrical dimensions of the resonator (length and mirror curvatures). Higher-order modes (Gauss-Laguerre modes or Gauss-Hermite modes) are powerorthogonal and their diameter as well as their beam-propagation factor increase as the mode order is increased. This implies that the beam quality of the stable resonator can be controlled by using intracavity apertures.

8.1.1.2 Linear unstable resonators

In linear unstable resonators, light travels between a confined high-reflecting mirror and a larger high-reflecting end mirror. The field is magnified with each roundtrip and the output coupling is produced by the consequent overfill of the confined high reflector. The annular near field generates side lobes in the far field due to hard-edge diffraction at the output coupler. The diffraction can be minimized by using Variable Reflectivity Mirrors (VRM) that exhibit maximum reflectance at the center and a gradual decrease towards the perimeter. In contrast to stable resonators, the diameter of unstable resonator modes is determined by the size of the output coupling mirror. The mode diameter does not depend on the mode order and since all transverse modes share the same gain area, only the lowest-loss mode is oscillating. This enables one to fill large-diameter active media in fundamental-mode operation by adapting the size of the output coupler to the size of the active medium. Due to the relatively high output coupling, unstable resonators can only be used efficiently in medium-to-high-gain lasers, i.e. high-power CO_2 lasers, excimer lasers, pulsed solid-state lasers, and military chemical lasers generating output powers in the MW range.

8.1.1.3 Ring resonators

In ring resonators, the beam is not incident normal to the mirror surfaces, but propagates in a ring-like pattern. In general, the field propagates in two directions but unidirectional operation can be achieved through the use of polarization rotators (stable resonators) or apertures (unstable resonators), both of which induce additional loss to the field of one propagation direction. The unidirectionality prevents spatial hole burning of the gain which facilitates single-mode operation (e.g. in dye lasers and solid-state lasers). In monolithic solid-state ring lasers, unidirectional operation provides very high stability and narrow linewidths.

8.1.1.4 Waveguide resonators

In waveguide resonators, the beam is guided within the active medium through grazing-incidence reflections off the outer surfaces and the feedback is achieved by resonator mirrors on both waveguide ends. The beam properties of waveguide lasers are determined by the degree of mode matching between the waveguide modes and the modes of the external stable resonator. Fundamental-mode operation is attained by adapting the free-space Gaussian beam diameter to the diameter of the fundamental waveguide mode. Waveguide resonators are used in laser systems that exhibit a power limitation due to inefficient cooling of the active media if conventional geometries (tube, rod) are used (e.g. sealed-off gas lasers). In addition to the improved thermal properties, the maximum output power of planar waveguides is scaled by area rather than length as in conventional geometries. Planar waveguides are widely used in sealed-off kW CO_2 and CO lasers.

8.1.1.5 Reviewing the basic properties of all optical resonators

The basic properties of all optical resonators can be reviewed without taking nonlinear interaction in the active medium into account. Mode structure, diffraction losses, misalignment sensitivity, and beam quality can be evaluated by finding the eigenmodes and eigenvalues of Fresnel-integrals for the round trip inside the resonator. Leaving the active medium out of the treatment is the classical approach to the subject since the gain generally only perturbs the physical properties of the resonator rather than completely changing them. Despite its more subtle impact the active medium is, in fact, important to assess because it affects the mode properties due to gain saturation, thermal lensing, and birefringence. Incorporation of the active medium into the theory is also required to calculate output power, optimum output coupling, and power stability. In general, the amplification of the electric field inside the active medium is described by coupled rate equations for the field and the inversion density, and the field propagation is performed by diffraction integrals. The laser properties can then be calculated using numerical methods.

8.1.2 Classification of optical resonators

An optical resonator usually consists of two spherical mirrors with radius of curvatures ρ_1 , ρ_2 separated by a distance L (Fig. 8.1.1). It is customary to replace each mirror by two lenses with focal length $f_i = \rho_i$ and locate the reference planes in between. The reference planes are the mirror surfaces, which means a plane wave in this representation is actually a spherical wave having the curvature of the mirror in real space. The ray-transfer matrix for the round trip inside an arbitrary optical resonator starting at mirror 1 is given by:

$$\mathsf{M}_{\text{RES}} = \begin{pmatrix} 2g_1g_2 - 1 & 2Lg_2\\ \frac{(2g_1g_2 - 1)^2 - 1}{2Lg_2} & 2g_1g_2 - 1 \end{pmatrix}$$
(8.1.1)

with

$$g_i = 1 - \frac{L}{\rho_i}$$
; $i = 1, 2$

The parameters g_1 and g_2 are called the g-parameters of the optical resonator. Note that the radius of curvature is positive for a concave mirror (focusing mirror) and negative for a convex mirror! According to (8.1.1), the imaging properties of two-mirror resonators are fully defined by the g-parameters and the mirror spacing L. We can further simplify the ray-transfer matrix M_{RES} by introducing the equivalent G-parameter $G = 2g_1g_2 - 1$:

$$\mathsf{M}_{\mathrm{RES}} = \begin{pmatrix} G & 2Lg_2 \\ \frac{G^2 - 1}{2Lg_2} & G \end{pmatrix} . \tag{8.1.2}$$

This matrix can be used to find spherical wavefronts at mirror 1 whose radii of curvature reproduce themselves after one roundtrip. In general, two reproducing spherical waves exist which are defined by the eigenvectors of the round-trip ray-transfer matrix. The corresponding eigenvalue is the magnification by which the lateral dimension of the wave is increased per round trip. The



Fig. 8.1.1. A round trip in an optical resonator can be described as a transit in an equivalent lens waveguide. Each mirror has been replaced by a lens pair with focal length $f_i = \rho_i$.

eigenvalues and the corresponding self-reproducing wave curvatures for matrix (8.1.2) read (note that R_a and R_b refer to the dotted planes in Fig. 8.1.1):

$$\mu_a = G + \sqrt{G^2 - 1}$$
, (8.1.3) $R_a = \frac{+2Lg_2}{\sqrt{G^2 - 1}}$, (8.1.4)

$$\mu_b = G - \sqrt{G^2 - 1}$$
, (8.1.5) $R_b = \frac{-2Lg_2}{\sqrt{G^2 - 1}}$. (8.1.6)

According to these eigensolutions we can distinguish between three different types of optical resonators:

1. |G| > 1, equivalent to $|g_1g_2| > 1$

We can find two spherical waves with real values for the radii of curvature $R_{a,b}$ reproducing themselves inside the resonator. If the beam diameter on mirror 1 is d, the spherical wave with radius of curvature R_a increases its diameter by $|\mu_a|$ every round trip (Fig. 8.1.2a). This eigensolution is called the divergent wave. The second eigensolution, the convergent wave, leads to a decrease of the beam diameter by $|\mu_b|$ per round trip. These resonators are referred to as unstable resonators.

- 2. |G| = 1, equivalent to $|g_1g_2| = 1$ Both radii of curvature R_a and R_b are infinite and both eigenvalues are equal to 1, which means that a plane wavefront is coming back planar after the round trip without change in diameter. These resonators are referred to as the resonators on the stability boundaries. The plane-plane resonator in Fig. 8.1.2b is one example of such a resonator.
- 3. |G| < 1, equivalent to $|g_1g_2| < 1$ The radii of curvature and the eigenvalues are all complex numbers. This result can only be interpreted in such a way that in geometrical optics no eigensolutions can be found in this type of resonator. Eigensolutions exist, but they can only be calculated by applying diffraction theory. These resonators are called *stable resonators*.



Fig. 8.1.2. The three types of optical resonators with spherical mirrors: (a) unstable resonator, (b) plane-plane resonator as an example of a resonator on the stability boundaries, (c) stable resonator having no eigensolutions in geometrical optics.



Fig. 8.1.3. The stability diagram of optical resonators with two spherical mirrors [97Hod].

Optical resonators with two mirrors can be visualized in a diagram where the g-parameters represent the coordinate axes (Fig. 8.1.3). This diagram is referred to as the g-diagram or, more often, as the stability diagram. A resonator defined by the g-parameters g_1 and g_2 is represented by a point in the stability diagram. Unfortunately, this representation is not unambiguous because the mirror spacing is not included. A special class of resonators are the confocal resonators with both mirrors having a common focal point. The confocal condition reads:

$$g_1 + g_2 = 2g_1g_2 . ag{8.1.7}$$

8.1.3 Unconfined stable resonators

The goal of a theoretical resonator description is to find steady-state distributions of the electric field on the resonator mirrors. These so-called eigensolutions $E_i(x, y)$ on mirror *i* will reproduce itself after each round trip (Fig. 8.1.4). The round trip in the resonator is described mathematically by the Kirchhoff integral equation [61Boy, 63Fox, 65Sie, 69Bau, 70Col]:

$$\gamma E_i(x_2, y_2) = i \frac{\exp[-ikL]}{2Lg_j \lambda_0} \int \int E_i(x_1, y_1) \cdot \exp\left[\frac{-i\pi}{2Lg_j \lambda_0} \left(G(x_1^2 + y_1^2 + x_2^2 + y_2^2) -2(x_1x_2 + y_1y_2)\right)\right] dx_1 dy_1$$
(8.1.8)

with

$$\begin{split} G &= 2g_1g_2 - 1; \, i, j = 1, 2; \, i \neq j, \\ L &= nL_0: \text{optical resonator length}, \\ L_0: \text{geometrical length}, \\ \lambda_0: \text{vacuum wavelength}, \\ k &= 2\pi/\lambda_0: \text{wave number}. \end{split}$$

The solutions to this integral equation represent the eigenmodes of the optical resonator. In general an infinite number of eigenmodes exist. The field distributions of the eigenmodes do not change their shapes but they might experience a decrease in amplitude due to diffraction losses. This is taken into account by the complex eigenvalue γ . The loss factor $V = \gamma \gamma^*$ represents the



Fig. 8.1.4. A steady-state field distribution E(x, y) must reproduce itself after each round trip. On the right side the round trip is presented as a transit in the equivalent resonator.

fraction of the initial power hitting the mirror after the round trip. The loss factor V is related to the loss ΔV via $\Delta V = 1 - V$. If both resonator mirrors are unconfined and perfectly reflecting, no power is lost during the round trip and the loss factor, therefore, is equal to 1.0. The condition $\gamma = 1.0$, referred to as the *resonance condition* of the optical resonator, yields the resonance frequencies.

8.1.3.1 Transverse mode structures

8.1.3.1.1 Gauss–Laguerre and Gauss–Hermite modes

Equation (8.1.8) can be solved analytically yielding an infinite number of eigensolutions. Which of these eigensolutions will actually be observed in the resonator depends on the geometry and the size of the mirrors. In reality, the mirrors will have a finite size with a shape that usually is round or rectangular. This boundary condition is taken into account by choosing those eigensolutions that exhibit circular or rectangular symmetry, respectively. The solutions to (8.1.8) for the two symmetries are described below (the index i denotes the mirror) [61Fox, 86Sie, 88Mil], see also Vol. VIII/1A1, Sect. 3.1.3.3:

(a) circular symmetry (Gauss-Laguerre modes):

$$E_{p\ell}^{(i)}(r,\Phi) = E_0 \left[\frac{\sqrt{2} r}{w_i}\right]^{\ell} L_{p\ell} \left(\frac{2r^2}{w_i^2}\right) \exp\left[\frac{-r^2}{w_i^2}\right] \cdot \begin{cases} \cos(\ell\Phi)\\\sin(\ell\Phi) \end{cases}, \tag{8.1.9}$$

$$\gamma = \exp\left[ik\left(2L - \frac{\lambda_0}{\pi}(2p + \ell + 1)\operatorname{acos}\sqrt{g_1g_2}\right)\right]$$
(8.1.10)

with

 $L_{pl}(t)$: Laguerre polynomial of order p, l; p, l: integer, r, Φ : radial and azimuthal coordinate, $k = 2\pi/\lambda_0$: wave number; λ_0 : wavelength in vacuum, $L = nL_0$: optical mirror spacing.

The Laguerre polynomials can be found in mathematical handbooks [64Abr, 66Mag]. For low orders p, l they read:

$$\begin{split} &L_{0\ell}(t) = 1 , \\ &L_{1\ell}(t) = \ell + 1 - t , \\ &L_{2\ell}(t) = \frac{(\ell+1)(\ell+2)}{2} - (\ell+2) t + \frac{1}{2} t^2 , \\ &L_{3\ell}(t) = \frac{(\ell+1)(\ell+2)(\ell+3)}{6} - \frac{(\ell+2)(\ell+3)}{2} t + \frac{(\ell+3)}{2} t^2 - \frac{1}{6} t^3 ; \end{split}$$

$$E_{mn}^{(i)}(x,y) = E_0 \exp\left[\frac{-(x^2+y^2)}{(w_{00}^{(i)})^2}\right] H_m\left(\frac{\sqrt{2} x}{w_{00}^{(i)}}\right) H_n\left(\frac{\sqrt{2} y}{w_{00}^{(i)}}\right) , \qquad (8.1.11)$$

$$\gamma = \exp\left[ik\left(2L - \frac{\lambda_0}{\pi}(m+n+1)\operatorname{acos}\sqrt{g_1g_2}\right)\right] \,. \tag{8.1.12}$$

The Hermite polynomials H_m can also be looked up in mathematical handbooks. For low-order numbers m, n they read:

$$\begin{split} H_0(t) &= 1 ,\\ H_1(t) &= 2 t ,\\ H_2(t) &= 4 t^2 - 2 ,\\ H_3(t) &= 8 t^3 - 12 t ,\\ H_4(t) &= 16 t^4 - 48 t^2 + 12 ,\\ H_5(t) &= 32 t^5 - 160 t^3 + 120 t . \end{split}$$

In both symmetries the loss factor V is equal to 1.0 since the mirrors are not limited by an aperture and, consequently, no power can leak out of the resonator. Figures 3.1.6 and 3.1.7 in Vol. VIII/1A1 present intensity distributions as a function of the order numbers p, l and m, n, calculated with (8.1.9) and (8.1.11), respectively.

The steady-state electric field distribution is characterized by the indices plq and mnq. In rectangular symmetry the first two indices represent the number of nodal lines of the intensity distribution in the corresponding direction. In circular symmetry the intensity distributions exhibit p radial and l azimuthal nodes at which the intensity is equal to zero. The index q represents the number of half wavelengths fitting into the mirror spacing (longitudinal mode order).

A single-frequency steady-state field distribution oscillating inside the resonator is called an eigenmode of the resonator. The eigenmodes are characterized by the transverse mode structure (transverse mode index p, l or m, n) and the axial mode order q. The notations for the eigenmodes are TEM_{plq} and TEM_{mnq} where the abbreviation TEM represents the fact that the electric and the magnetic field vectors are perpendicular to each other and to the wave vector \mathbf{k} (Transverse Electro Magnetic). This is not entirely true since the diffraction generates small field components in the direction of the wave propagation. Only in the limit of large beam radii (large Fresnel numbers) is the field truly transverse. The notation TEM was adopted from the modes of waveguides which in fact are transversal. Although not physically correct, it is customary to refer to the modes of open resonators as TEM modes as well. In general, the axial mode index q is not used and the mode structure is specified by the order numbers p, l and m, n (q is on the order of 10^6).

The lateral extent of the eigenmodes on mirror i is determined by the beam radius $w_{00}^{(i)}$ of the TEM₀₀ mode. The beam radii on the two mirrors depend on the mirror spacing and the g-parameters of the resonator:

$$w_{00}^{(i)^2} = \frac{\lambda L}{\pi} \sqrt{\frac{g_j}{g_i(1 - g_1 g_2)}} ; \quad i, j = 1, 2 ; \quad i \neq j .$$
(8.1.13)

As can be seen in Fig. 8.1.5, the size of the intensity distributions increases as the mode order is increased. In both symmetries the TEM_{00} mode has the same shape (Figs. 3.1.6 and 3.1.7 in Vol. VIII/1A1); the intensity distribution is Gaussian. At a distance $r = w_i$ from the center of gravity, the intensity has decreased by a factor of $1/e^2$ and 86.5% of the beam power is contained within the corresponding circle. The lateral extent of the TEM_{00} mode is therefore defined by the beam radius w_i , also referred to as the Gaussian beam radius. The TEM_{00} mode is the mode with the smallest size that can oscillate in a stable resonator. It is generally referred to as the



Fig. 8.1.5. Measured (solid line) and calculated (dashed line) one-dimensional intensity distributions of transverse modes TEM_{mn} in rectangular symmetry. The horizontal lines indicate the beam radii w_{m0} defined by the second intensity moments.

fundamental mode or the Gaussian beam. For transverse modes of higher order pl or mn, the beam radii are defined via the second-order intensity moments. This definition of the beam radii enables one to calculate the propagation of arbitrary field distributions through ABCD-type optics by applying the generalized ABCD law. The beam radii on mirror i read [79Nem, 80Car]:

circular symmetry:

$$w_{pl}^{(i)} = w_{pl}^{(i)} \sqrt{2p + l + 1} , \qquad (8.1.14)$$

rectangular symmetry:

$$w^{(i)} - w^{(i)}\sqrt{2m+1}$$
 (*x*-direction) (8.1.15)

$$w_{mn} = w_{00} \sqrt{2m + 1} \quad (x \text{-diffection}),$$
 (8.1.13)

$$w_{mn}^{(i)} = w_{00}^{(i)} \sqrt{2n+1} \quad (y\text{-direction}) .$$
 (8.1.16)

The cross-sectional area of the Gaussian beam radius scales linearly with the wavelength and the resonator length. It is for this reason that CO_2 lasers with their large emission wavelength of 10.6 µm exhibit relatively large fundamental mode diameters on the order of 10 mm. On the other hand, diode lasers have very small Gaussian beam radii on the order of µm due to their short mirror spacing in the sub-mm range. Furthermore, at a fixed wavelength and a fixed resonator length, the Gaussian beam radius will decrease as the resonator design is chosen closer to a stability limit in the g-diagram. This fact is visualized in Fig. 8.1.6 which presents curves of constant cross-sectional area of the Gaussian beam at mirror 1. The beam radius at mirror 2 can be obtained from this graph by switching g_1 and g_2 .

The transverse modes that are actually observed in a laser resonator are mainly determined by the size of the mirrors. The resonator mirrors are generally limited either by apertures or by the active medium itself. Only those transverse modes whose beam radii are smaller than the radii of the mirror apertures can be observed. If the gain of the laser medium is increased, more transverse modes will, however, reach the laser threshold. As a rule of thumb, a higher-order transverse mode will oscillate if the radius of the aperture a is greater than $0.9 \dots 1.0$ times the beam radius. Fundamental-mode operation requires an aperture radius a of $1.0 \dots 1.4$ times the Gaussian beam radius. Upper values and lower values correspond to low-gain and to high-gain active media, respectively.

Lasers can be forced to oscillate only in the fundamental mode by inserting apertures into the resonator with a diameter close to the diameter of the Gaussian beam at the aperture plane. If the apertures are considerably larger than the Gaussian beam, all higher transverse modes fitting into the aperture will oscillate. It is for this reason that in multimode lasers one does generally not observe the characteristic intensity distributions of Fig. 8.1.7. Since all of these modes exhibit virtually no loss, they are oscillating simultaneously resulting in a more or less homogenous beam



Fig. 8.1.6. Curves of constant cross-sectional area of the Gaussian beam at mirror 1 in the stability diagram. The cross-sectional area is normalized to $\lambda L/\pi$. The beam radius at mirror 2 can be obtained by switching g_1 and g_2 .

Fig. 8.1.7. Photographs of intensity distributions of different transverse modes at the output coupling mirror of a HeNe laser with a stable resonator. Moveable intracavity cross-wires were used to select modes with rectangular symmetry. The lower right photograph shows the lowestorder donut mode (without cross-wires) [97Hod].

profile. The more modes that participate in this process the more homogenous the laser beam becomes (Fig. 8.1.8). This multimode behavior is supported by the fact that different modes exhibit their intensity maxima in different areas of the active medium. The gain at the nodal lines of the intensity distribution is not depleted and can then be used by a different mode that has its intensity peaks in these vacant areas.



Fig. 8.1.8. Photograph of the beam profile of a Nd:YAG rod laser ($\lambda = 1.064 \ \mu m$) with a stable resonator in multimode operation (maximum of 2p+l+1 is about 60) [97Hod].

8.1.3.1.2 Hybrid modes

In circularly symmetric laser resonators one can quite often observe intensity distributions of modes that exhibit an annular intensity profile with almost zero intensity in the center. These beam profiles are generated by a superposition of two circularly symmetric transverse modes of the same order pl which are both linearly polarized and oscillate rotated by an angle of 90° with respect to each other (Fig. 8.1.9). There are four different ways to combine two linearly polarized modes resulting in different polarization states of the sum mode. The superposition always yields the same annular intensity profile with p + 1 rings with maximum intensity and a characteristic hole in the center. It is quite obvious why these modes are called donut modes. Sometimes they are also referred to as hybrid modes. Hybrid modes are marked by an asterisk next to the mode order numbers. Due to the different polarizations, the two transverse modes do not interfere and the intensity distribution of the sum is given by the sum of the individual mode intensity profiles. Application of (8.1.9) yields for the radial intensity distribution of hybrid modes at mirror i:

$$I_{p\ell^*}(t) = I_0 t^{\ell} \exp[-t] [L_{p\ell}(t)]^2$$
(8.1.17)

with

$$t = 2 \left[r/w_{00}^{(i)} \right]^2$$



Fig. 8.1.9. Hybrid modes are generated by the superposition of two linearly polarized circularly symmetric transverse modes of the same order pl which oscillate rotated by 90° with respect to each other. The graph shows two of the four possible ways to superimpose the mode structures. The lower graphs present the radial intensity distributions of the three lowest-order hybrid modes TEM_{01*}, TEM_{11*}, and TEM_{21*}.

8.1.3.2 Beam propagation of stable resonator modes

8.1.3.2.1 Fundamental mode

The Gaussian beam radius $w_{00}^{(i)}$ at mirror *i* determines the lateral dimension of the mode. For stable resonators, the intensity distributions at any plane inside or outside the resonator exhibit the same shape as the distributions on the mirrors; only the beam radius changes with the propagation.

For the fundamental mode, or Gaussian beam, the beam radius as a function of the distance can be calculated by using the ABCD law. If z denotes the distance along the optical axis from the position of the beam waist with radius w_0 , the Gaussian beam radius $w_{00}(z)$ inside the resonator reads (Fig. 8.1.10):

$$w_{00}(z) = w_0 \sqrt{1 + \left(\frac{z}{z_0}\right)^2} \tag{8.1.18}$$

with

$$w_0^2 = \frac{\lambda L}{\pi} \frac{\sqrt{g_1 g_2 (1 - g_1 g_2)}}{|g_1 + g_2 - 2g_1 g_2|} \quad (waist \ radius) , \qquad (8.1.19)$$

$$z_0 = \frac{\pi w_0^2}{\lambda} \quad (Rayleigh \ range) , \qquad (8.1.20)$$

$$L_{01} = L \frac{(1-g_1)g_2}{|g_1+g_2-2g_1g_2|} \quad (waist \ position) \ . \tag{8.1.21}$$

At the distance of one Rayleigh range z_0 from the position of the beam waist the beam radius has increased by a factor of $\sqrt{2}$, which means that the cross-sectional area of the beam has doubled. The waist position L_{01} is the distance of the minimum beam radius (beam waist) from mirror 1. If L_{01} is positive the waist is located to the right of mirror 1 (as shown in Fig. 8.1.10), for a negative L_{01} the waist is found to the left of mirror 1.



Fig. 8.1.10. Propagation of a Gaussian beam in a stable resonator.

The Gaussian beam is fully determined by the beam radius, the Rayleigh range, and the location of the beam waist. The divergence angle (half cone angle) θ_0 is obtained from these quantities by using the relation:

$$\theta_0 = \frac{\lambda}{\pi \, w_0} = \frac{w_0}{z_0} \,. \tag{8.1.22}$$

The beam parameter product $w_0\theta_0$ is a constant of the Gaussian beam as long as the beam propagates through ABCD-type optical systems (aberration-free optics). By combining (8.1.19) and (8.1.22), the angle of divergence θ_0 for a general stable resonator is given by:



Fig. 8.1.11. Curves of constant angle of divergence θ_0 of the Gaussian beam in the *g*-diagram. The parameter is the normalized angle of divergence $\theta_0 \sqrt{(L/\lambda)}$, which is 1.0 for the confocal resonator.

$$\theta_0^2 = \frac{\lambda}{\pi L} \frac{|g_1 + g_2 - 2g_1g_2|}{\sqrt{g_1g_2(1 - g_1g_2)}} \,. \tag{8.1.23}$$

The normalized angle $\theta_0 \sqrt{(L/\lambda)}$ depends only on the g-parameters of the resonator (Fig. 8.1.11). For a constant resonator length L, the angle of divergence becomes very small if the resonator is chosen close to the hyperbola in the first quadrant. The corresponding beam waist radius w_0 is, of course, very large for these resonators since the beam parameter product $w_0\theta_0$ is a constant. The largest angles of divergence are found for negative g-parameters at the stability limit $g_2 = 1/g_1$, where the concentric resonators are located. Since the large angles of divergence are generated by extremely small intracavity beam waists, these resonators have a small mode volume and have therefore found only limited application in laser systems.

The radius of curvature $R^{(i)}$ of the Gaussian beam at mirror *i* is always equal to the radius of curvature ρ_i of the mirror. The general expression for the radius of curvature *R* as a function of the propagation distance *z* reads:

$$R(z) = z_0 \left[\frac{z}{z_0} + \frac{z_0}{z} \right] .$$
(8.1.24)

It can be easily verified that the mirror surfaces indeed represent surfaces of constant phase of the Gaussian beam by setting $z = L_{01}$ and $z = L - L_{01}$ in (8.1.24). The radius of curvature exhibits a minimum at the Rayleigh range z_0 . The nonlinearity of R(z) is the reason for the deviation of the imaging condition of Gaussian beams from Newton's law of geometrical optics. The general expression for the electric field of a Gaussian beam as a function of the distance z from the waist is given by:

$$E(x, y, z) = \frac{E_0}{\sqrt{1 + (z/z_0)^2}} \exp\left[\frac{-(x^2 + y^2)}{w_{00}^2(z)} - \frac{\mathrm{i}k(x^2 + y^2)}{2R(z)}\right] \exp\left[-\mathrm{i}\operatorname{atan}\left(\frac{z}{z_0}\right)\right] \exp[\mathrm{i}kz] .$$
(8.1.25)

The last term is referred to as the Gouy phase shift. This additional phase is the reason why the resonance frequency has an additional term that depends on the g-parameters of the resonator. By introducing the q-parameter of the Gaussian beam:

$$\frac{1}{q(z)} = \frac{1}{R(z)} - \frac{i\lambda}{\pi w_{00}^2(z)}$$
(8.1.26)

the field distribution of a Gaussian beam can be written as:

$$E(x, y, z) = \frac{E_0}{\sqrt{1 + (z/z_0)^2}} \exp\left[-\frac{ik(x^2 + y^2)}{2q(z)}\right] \exp\left[-i \operatorname{atan}\left(\frac{z}{z_0}\right)\right] \exp[ikz] .$$
(8.1.27)

Propagation of the Gaussian beam through ABCD-type optics can be accomplished by using the ABCD-law for the *q*-parameter:

$$q_2 = \frac{Aq_1 + B}{Cq_1 + D} , \qquad (8.1.28)$$

where the index 1 and 2 denote the two reference planes between which the propagation occurs. A transverse distribution of the form

$$E(x_1, y_1) = E_0 \exp\left[-\frac{\mathrm{i}k(x_1^2 + y_1^2)}{2q_1}\right]$$
(8.1.29)

transforms into

$$E(x_2, y_2) = \frac{E_0}{A - B/q_1} \exp\left[-\frac{\mathrm{i}k(x_2^2 + y_2^2)}{2q_2}\right] . \tag{8.1.30}$$

8.1.3.2.2 Higher-order modes

The treatment of the propagation of higher-order modes in the resonators becomes quite simple given the preceding detailed discussion of the propagation of Gaussian beams. In both circular and rectangular symmetries the propagation of the fundamental mode and of the higher-order modes is similar. If w_0 and θ_0 denote the waist radius and the angle of divergence (half angle) of the fundamental mode, the beam radius as a function of the propagation distance from the waist position reads as follows:

(a) circular symmetry:

$$w_{p\ell}(z) = w_0 \sqrt{2p + \ell + 1} \sqrt{1 + \left(\frac{z}{z_0}\right)^2} \quad (beam \ radius \ at \ location \ z) \ , \tag{8.1.31}$$

$$w_{p\ell} = w_0 \sqrt{2p + \ell + 1} \quad (waist \ radius) , \qquad (8.1.32)$$

$$\theta_{p\ell} = \theta_0 \sqrt{2p + \ell + 1} \quad (angle \ of \ divergence) \ . \tag{8.1.33}$$

(b) rectangular symmetry:

$$w_m(z) = w_0 \sqrt{2m+1} \sqrt{1 + \left(\frac{z}{z_0}\right)^2} \quad (beam \ radius \ at \ location \ z) \ , \tag{8.1.34}$$

$$w_m = w_0 \sqrt{2m+1} \quad (waist \ radius) , \qquad (8.1.35)$$

$$\theta_m = \theta_0 \sqrt{2m+1}$$
 (angle of divergence). (8.1.36)

The Rayleigh range is the same as for the Gaussian beam, see Fig. 8.1.12:

$$z_0 = \frac{w_m}{\theta_m} = \frac{w_0}{\theta_0} = \frac{\pi \, w_0^2}{\lambda} \,. \tag{8.1.37}$$



Fig. 8.1.12. Beam propagation of higher-order modes. All modes have the same Rayleigh range. At any plane the beam radii are $\sqrt{2p + \ell + 1}$ larger than the Gaussian beam radius.

The beam parameter product $w\theta$ is a constant of the beam with

$$w_{p\ell}\theta_{p\ell} = (2p+\ell+1)\frac{\lambda}{\pi} = M^2\frac{\lambda}{\pi},$$
(8.1.38)

$$w_m \theta_m = (2m+1)\frac{\lambda}{\pi} = M^2 \frac{\lambda}{\pi}$$
 (8.1.39)

The term M^2 is referred to as the beam propagation factor of the mode. In a laser resonator, the maximum propagation factor corresponding to the highest-order transverse mode determines the beam quality of the laser beam. If N apertures with radii a_i are located inside the resonator at positions $z_1 \dots z_N$ and $w_{00}(z_i)$ are the beam radii of the Gaussian beam at the apertures, the maximum propagation factor, to a good approximation, can be calculated using the relation

$$M_{\max}^2 = \min\left(\left[\frac{a_1}{w_{00}(z_1)}\right]^2, \left[\frac{a_2}{w_{00}(z_2)}\right]^2, \dots, \left[\frac{a_N}{w_{00}(z_N)}\right]^2\right),$$
(8.1.40)

where "min" is the minimum term in the group.

The electric field distribution of individual transverse modes as a function of the propagation distance z reads:

(a) in circular symmetry:

$$E(r,\Phi,z) = \frac{E_0}{\sqrt{1+(z/z_0)^2}} \exp\left[\frac{-r^2}{w_{00}^2(z)} - \frac{\mathrm{i}kr^2}{2R(z)}\right] \left[\frac{\sqrt{2}r}{w_{00}(z)}\right]^c L_{p\ell} \left[\frac{2r^2}{w_{00}(z)^2}\right] \begin{cases} \cos(\ell\Phi) \\ \sin(\ell\Phi) \end{cases} \\ \times \exp\left[-\mathrm{i}(2p+\ell+1) \operatorname{atan}\left(\frac{z}{z_0}\right)\right] \end{cases},$$
(8.1.41)

- 0

(b) in rectangular symmetry:

$$E(x, y, z) = \frac{E_0}{\sqrt{1 + (z/z_0)^2}} \exp\left[\frac{-(x^2 + y^2)}{w_{00}^2(z)} - \frac{ik(x^2 + y^2)}{2R(z)}\right] H_m\left[\frac{\sqrt{2} x}{w_{00}(z)}\right] H_n\left[\frac{\sqrt{2} y}{w_{00}(z)}\right] \times \exp\left[-i(m+n+1) \operatorname{atan}\left(\frac{z}{z_0}\right)\right].$$
(8.1.42)

The beam radius $w_{00}(z)$ and the radius of curvature R(z) are those of the Gaussian beam. The Gouy phase shift (last term) generates the frequency difference between modes of different transverse order.

By introducing the q-parameter

$$\frac{1}{q(z)} = \frac{1}{R(z)} - \frac{i\lambda M^2}{\pi w_{p\ell}^2(z)}$$
(8.1.43)

160

the propagation of higher-order modes from a plane 1 to a plane 2 through ABCD-type optics can be calculated by using the ABCD-law for the q-parameter (8.1.28).

8.1.3.3 Beam quality and resonator parameters

The magnitude of the beam parameter product of a stable laser resonator depends on the number of transverse modes oscillating. The beam propagation factor M^2 is determined by the radius and the location of the smallest aperture (with respect to the Gaussian beam radius) inside the resonator. To a very good approximation, the beam propagation factor can be calculated by using (8.1.40). If b denotes the radius of the active medium with length l and the distances from the endfaces to the nearest mirror are l_1 and l_2 (Fig. 8.1.13), the beam propagation factor as a function of the resonator parameters reads:

$$M^{2} = \frac{\pi b^{2}}{\lambda L} \frac{|g_{1} + g_{2} - 2g_{1}g_{2}|}{\sqrt{g_{1}g_{2}(1 - g_{1}g_{2})}} \times \left[1 + \frac{\left[[g_{1} + g_{2} - 2g_{1}g_{2}][X - Lg_{2}(1 - g_{1})/(g_{1} + g_{2} - 2g_{1}g_{2})]\right]^{2}}{L^{2}(g_{1}g_{2}(1 - g_{1}g_{2}))}\right]^{-1}$$
(8.1.44)

with

$$X = l_1 + l$$
, if $L_{01} < l_1 + l/2$ (see Fig. 8.1.13 and (8.1.21)),
 $X = l_1$ otherwise,
 $L = l_1 + l_2 + l$.



Fig. 8.1.13. The size of the active medium determines the number of transverse modes. The ratio of the radius b to the largest Gaussian beam radius inside the medium w_{00} determines the beam propagation factor. L_{01} is the distance of the beam waist from mirror 1, given by (8.1.21).

Figure 8.1.14 presents the beam parameter product as a function of the g-parameters for stable resonators of equal length L = 1 m and different locations of the active medium with radius b = 5 mm. The best beam quality is attained near the stability limits at positive g-parameters since the Gaussian beam radii go to infinity in this area. If the active medium is placed in the middle of the resonator, the concentric resonator generally exhibits the worst beam quality. Note that the beam parameter product in multimode operation does not depend on the wavelength! This means that the size of the focal spot does not depend on the type of laser material used. A smaller wavelength will generate a smaller-radius Gaussian beam which in turn increases the number of transverse modes fitting into the medium. Therefore, the product $M^2\lambda$ is a constant of the resonator.



Fig. 8.1.14. Beam parameter products in the stability diagram for different positions of the active medium (b = 5 mm, l = 100 mm). The resonator length is always L = 1 m. For each graph the maximum beam parameter product is given [97Hod].

8.1.3.4 Resonance frequencies

The development of steady-state field distributions in an unconfined optical resonator requires that both the amplitude and the phase of the electric field E(x, y) are reproduced after each round trip. According to (8.1.8) this requirement is met if the eigenvalue γ is equal to 1.0, which is referred to as the *resonance condition*. Together with (8.1.10) and (8.1.12), the resonance condition yields the resonance frequencies ν of the eigenmodes [61Fox]:

circular symmetry:

$$\nu_{plq} = \frac{c_0}{2L} \left[q + \frac{2p + \ell + 1}{\pi} \, \operatorname{acos} \sqrt{g_1 g_2} \right] \,, \tag{8.1.45}$$

rectangular symmetry:

$$\nu_{mnq} = \frac{c_0}{2L} \left[q + \frac{m+n+1}{\pi} \, \cos \sqrt{g_1 g_2} \right] \tag{8.1.46}$$

with

 c_0 : speed of light in vacuum,

 $L = nL_0$: optical mirror spacing (L_0 : geometrical spacing, n: index of refraction inside the resonator).

The resonance frequencies thus depend on both the axial and the transverse mode order. In contrast to the plane-plane resonator FPI ($g_1 = g_2 = 1$) whose resonance frequencies are only determined by the axial mode index, each axial mode of stable resonators is subdivided into a sequence of frequencies corresponding to different transverse modes. This separation is controlled by the *g*-parameters of the resonator mirrors (Fig. 8.1.15). As the origin of the *g*-diagram is approached, the frequency gap between different transverse modes having the same axial mode order becomes wider. In the limit of the stable confocal resonator ($g_1 = g_2 = 0$), the frequency gap equals $c_0/(4L)$ which is half the axial mode distance.



Fig. 8.1.15. Resonance frequencies of the modes of stable resonators with rectangular symmetry for different g-parameter products. The indices mn of some transverse modes are shown for the axial mode of order q + 2.

8.1.4 Aperture-limited stable resonators

In addition to selecting the number of oscillating transverse modes, an aperture inside the resonator will also generate diffraction losses which are the result of a changed transverse mode structure. If the aperture radius is decreased so that it becomes smaller than the fundamental-mode beam radius, the angle of divergence will increase resulting in an increased power fraction hitting the aperture after each round trip. The intensity distribution at the mirrors is no longer Gaussian and the beam propagation inside the resonator does not follow the Gaussian beam propagation rules. The losses of the fundamental mode will decrease as the aperture radius is increased and the intensity profile will approach that of a Gaussian beam. However, decreasing the diffraction loss by opening the aperture will also enable higher-order modes to reach laser threshold.

8.1.4.1 Resonators with one aperture

If only one aperture is placed inside the resonator directly in front of mirror 1 (Fig. 8.1.16), the field distributions at mirror 1 are eigensolutions of the Kirchhoff integral equation (8.1.8) for the round trip [61Boy, 61Fox, 63Fox, 65Li, 65Heu, 65Sie, 66Kog1, 66Kog2]. In circular geometry with an aperture radius a, the integral equation for the fields of the transverse eigenmodes reads:

$$\gamma_{pl} E_{pl}(r_2, \Phi_2) = i \frac{\exp[-ikL]}{2L g_2 \lambda_0} \\ \times \int_{0}^{2\pi} \int_{0}^{a} E_{pl}(r_1, \phi_1) \cdot \exp\left[\frac{-i\pi}{2L g_j \lambda_0} \left(G(r_1^2 + r_2^2) - 2r_1 r_2 \cos(\phi_1 - \phi_2)\right)\right] r_1 dr_1 d\phi_1$$
(8.1.47)

with

 $G = 2g_1 g_2 - 1,$ $L = n L_0$: optical resonator length (L_0 : geometrical length), λ_0 : vacuum wavelength, $k = 2\pi/\lambda_0$: wave number.



Fig. 8.1.16. Stable resonator with aperture-limited mirror 1 (circular symmetry).

In contrast to the unconfined resonator, the loss factors $V = |\gamma_{pl}|$ are now lower than 1.0 since the eigenmodes exhibit diffraction losses at the aperture. The diffraction loss and the mode structure depend only on two parameters:

- 1. the equivalent G-parameter: $G = 2g_1g_2 1$,
- 2. the effective Fresnel number: $N_{\text{eff}} = a^2/(2Lg_2\lambda_0)$.

The effective Fresnel number is proportional to the area of the aperture. Its relationship to the Gaussian beam radius $w_{00}^{(1)}$ at mirror 1 is given by:



Fig. 8.1.17. Loss factor per round trip of different transverse modes in circular symmetry for G = 0.0 as a function of the effective Fresnel number $N_{\rm eff}$. The aperture radius is equal to the Gaussian beam radius for $N_{\rm eff} = 1/\pi \approx 0.318$.

Fig. 8.1.18. Calculated loss factors per round trip of the fundamental mode in stable resonators with a circular aperture of radius a as a function of the equivalent G-parameter $G = 2g_1g_2 - 1$. The aperture radius is shown normalized to the Gaussian beam radius w.

$$\left|\frac{a}{w_{00}^{(1)}}\right|^2 = \pi N_{\text{eff}} \sqrt{1 - G^2} \,. \tag{8.1.48}$$

Stable resonators in fundamental-mode operation with an aperture radius that is 1.2...1.5 times larger than the Gaussian beam radius exhibit an effective Fresnel number around 1.0. All resonators having the same absolute value of the equivalent g-parameter and the same absolute value of the effective Fresnel number exhibit the same eigenmodes and loss factors. These resonators are referred to as equivalent resonators. The integral equation (8.1.47) generally cannot be solved analytically. Figure 8.1.17 presents the numerically calculated loss factors of different low-order transverse modes as a function of the effective Fresnel number for resonators with G = 0.0. In most lasers only the loss of the fundamental mode is of practical interest. The loss of the fundamental mode determines both the laser threshold and the efficiency, whereas the losses of higher transverse modes affect only the output power. Figure 8.1.18 shows calculated loss factors per round trip in fundamental-mode operation as a function of the equivalent G-parameter, a comparison with experimental data is shown in Fig. 8.1.19.



Fig. 8.1.19. Measured loss factors per round trip as a function of the ratio of the aperture radius a to the Gaussian beam radius w for two different *G*-parameters. The solid line represents the numerically calculated loss factor [88Ozy].

In order to minimize the fundamental-mode loss the aperture radius has to be chosen as large as possible. However, to prevent the next transverse mode from oscillating it is necessary to choose an aperture radius of less than $1.2 \ldots 1.4$ times the Gaussian beam radius. The larger |G| and the higher the gain of the medium, the closer the aperture radius can approach the Gaussian beam radius. The loss per round trip of the fundamental mode can be approximated by using the empirical relation [87Sil, 87Gro]:

$$\Delta V = \exp\left[-\alpha N_{\text{eff}}^{\beta}\right] \,. \tag{8.1.49}$$

The parameters α and β for the TEM₀₀ mode read:

G	0.0	0.2	0.4	0.5	0.6	0.7	0.8	0.85	0.9	0.95	0.99
α	8.4	6.9	4.9	4.4	3.8	3.5	2.9	2.6	2.3	2.0	1.83
β	1.84	1.66	1.38	1.34	1.34	1.27	1.16	1.08	1.01	0.86	0.59

8.1.4.2 Resonators with two apertures

This case is much more complicated to deal with since the mode properties depend on three parameters rather than only on two parameters. In order to calculate the loss factor and the mode structure, the Kirchhoff integral has to be applied to both transits inside the resonator. If the aperture radii at mirror 1 and 2 are a_1 and a_2 , respectively (Fig. 8.1.20), the loss per round trip of each transverse mode is a function of the modified g-parameters g_1a_1/a_2 , g_2a_2/a_1 , and the Fresnel number N with

$$N = \frac{a_1 a_2}{\lambda L} \ . \tag{8.1.50}$$

It is, therefore, difficult to present a general overview of the mode properties of these resonators. Furthermore, the number of publications dealing with double-aperture resonators is very limited [65Li, 65Str, 83Pic, 83Rig]. The loss per round trip is given by $\Delta V_{\rm R} = 1 - |\gamma_{p\ell}|^2$. It is common to



Fig. 8.1.20. Stable resonator with two aperture-limited mirrors.

Fig. 8.1.21. Diffraction losses per transit of stable resonators in circular symmetry with both mirrors limited by apertures with radius $a_1 = a_2 = a$ as a function of the Fresnel number $a^2/(\lambda L)$. The losses for the fundamental mode and the next-order transverse mode are shown. The curve parameter is the *g*-parameter of mirror 2 (after [65Li]).

define the loss per transit via $\Delta V_{\rm T} = 1 - \sqrt{|\gamma_{p\ell}|^2}$, which represents the average loss per transit since the losses are generally different for the two directions. Diffraction losses per transit of stable resonators in circular symmetry with both mirrors limited by apertures with radius $a_1 = a_2 = a$ as a function of the Fresnel number $a^2/(\lambda L)$ are shown in Fig. 8.1.21.

8.1.5 Misalignment sensitivity

The misalignment sensitivity of a resonator is defined as the sensitivity with which the diffraction losses or the output power are changed due to mirror tilt [77Ber, 80Hau, 86DeS, 86Mag, 88Hau2, 92Hod2]. In this section we will only discuss the influence of the mirror misalignment on the resonator losses. The geometrical effect of a mirror tilt is shown in Fig. 8.1.22 for a resonator with one aperture-limited mirror. Rotation of mirror j by an angle α_j results in a rotation of the optical axis by an angle θ_j with the center of curvature of mirror i being the pivot point. As in the aligned resonator, the optical axis is defined by the line going through the centers of curvature of



Fig. 8.1.22. Geometry of mirror misalignment in an optical resonator. Rotation of mirror 2 by α_2 results in a shift of the optical axis by Δ_{12} at mirror 1 and by Δ_{22} at mirror 2. The pointing stability θ and the shifts can be calculated by using (8.1.51)–(8.1.53). These equations are also applicable to unstable resonators.

Fig. 8.1.23. Measured loss factors per round trip of a stable resonator as a function of the angle of misalignment of mirror 2. Mirror 1 is limited by an aperture with radius a (G = 0.34, L = 0.7 m, a = 0.55 mm, $\lambda_0 = 1.064$ µm).

the mirrors. The angle of rotation of the optical axis θ_j , also referred to as the pointing stability, and the shifts Δ_{jj} , Δ_{ij} of the intersecting points of the optical axis on mirror j and mirror i, respectively, read [80Hau]:

$$\theta_j = \alpha_j \, \frac{1 - g_i}{1 - g_1 g_2} \,, \tag{8.1.51}$$

$$\Delta_{jj} = \alpha_j \; \frac{Lg_i}{1 - g_1 g_2} \;, \tag{8.1.52}$$

$$\Delta_{ij} = \alpha_j \; \frac{L}{1 - g_1 g_2} \;, \tag{8.1.53}$$

where g_i is the g-parameter of mirror i, and i, j = 1, 2 with $i \neq j$. The transverse eigenmodes will keep oscillating parallel to the optical axis whether the resonator is aligned or not. Furthermore, as far as stable resonators are concerned, the mode structure also stays almost symmetric to the optical axis. Since the optical axis comes closer to one side of the aperture if one or both mirrors are tilted, additional diffraction losses are generated. The effective aperture radius is decreased by the shift of the optical axis. For all linear resonators, stable as well as unstable ones, the loss factor decreases parabolically with the angle of rotation for small mirror tilts:

$$V(\alpha) = V(0) \left[1 - 0.1 \left(\frac{\alpha}{\alpha_{10\%}} \right)^2 \right]$$
(8.1.54)

with V(0) being the loss factor for the aligned resonator (Fig. 8.1.22). Measured loss factors per round trip of a stable resonator as a function of the angle of misalignment of mirror 2 are shown in Fig. 8.1.23.

The angle $\alpha_{10\%}$ denotes the angle of misalignment at which the loss factor has decreased by 10% and consequently the losses have increased by 10%. This angle is used to define the misalignment sensitivity of optical resonators. A low misalignment sensitivity is equivalent to a small 10%-angle $\alpha_{10\%}$. Typical values of $\alpha_{10\%}$ for a stable resonator in fundamental-mode operation are on the order

of 50 μ rad. The exact value of the 10%-angle depends on the *g*-parameters of the mirrors, the Fresnel number, the resonator length, and the aperture radii. All resonators exhibit two 10%-angles, each corresponding to the tilt of one mirror. It is customary to define an average misalignment sensitivity by taking the geometrical mean value of the two 10%-angles:

$$\alpha_{10\%} = \frac{1}{2} \sqrt{\alpha_{10\%,1}^2 + \alpha_{10\%,2}^2} \tag{8.1.55}$$

with the additional indices indicating the corresponding mirror. This mean angle defines the average angle by which both mirrors can be rotated simultaneously before a 10% increase in diffraction losses is generated.

8.1.5.1 Fundamental-mode operation

Let both aperture radii a_1, a_2 be adapted to the Gaussian beam radii $w_{00}^{(1)}, w_{00}^{(2)}$ at the mirrors with $a_i = s w_{00}^{(i)}$ and s = 1.2...1.6. If mirror *i* is misaligned by an angle α_i , a first-order perturbation analysis of the diffraction integral yields for the loss factor per round trip [80Hau]:

$$V(\alpha_i) = V(0) \left[1 - \alpha_i^2 \frac{s^2}{\exp\left[2s^2\right] - 1} D_i^2 \right]$$
(8.1.56)

with

$$D_i^2 = \frac{\pi L^*}{\lambda_0} \sqrt{\left(\frac{g_j}{g_i}\right)} \frac{1 + g_1 g_2}{(1 - g_1 g_2)^{1.5}} , \qquad (8.1.57)$$

 $L^* = L_0 - (n-1)l/n$: effective resonator length, L_0 : geometrical resonator length,

L₀. geometrical resonator length,

 $n\colon \mathrm{index}$ of refraction of active medium,

 $l\colon \text{length}$ of active medium.

The misalignment of a resonator with both mirrors aperture-limited is shown in Fig. 8.1.24.



Fig. 8.1.24. Misalignment of a resonator with both mirrors aperture-limited [80Hau] (© OSA 1980).

If both mirrors are misaligned, the mean 10%-angle is given by

$$\alpha_{10\%} = \sqrt{0.1} \ \frac{\sqrt{\exp\left[2s^2\right] - 1}}{s\sqrt{D_1^2 + D_2^2}} = \sqrt{0.1} \ \frac{\sqrt{\exp\left[2s^2\right] - 1}}{sD}$$
(8.1.58)

with

$$D = \sqrt{\left(\frac{\pi L}{\lambda_0}\right) \frac{1 + g_1 g_2}{(1 - g_1 g_2)^{1.5}} \frac{|g_1 + g_2|}{\sqrt{g_1 g_2}}}.$$
(8.1.59)



Fig. 8.1.25. Curves of constant misalignment sensitivity D for resonators with both apertures adapted to the Gaussian beam radius. D_0 is the sensitivity of the confocal resonator [80Hau] (© OSA 1980).

We define D as the misalignment sensitivity of the resonator since a small value of D results in a large 10%-angle (Fig. 8.1.25). The misalignment sensitivity in Fig. 8.1.25 is normalized to the sensitivity D_0 of the symmetric confocal resonator with $g_1 = g_2 = 0$, given by

$$D_0 = \sqrt{\frac{2\pi L}{\lambda_0}} \ . \tag{8.1.60}$$

The misalignment sensitivity is only a function of the resonator length and the g-parameters. Lowest misalignment sensitivities again are attained for resonators near the axes of the stability diagram.

8.1.5.2 Multimode operation

With both mirrors aligned the modes fill the entire medium since it represents the only modeselecting aperture in the resonator. As the optical axis is rotated due to misalignment, the mode structure gets cut at one edge of the medium (lower left corner in Fig. 8.1.26) resulting in a decrease of the mode volume [86DeS, 92Hod2]. However, the optical axis remains the center of gravity of the mode structure which means that the beam gets clipped on both sides although only one side is affected by an edge. For a radius b of the active medium and small angles of misalignment, the cross-sectional area $A_b(\alpha)$ of the beam in the misaligned resonator reads



Fig. 8.1.26. A misaligned stable resonator in multimode operation. The hatched area indicates the mode volume [92Hod2] (© Taylor & Francis 1992).

$$A_b(\alpha) = \pi b^2 \left[1 - 0.1 \left(\frac{\alpha}{\alpha_{10\%}} \right) \right] , \qquad (8.1.61)$$

where $\alpha_{10\%,i}$ is the angle at which the mode cross section and, consequently, the output power have decreased by 10%, when mirror *i* is tilted (Fig. 8.1.26):

$$\alpha_{10\%,i} = \frac{0.025 \pi b}{L_{\text{eff}}} \frac{|1 - g_1 g_2|}{|1 - x_j / \rho_j|}, \quad i, j = 1, 2, \quad i \neq j$$
(8.1.62)

with

$$\begin{aligned} x_j &= \ell_j - \ell(n-1)/(2n) , \text{ if } \rho_j > \ell_j + \ell/(2n) , \\ x_j &= \ell_j + \ell(n+1)/(2n) \text{ else,} \\ L_{\text{eff}} &= \ell_1 + \ell_2 + \ell/n : \text{ effective resonator length.} \end{aligned}$$

The dependence of the measured 10%-angle $\alpha_{10\%,2}$ on the effective resonator length L_{eff} is shown in Fig. 8.1.27 for four stable resonators.



Fig. 8.1.27. Dependence of the measured 10%-angle $\alpha_{10\%,2}$ for four stable resonators on the effective resonator length L_{eff} (pulsed Nd:YAG rod laser in single-shot operation, b = 3.17 mm, l = 75 mm). The theoretical curves calculated with (8.1.62) are marked by solid lines. Resonator data are presented in each graph. In the unstable region (top right graph), (8.1.62) does not apply [92Hod2] (© Taylor & Francis 1992).

8.1.6 Unstable resonators

If the g-parameters of the resonator mirrors fulfill the relations $g_1g_2 < 0$ or $g_1g_2 > 1$, the radiation inside the resonator cannot be characterized by Gauss-Hermite or Gauss-Laguerre polynomials. These resonators are referred to as *unstable resonators* [65Sie, 72Ana], whereby the term "unstable" accounts for the fact that a Gaussian beam launched into the resonator will increase its beam radius after each round trip and, therefore, does not represent a "stable" eigensolution of the resonator (Fig. 8.1.28). By using the equivalent *G*-parameter $G = 2g_1g_2 - 1$ we can characterize the different resonator schemes (as already discussed in Sect. 8.1.2):

- stable resonators: 0 < |G| < 1,
- resonators on the stability limits: |G| = 1,
- unstable resonators: |G| > 1.

See Fig. 8.1.3 for the stability diagram of optical resonators.

Furthermore, one is able to distinguish *positive-branch* unstable resonators (G > 1) and *negative-branch* unstable resonators (G < -1). The steady-state field distributions on the mirrors of unstable resonators, similar to stable resonators, are solutions of the Kirchhoff integral equation (see (8.1.8)). However, in contrast to stable resonators, the beam propagation inside unstable resonators, to a good approximation, can be described by the propagation laws of geometrical optics.



Fig. 8.1.28. No Gaussian beam can reproduce itself in an unstable resonator. The beam radii at the mirrors increase with each round trip.

8.1.6.1 Beam propagation

8.1.6.1.1 Characterization of unstable resonators

Unstable resonators are characterized by the presence of spherical waves inside the resonator that reproduce themselves after each round trip [65Sie, 66Kah, 72Ana]. In every unstable resonator we can find two spherical waves whose radii of curvature at any plane inside the resonator are reproduced after a round trip. If R_+ and R_- denote the radii of curvature at mirror 1 (after being reflected off the mirror) of the two spherical waves (Fig. 8.1.29), the following relation holds:

$$R_{\pm} = \frac{\pm 2Lg_2}{|G| \pm \sqrt{G^2 - 1} - 2g_2 + 1} \,. \tag{8.1.63}$$



Fig. 8.1.29. The two self-reproducing spherical waves with radii of curvature R_+ and R_- at mirror 1 (after the reflection). After one round trip the beam diameter is changed by a factor $M_+(>1)$ and $|M_-|$ ($M_- < 1$), respectively.

Fig. 8.1.30. Beam propagation inside an unstable resonator (diverging wave).

If mirror 1 is limited by an aperture with radius a, which means that the beam starting at mirror 1 has a diameter of 2a, the radii of curvature R_+ and R_- reproduce themselves after the round trip, but the beam diameter is magnified by a factor M_+ and M_- , respectively, with

$$M_{\pm} = |G| \pm \sqrt{G^2 - 1}$$
 and $|M_{+}M_{-}| = 1$. (8.1.64)

The spherical wave with radius R_+ increases the beam radius after each round trip by the factor $|M_+|$, called the *magnification*. Since $|M_+| > 1$ holds, the corresponding spherical wave is referred to as the *diverging wave*. If the power starting at mirror 1 (inside the aperture) is given by P_0 , only the power

$$P_1 = \frac{1}{M_+^2} P_0 \tag{8.1.65}$$

hits the mirror inside the aperture after the round trip, provided that the intensity profile is homogeneous. The loss ΔV per round trip and the loss factor V per round trip are thus given by

$$\Delta V = 1 - \frac{1}{M_+^2}, \quad V = 1 - \Delta V = \frac{1}{M_+^2}.$$
(8.1.66)

The loss factor represents the power fraction that stays inside the resonator after each round trip.

In contrast to the diverging wave, the converging wave (R_{-}) decreases the beam diameter after each round trip by a factor $|M_{-}|$ with $|M_{-}| < 1$. The power P_0 starting at mirror 1 is conserved, but with every round trip the beam radius at mirror 1 is continuously decreased by $|M_{-}|$ so that no steady-state beam radius can be established on the mirror. After a few round trips in the resonator, the beam radius of the converging wave reaches its minimum value given by the diffraction limit and will then expand again. The converging wave transforms itself into a diverging wave due to diffraction. It is for this reason that the beam propagation in an unstable resonator is characterized by the divergent wave only. However, the convergent wave may have an influence on the mode properties of unstable resonators, if it is continuously excited by reflection off apertures or the endfaces of the active medium [72Ana, 90Hod2].

Figure 8.1.30 depicts the beam propagation in an unstable resonator. Instead of limiting mirror 1 by an aperture, the size of the highly reflecting area now defines the beam size on the mirror. The

(8.1.69)



Fig. 8.1.31. Technical realization of the outputcoupling mirror of unstable resonators.

laser beam is generated by output coupling around the reflective spot on the mirror. In circular symmetry, the near field exhibits the shape of an annulus with inner radius a and outer radius Ma. The size of mirror 2 is chosen such that no power is coupled out at this side of the resonator. The coatings on both mirrors are highly reflecting for the desired wavelength of laser emission. The following relations hold:

g-parameters:

$$g_i = 1 - \frac{L}{\rho_i}, \quad i = 1, 2,$$
 (8.1.67)

equivalent *G*-parameter:
$$G = 2g_1g_2 - 1, \qquad (8.1.68)$$

magnification (round trip):

magnification (transit):

$$M' = g_1 + \frac{\sqrt{G^2 - 1}}{2g_2} , \qquad (8.1.70)$$

 $M = |G| + \sqrt{G^2 - 1}$,

radius of curvature of the spherical wave at mirror 1:

- traveling towards the mirror:

$$R_{\rm t} = \frac{L}{M' + 1 - 2g_1} , \qquad (8.1.71)$$

- reflected off the mirror:

$$R_{\rm r} = \frac{2Lg_2}{M+1-2q_2}, \qquad (8.1.72)$$

loss factor per round trip:

$$M + 1 - 2g_2$$
, (0.1.72)
 $M = 1$ (0.1.72)

$$V = \frac{1}{M^2} . (8.1.73)$$

Three different ways to accomplish this special output-coupling scheme in unstable resonators exist (Fig. 8.1.31): The high-reflecting, confined mirror can be held in place by thin pins or mounted in the bore of a highly transmitting substrate (b). More common is the application of a high-reflecting coating on an AR-coated substrate as depicted in (a). For CO_2 lasers and other lasers emitting in the wavelength range on the order of 10 μ m, the output coupling by means of a scraper is a well established technique (c).

8.1.6.1.2 Resonator schemes

Unstable resonators can be subdivided into two different classes:

1. $g_1g_2 > 1$, positive branch

These resonators exhibit either zero or two focal points inside the resonator. The centers of curvatures of the spherical waves at the mirrors are located either outside $(g_1 > 0, g_2 > 0)$ or inside $(g_1 < 0, g_2 < 0)$ the resonator.

2. $g_1g_2 < 0$, negative branch

One center of curvature of the spherical wave is located inside the resonator. The *negative-branch unstable resonators* therefore exhibit a focal spot in the resonator. Due to possible damage of the active medium by the high intracavity intensities, these resonators are only used in high-gain gas lasers. Compared to positive-branch resonators they exhibit a much lower misalignment sensitivity.

Resonators for which the relation

$$g_1 + g_2 = 2g_1g_2 \tag{8.1.74}$$

holds, are referred to as *confocal resonators* (Fig. 8.1.32). For confocal resonators, the focal points of the two resonator mirrors are on top of each other, which means that the unstable resonator acts like a telescope. This special mirror arrangement provides a collimated output beam. The mode volume can thus be easily adapted to the volume active medium yielding optimum fill factors.

Figure 8.1.33 shows curves of constant magnification in the stability diagram. The broken lines indicate the confocal resonators. The curves with magnification |M| = 1 are the stability limits.



Fig. 8.1.32. Beam propagation in a positivebranch confocal unstable resonator with magnification |M| = 2.

Fig. 8.1.33. Curves of constant magnification in the stability diagram. The broken lines indicate the confocal resonators. The curves with magnification |M| = 1 are the stability limits.

8.1.6.2 Mode structures and losses

The steady-state field distributions on the mirrors of unstable resonators can be calculated similarly to those of stable resonators by applying the integral equation (8.1.47) (Fig. 8.1.34) [65Sie, 67Sie, 74Sie, 75Szi, 87Oug, 88Hau1]. For both stable and unstable resonators the same integral equations can be used, the only difference is that the absolute value of the *G*-parameter is now greater than 1.0. Similar to the treatment of stable resonator modes, the integral equations can be simplified by separating the coordinates. The intensity distributions of the eigenmodes and the loss factor $V = |\gamma|^2$ again depend on the absolute value of the equivalent *G*-parameter *G* and the absolute value of the effective Fresnel number $N_{\text{eff}} = a^2/(2Lg_2\lambda)$. Unstable resonators exhibit an infinite set of TEM modes with mode indices *pl* and *mn*, but the mode properties are completely different as compared to the modes of stable resonators (Fig. 8.1.35). In contrast to stable-resonator modes, the loss factor does not always increase as the aperture radius is increased, leading to a periodic sequence of loss factor maxima.

Since different transverse modes exhibit their maxima at different Fresnel numbers, the loss factor curves cross, which means that at this point two transverse modes have the same diffraction loss. For transverse modes without azimuthal structure (l = 0) the loss factor is higher as compared to other modes and also higher than the geometrical loss factor of $1/M^2$. The reason for the lower loss becomes apparent if we investigate the intensity profiles of the modes at the loss-factor maxima. Since the profiles are more centered as compared to the homogeneous profile assumed in the preceding section, a higher power fraction hits the output coupler again after the round trip resulting in a higher loss factor.

It is important to note that the beam radii of different transverse modes are the same. This property, together with the difference in diffraction loss between transverse modes, determines



Fig. 8.1.34. Calculation of the field distribution at the plane of the output coupler for unstable resonators in circular symmetry. After the round trip the shape of the field at mirror 1 is still the same.

Fig. 8.1.35. Calculated loss factor per round trip as a function of the effective Fresnel number N_{eff} for unstable resonators in circular symmetry with magnification M = 2. The radial intensity distributions at the plane of the output coupler are shown for the first three loss-factor maxima.

177

the special oscillation behavior of unstable resonators. The mode with the lowest loss will start oscillating first and deplete the gain in the same area of the active medium that might be used by other transverse modes. These modes, however, exhibit too high of a loss for the leftover gain to reach the laser threshold. Except for the operation at mode-crossing points, unstable resonators thus oscillate in a single transverse mode with no azimuthal structure (l = 0).

For practical applications, therefore, it is sufficient to determine the loss and the mode structure only for the lowest-loss modes. The calculated loss factors per round trip of the lowest-loss mode for unstable resonators in circular and rectangular symmetry are presented in Fig. 8.1.36. In Fig. 8.1.37 measured and calculated loss factors per round trip of unstable resonators in circular symmetry with magnification M = 4 are shown. In these graphs, the loss factor is plotted versus the equivalent Fresnel number $N_{\rm eq}$, which is related to the effective Fresnel number via

$$N_{\rm eq} = N_{\rm eff} \sqrt{G^2 - 1} \,. \tag{8.1.75}$$

By using the equivalent Fresnel number as a mode parameter, the mode-crossing points are characterized by integral values of N_{eq} . In order to attain single-transverse-mode operation in an unstable resonator, it, therefore, is recommended to choose half-integral values of the equivalent Fresnel number because the loss difference between modes is a maximum in these areas (Fig. 8.1.38).



Fig. 8.1.36. Calculated loss factor per round trip of the lowest-loss modes as a function of the equivalent Fresnel number N_{eq} . (a) circular symmetry with mirror radius a, (b) rectangular symmetry with mirror width 2a, one-dimensional loss factor V_x , the total loss factor is $V = V_x V_y$ [92Hod1] (Chapman & Hall 1992).



Fig. 8.1.37. Measured and calculated loss factors per round trip of unstable resonators in circular symmetry with magnification M = 4 as a function of the equivalent Fresnel number (pulsed Nd:YAG laser) [88Hau1] (© AIP 1988).

Landolt-Börnstein New Series VIII/1A2


Fig. 8.1.38. Measured radial intensity distributions at the plane of the output coupler of unstable resonators in circular symmetry with G = 1.76 (M = 3.21) and different equivalent Fresnel numbers. The profile within the radius *a* is reflected by the mirror, the outer portion of the mode is coupled out. The dotted lines are the numerically calculated distributions (pulsed Nd:YAG laser) [88Hau1] (\bigcirc AIP 1988).

8.1.6.3 Beam quality

8.1.6.3.1 Circular symmetry

The beam quality of unstable resonators is determined by the diffraction at the confined outputcoupling mirror. If we assume that the outcoupled beam exhibits a plane phase front, the intensity distribution in the focal spot is given by the intensity distribution in the far field of a homogeneously illuminated annular ring with inner radius a and outer radius Ma. Application of the diffraction integral in the Fraunhofer approximation yields the far-field intensity distribution (Fig. 8.1.39):

$$I(\theta) = I(0) \frac{M^2}{M^2 - 1} \left[\frac{J_1(2\pi M a\theta/\lambda)}{\pi M a\theta/\lambda} - \frac{1}{M^2} \frac{J_1(2\pi a\theta/\lambda)}{\pi a\theta/\lambda} \right] , \qquad (8.1.76)$$

where J_1 is the Bessel function of order 1. For high magnifications $(M \gg 1)$, the far-field intensity distribution of a round aperture is obtained with a Full-Width-Half-Maximum (FWHM) diameter of the central peak of

$$\Delta \theta = 0.51 \frac{\lambda}{Ma} \,. \tag{8.1.77}$$

The power fraction in the side lobe and the spot radius r decrease as the magnification of the resonator is increased. It is for this reason that the magnification of unstable resonators is generally



Fig. 8.1.39. Far-field intensity distributions of a homogeneously illuminated ring with inner radius a and outer radius Ma in the limits of low and high magnification M (calculated with (8.1.76)).



Fig. 8.1.40. Calculated beam parameter products $d_0 \Phi/4\lambda$ (d: beam width, Φ : full angle of divergence, both defined via 86.5% power content) and power fractions in the center peak and in the side lobe as a function of the magnification (using the far-field distribution of a homogeneously illuminated ring with inner radius a and outer radius Ma).

chosen as high as possible with an upper limit given by the gain of the medium. Unstable resonators are mostly applied to active media with sufficiently high small-signal gain $(g_0 l > 1.5)$. For low-gain media the optimum magnification would be too low to have the major power fraction going into the central peak.

In the limit of high magnifications, the beam parameter product approaches the value for a round beam with a homogeneous intensity distribution (Fig. 8.1.40). Compared to a Gaussian beam, the beam parameter product of unstable resonators is at least 2.5 times higher. For typical magnifications between M = 2 and M = 4, the beam quality of unstable resonators is about three times worse as compared to a Gaussian beam. The model used to calculate the beam quality assumes that the near field is a perfect annular ring. In reality the mode structure in the near field exhibits variations in both the amplitude and the phase. This diffraction effect leads to a lower power fraction in the side lobe and to slightly lower beam parameter products. This is mainly due to a less steep slope in the outer area of the beam.

8.1.6.3.2 Rectangular symmetry

A similar geometrical treatment can be performed for unstable resonators in rectangular symmetry. The one-dimensional far field is given by the Fourier transform of a double slit with a slit width of (M-1)a and a slit separation (center-to-center) of (M+1)a. The angular intensity distribution in one dimension reads

$$I(\theta) = I(0) \left[\frac{\sin[\pi \,\theta(M-1)a/\lambda] \, \cos[\pi \,\theta(M+1)a/\lambda]}{\pi \,\theta(M-1)a/\lambda} \right]^2 \,. \tag{8.1.78}$$

For the same magnification, the intensity distribution at the focus of unstable resonators in rectangular symmetry exhibits much higher side lobes as compared to circular-symmetric unstable resonators (Fig. 8.1.41).



Fig. 8.1.41. Measured (solid lines) and calculated (broken lines) intensity distributions at the focus of unstable resonators for Nd:YAG lasers ($\lambda = 1.064 \ \mu m$). (a) circular symmetry, $a = 1.5 \ mm$, (b) rectangular symmetry, $a = 2.0 \ mm$ [92Hod1].

8.1.6.4 Unstable resonators with variable-reflectivity mirrors

The power content in the side lobes of the far field of unstable resonators can be decreased by using output-coupling mirrors with a variable reflectivity profile [65Vak, 70Zuc, 84McC, 88Sne, 88DeS2, 88DeS1, 92Mag, 93Hod1, 93Hod2]. The reflectivity profile which generally exhibits a maximum at the center improves the beam quality by two effects: The diffraction of the power into the side lobes is considerably decreased by replacing the hard edge by a continuous transition from high to low reflectance (Fig. 8.1.42). Furthermore, the center reflectivity and the shape of the profile can be varied to generate a specified intensity profile in the near field (e.g. flat-top beams). The specific shape of the reflectivity profile is only of secondary importance, especially if the center



Fig. 8.1.42. Calculated intensity distributions in the near field and the far field of a confocal unstable resonator with M = 2 and $N_{eq} = 0.5$ for different reflectivity profiles of the output coupler.



Fig. 8.1.43. Super-Gaussian reflectivity profiles for different indices n as a function of the radius.

reflectivity is chosen lower than 100%. However, in order to maximize the fill factor, steeper profiles are preferred since the mode with its correspondingly steeper slopes can then be better adapted to the active medium. From 1985 to 1994 much effort was spent to develop coating techniques for generating Variable-Reflectivity Mirrors (VRM) exhibiting Gaussian, super-Gaussian (Fig. 8.1.43), and parabolic reflectivity profiles [85Lav, 88DeS1, 89Emi, 90Pie, 93Dup, 93Pie, 94Bos3, 94Bos2, 94Bos1].

The VRMs most commonly used exhibit a *super-Gaussian reflectivity profile* which in circular symmetry reads:

$$R(r) = R_0 \exp\left[-2\left(\frac{r}{w}\right)^n\right]$$
(8.1.79)

with

 R_0 : center reflectivity, w: profile radius, n: super-Gaussian index, r: radial coordinate.

The super-Gaussian index n determines the shape of the reflectivity profile. For n = 2 the profile is Gaussian and with increasing n the slopes become steeper until the hard-edge profile is approached in the limit $n \to \infty$ (Fig. 8.1.43). The mode properties can be evaluated by using the stationary condition for the electric field. The corresponding intensity profile is given by

$$I(r) = I_0 \exp\left[-2\left(\frac{r}{w_M}\right)^n\right]$$
(8.1.80)

with

$$w_M = w (M^n - 1)^{1/n}, \quad M > 1.$$
 (8.1.81)

The total loss factor per round trip reads:

$$V = \gamma \gamma^* = \frac{R_0}{M^2} \,. \tag{8.1.82}$$

Thus, the intensity distribution at the output coupler inside the resonator is also super-Gaussian with a beam radius w_M . The intensity profile of the near field can be obtained by multiplying



Fig. 8.1.44. Calculated loss factor per round trip of the lowest-loss mode (l = 0) as a function of the equivalent Fresnel number for circularly symmetric unstable resonators with magnification M = 2 using different super-Gaussian indices n [92Bow].

Fig. 8.1.45. Measured beam parameter products $d \Phi/4$ (d: beam width, Φ : full angle of divergence, both defined via 86.5% power content) in the x-(width) and the y-direction (height) for a 2.3 kW Nd:YAG slab laser system with one oscillator slab and two amplifier slabs of dimensions $7 \times 26 \times 179 \text{ mm}^3$. The unstable resonator comprises a VRM with profile radii of 2 mm and 7 mm in the x- and y-direction, respectively. Total pump energy: 3,400 J; pulse length: 4 ms. Data are shown for one (osc.), two (osc./amp.) and three slabs (osc./amp./amp.) [93Hod3] (© OSA 1993).

(8.1.80) with the intensity transmission 1 - R(r) of the VRM. In this approach the loss factor per round trip is the same for all super-Gaussian indices and equal to the geometrical loss factor! This result can be verified if diffraction integrals are used to propagate the field inside the resonator (Fig. 8.1.44). Only for low equivalent Fresnel numbers can the beam propagation not be described by geometrical optics.

Owing to their excellent beam qualities, unstable resonators with variable-reflectivity mirrors have found application in both gas and solid-state lasers. VRMs are commercially available for a variety of laser wavelengths ranging from the visible to the infrared at 10.6 μ m. In solid-state lasers, these resonators are mostly applied to low to medium power *Q*-switch systems [88Sne, 88Lav], but successful operation in free-running lasers with output powers in the kW range has also been reported [92Mag, 93Hod1, 93Hod2, 93Hod3, 94Bos2] (Fig. 8.1.45). If a flat-topped intensity distribution in the near field is required (e.g. for efficient filling of amplifier stages), the center reflectivity and the reflectivity profile can be used as design parameters. In order to attain optimum filling of the active medium by the laser mode an optimization of the reflectivity profile is required. Since the adaptation of the beam width to the width of the active medium is easier to accomplish if the mode exhibits steep slopes, super-Gaussian profiles with high index n provide better means to maximize the output power. For an unstable resonator with a Gaussian VRM, a complete filling of the active medium will generate side lobes in the far field. However, for super-Gaussian profiles with n > 2, the super-Gaussian field distributions at the output coupler are not eigensolutions of the paraxial wave-equation. The intensity distributions inside the resonator, therefore, will change when the field propagates inside the resonator. This may result in ripples and hot spots. The same statement holds for the field transmitted through the VRM.

8.1.6.5 Applications of unstable resonators

The main advantage of unstable resonators is the adaptation of the beam radius inside the resonator by changing the mirror radius *a*. Since single-transverse-mode operation is not linked to the size of the mirror, as is the case for stable resonators, beam qualities close to the diffraction limit can be attained even for gain media with very large cross sections. In contrast to stable resonators, a high mode volume and a high beam quality can be realized simultaneously. Unfortunately, there are several drawbacks which limit the application of unstable resonators.

First of all, the power content in the side lobes of the far field might cause problems in specific applications. In order to minimize the power content in the side lobes the magnification has to be chosen as high as possible. This, however, requires an active medium with a high gain. It therefore makes no sense to use an unstable resonator in a HeNe laser, since the gain is too low to obtain a good laser efficiency. Secondly, the output power of unstable resonators generally is $20\% \dots 30\%$ lower as compared to a stable resonator in multimode operation (Fig. 8.1.46). This is due to a lower fill factor in combination with additional diffraction losses induced by the rim of the active medium. If one is more interested in the output power than in attaining an excellent beam quality, it is therefore more advantageous to use a stable resonator in multimode operation. This preference of multimode stable resonators is also due to their easier alignment and their lower sensitivity to mirror misalignment.

Unstable resonators have found application in high-power CO₂ lasers [69Kru, 71Loc, 72Fre, 73Wis, 77Isa, 88Coo] and excimer lasers [79Rea, 85Iza, 88Lit], both of which generally exhibit a high gain and a large cross section of the active medium. Unstable resonators are also used in pulsed solid-state rod lasers [76Ewa, 77Her, 80Giu, 86Har, 86Upp1, 87Apa, 88Sne, 89Par, 90Hod1, 90Hod2, 90Hod3, 90Yas, 92Hod1, 91Mag, 92Mag, 93Hod1, 93Hod2, 94Bos2, 94Bur]. Unfortunately, the thermal lensing of the solid-state laser materials deteriorates the beam quality and makes beam handling quite difficult. In addition, unstable resonators are sensitive to aberrations of the thermal



Fig. 8.1.46. Maximum extraction efficiency as a function of the small-signal gain for different resonators. Calculation based on diffraction integrals in combination with rate equations [97Hod].

lens resulting in a deterioration of the beam quality in high-power solid-state lasers. It is for these reasons that most applications of unstable resonators for solid-state rod lasers are limited to low-power Q-switch systems and free-running systems with less than 100 W average power. Higher power and good beam quality can be realized with zig-zag lasers due to the first-order compensation of the thermal lens. Unfortunately, the zig-zag slab geometry is too complicated to be considered a reasonable choice for a commercial product. The utilization of unstable resonators in diode lasers [89Ruf], dye lasers [95Cha], and free-electron lasers [84Moo, 86Skl, 87Skl] has also been reported.

8.1.7 Output power of stable resonators

8.1.7.1 Calculation of the output power of stable resonators

In the resonator model used (Fig. 8.1.47) it is assumed that both the forward-traveling beam with intensity $I^+(z)$ and the backward-traveling beam with intensity $I^-(z)$ cover the same area of the active medium. The complete overlap of the two counterpropagating beams is characteristic for stable resonators. During a round trip the intensity is decreased due to diffraction losses (loss factors $V_1 \ldots V_4$), scattering, and absorption inside the medium (loss factor V_S), and by output coupling. In steady-state operation, these losses are compensated by the amplification process characterized by the small-signal gain coefficient g_0 .



Fig. 8.1.47. Resonator model used for the calculation of the output power.

The next assumption we make is that no spatial hole burning is present meaning that at any plane inside the medium the intensity I(z) is given by the sum of the two intensities $I^+(z)$ and $I^-(z)$. This is a reasonable approach for most lasers since the effect of spatial hole burning on the output power is smoothed out by atomic motion (gas lasers), energy migration, or axial multimode operation (solid-state lasers). However, in single-axial-mode solid-state or liquid dye lasers, spatial hole burning has an impact on the output power (it is lower by up to 30%). In these cases a realistic model for the output power requires the incorporation of the interference between the two counterpropagating fields resulting in the intensity profile

$$I(z) = I^{+}(z) + I^{-}(z) - 2\sqrt{I^{+}(z)I^{-}(z)} \cos[2kz]$$
(8.1.83)

with k being the wave number inside the medium. Without spatial hole burning, the differential equations for the intensities read [86Sie, 88Mil, 78Rig]:

$$\frac{\mathrm{d}I^{\pm}(z)}{\mathrm{d}z} = \pm \left[\frac{g_0}{\left(1 + \frac{I^+(z) + I^-(z)}{I_{\mathrm{S}}}\right)^X} - \alpha_0 \right] I^{\pm}(z) , \qquad (8.1.84)$$

where g_0 is the small-signal gain coefficient, I_S is the saturation intensity, α_0 is the loss coefficient, and X is equal to 1.0 (0.5) for homogeneous (inhomogeneous) broadening.

Equation (8.1.84) can be solved analytically if we assume that the intensity sum is constant with $I^+(z) + I^-(z) = 2I$. The mean intensity I now represents the intensity of the beam traveling towards the output coupling mirror. Equation (8.1.84) then yields for the factor GV_S by which the intensity is amplified in a transit through the medium

$$GV_{\rm S} = \exp\left[\frac{g_0\ell}{(1+\frac{2I}{I_{\rm S}})^X}\right] \exp\left[-\alpha_0\ell\right] \,. \tag{8.1.85}$$

By using the steady-state condition

$$G^2 R V_S^2 V_1 V_2 V_3 V_4 = 1 \tag{8.1.86}$$

the intensity I is found to be

$$I = \frac{I_{\rm S}}{2} \left[\left(\frac{g_0 \ell}{\left| \ln \sqrt{RV_{\rm S}^2 V_1 V_2 V_3 V_4} \right|} \right)^{1/X} - 1 \right] \,. \tag{8.1.87}$$

If the cross-sectional area of the beam is $A_{\rm b}$ (= πa_1^2 in Fig. 8.1.47), the output power is given by

$$P_{\rm out} = A_{\rm b} I (1-R) V_2 = A_{\rm b} I_{\rm S} \frac{(1-R) V_2}{2 \left(\left| \ln \sqrt{R V_{\rm S}^2 V} \right| \right)^{1/X}} \left[(g_0 \ell)^{1/X} - \left(\left| \ln \sqrt{R V_{\rm S}^2 V} \right| \right)^{1/X} \right] \quad (8.1.88)$$

with the round trip diffraction loss factor $V = V_1 V_2 V_3 V_4$.

This expression for the output power of stable resonators can be used to a very good approximation, if the output coupling is low (high reflectance R). In this case the sum of the two intensities inside the medium is almost constant as a numerically calculated example in Fig. 8.1.48 indicates. For lower reflectances, the z-dependence of the intensities has to be taken into account by solving the differential equation (8.1.84) using the boundary conditions at the mirrors. Unfortunately, the solution can only be found numerically due to the homogeneously distributed loss (loss coefficient α_0) inside the medium [78Rig, 80Eim, 80Sch]. However, for homogeneously broadened lasers (X = 1) it is possible to derive an analytical solution by setting α_0 equal to zero in (8.1.84) and taking into account the loss by multiplying the intensity at the high-reflecting mirror with the loss factor $V_{\rm S}^2 = \exp[-2\alpha_0 l]$. This method provides an analytical expression for the output power that is very close to the numerical solution (the difference is less than 0.5%). After a lengthy calculation the final expression reads [88Hod]:

$$P_{\rm out} = A_{\rm b} I_{\rm S} \left. \frac{(1-R)V_2}{1-RV_2 V_3 + \sqrt{RV} \left(1/(V_1 V_2 V_{\rm S}) - V_{\rm S} \right)} \left[g_0 \ell - \left| \ln \sqrt{RV_{\rm S}^2 V} \right| \right]$$
(8.1.89)

with

$$V = V_1 V_2 V_3 V_4$$

A comparison with (8.1.88) indicates that the z-dependence of the gain leads to a change in the fraction. For high reflectances R both equations provide similar output powers. Without lack of generality, we combine all losses into one and consider a resonator that exhibits a loss factor per transit $V_{\rm S}$. Equation (8.1.89) then reads:

$$P_{\rm out} = A_{\rm b} I_{\rm S} \, \frac{1 - R}{1 - R + \sqrt{R} \, (1/V_{\rm S} - V_{\rm S})} \left[g_0 \ell - \left| \ln \sqrt{R V_{\rm S}^2} \right| \right] \,. \tag{8.1.90}$$



Fig. 8.1.48. Gain coefficient g_0 and normalized intensity $I/I_{\rm S}$ as a function of the axial coordinate z for different reflectances R of the output coupler, calculated with (8.1.84) ($g_0l = 1.5$, $\alpha_0l = 0.053$, no diffraction losses).

Fig. 8.1.49. Normalized laser power $P_{\rm out}/(A_{\rm b}I_{\rm S})$ as a function of the small-signal gain g_0l for different reflectances of the output coupler according to (8.1.90) assuming a loss per transit of 5% ($V_{\rm S} = 0.95$).

Figure 8.1.49 presents the normalized output power $P_{\text{out}}/(A_{\text{b}}I_{\text{S}})$ as a function of the small-signal gain g_0l which for most laser materials is proportional to the pump power. Starting at the threshold small-signal gain $(g_0\ell)_{\text{th}} = |\ln(\sqrt{R}V_{\text{S}})|$ the power increases linearly with the small-signal gain and the slope of the curve becomes steeper as the reflectance R of the output coupler is decreased. The output power cannot exceed the power P_{OL} that is available in the medium in the form of inversion, with

$$P_{\rm OL} = AI_{\rm S}g_0\ell = \eta_{\rm excit}P_{\rm electr} \tag{8.1.91}$$

A: cross-sectional area of the medium,

 η_{excit} : excitation efficiency,

 P_{electr} : electrical pump power.

Only if the laser resonator exhibits no losses ($V_{\rm S} = 1.0$ and $R \to 1.0$ in (8.1.90)) and the laser beam fills the whole medium ($A_{\rm b} = A$) can all the available power $P_{\rm OL}$ be extracted from the active medium (broken line in Fig. 8.1.49). Note that (8.1.91) holds only for homogeneously broadened lasers. For inhomogeneous line broadening, the available power is a function of the laser intensity due to the broadening of the homogeneous line width.

The slope of the output power with respect to the electrical pump power is referred to as the slope efficiency η_{slope} with

$$\eta_{\text{slope}} = \frac{\mathrm{d}P_{\text{out}}}{\mathrm{d}P_{\text{electr}}} = \frac{\eta_{\text{excit}}}{AI_{\text{S}}} \frac{\mathrm{d}P_{\text{out}}}{\mathrm{d}(g_{0}\ell)} = \eta_{\text{excit}} \gamma \frac{1-R}{1-R+\sqrt{R}\left(1/V_{\text{S}}-V_{\text{S}}\right)},$$
(8.1.92)

where $\gamma = A_{\rm b}/A$ is the fill factor.

The slope efficiency is a function of the loss, the output coupling, and the ratio of the mode volume to the volume of the active medium. This ratio, called the fill factor, can be calculated in a first approach by using the beam diameter of the highest-order mode inside the active medium. If the active medium represents the only aperture inside the stable resonator and a high enough number of transverse modes are able to oscillate (let us say more than 10), the maximum fill factor of 1.0 is obtained. In fundamental-mode operation with the Gaussian beam radius w being adapted to the radius b of the active medium ($b \approx 1.3 w$), the fill factor typically is around 0.9.

Unless the resonator exhibits no losses, only a fraction of the maximally available power P_{OL} can be extracted in the form of a laser beam. This fraction is referred to as the extraction efficiency η_{extr} of the resonator:

$$\eta_{\text{extr}} = \frac{P_{\text{out}}}{P_{\text{OL}}} = \frac{1}{\eta_{\text{excit}}} \frac{P_{\text{out}}}{P_{\text{electr}}} .$$
(8.1.93)

The total efficiency of the laser resonator is given by:

$$\eta_{\rm tot} = \frac{P_{\rm out}}{P_{\rm electr}} = \eta_{\rm excit} \eta_{\rm extr} .$$
(8.1.94)

8.1.7.2 Optimum output coupling and maximum output power

If the reflectance of the output coupler is varied at a given small-signal gain or pump power, the output power exhibits a maximum at the optimum output coupling. This behavior is easy to understand considering the fact that the output power is zero at low reflectances (laser threshold is not reached) and at a reflectance of 100% (no power is coupled out of the resonator). Therefore, a maximum of the output power and the extraction efficiency must exist for a certain value of the output coupling. Going back to the approximate power formula (8.1.88) we can find the optimum reflectance R_{opt} and the maximum output power $P_{out,max}$ by setting the derivative $\delta P_{out}/\delta(\ln R)$ equal to zero. Again, we assume that the losses per transit are represented by the loss factor $V_{\rm S} = \exp[-\alpha_0 l]$, and we consequently set the diffraction loss factor V equal to 1.0. By using the approximation $1 - R \approx |\ln R|$, the following expressions are obtained [82Sve]:

8.1.7.2.1 Homogeneous line broadening

Maximum output power:

$$P_{\rm out,max} = A_{\rm b} I_{\rm S} \alpha_0 \ell \left[\sqrt{\frac{g_0 \ell}{\alpha_0 \ell}} - 1 \right]^2 = \left[\sqrt{\eta_{\rm excit} P_{\rm electr} A_{\rm b} / A} - \sqrt{A_{\rm b} I_{\rm S} \alpha_0 \ell} \right]^2.$$
(8.1.95)

Maximum extraction efficiency:

$$\eta_{\text{extr,max}} = \frac{\alpha_0 \ell}{g_0 \ell} \left[\sqrt{\frac{g_0 \ell}{\alpha_0 \ell}} - 1 \right]^2.$$
(8.1.96)

Optimum output coupling:

$$\ln R_{\rm opt} = -2\alpha_0 \ell \left[\sqrt{\frac{g_0 \ell}{\alpha_0 \ell}} - 1 \right] \,. \tag{8.1.97}$$

Although (8.1.88) is only an approximate expression for the output power, the extreme values given by (8.1.95)–(8.1.97) can, to a very good approximation, be used to optimize the power performance of a laser system. This can be easily verified in Fig. 8.1.50 in which the correct extreme values (found by solving numerically the differential equation (8.1.84)) are shown [80Sch]. Even for lasers with high small-signal gain ($g_0l > 3$) and high loss ($\alpha_0 l > 0.1$), which require a low reflectance for optimum performance, the difference between the exact extreme values and the ones given by (8.1.96) and (8.1.97) is negligible. Note that the graph presented in Fig. 8.1.50 assumes a fill factor of $\gamma = A_b/A = 1.0$. The optimum parameters for a given laser are represented by the intersecting point of the two curves with constant small-signal gain and constant loss. Again, we find that an extraction efficiency of 100% is only attainable if the laser exhibits no loss ($\alpha_0 l = 0$) and the output coupling is close to zero.

The losses have a less dramatic impact on the output power if the small-signal gain of the laser is high. This is why at the same average pump power, lasers provide a higher average output power in pulsed operation than in cw operation. For a repetition rate f and a pulse width Δt , the small-signal gain is higher by the factor $1/(f\Delta t)$ as compared to cw operation at the same average pump power. According to Fig. 8.1.50, the extraction efficiency is increased leading to a higher average output power. Furthermore, the extraction efficiency becomes less sensitive to changes in output coupling as Fig. 8.1.51 indicates. The realization of the optimum output coupling is thus much less critical than in a low-gain laser. Figure 8.1.52 presents experimental examples for output power and output energies as a function of the output coupling.



Fig. 8.1.50. Diagram for the determination of the optimum reflectance $R_{\rm opt}$ and the maximum extraction efficiency $\eta_{\rm extr,max}$ of homogeneously broadened lasers using the small-signal gain $g_0 l$ and the loss per transit $\alpha_0 l = -\ln V_{\rm S}$ (fill factor $\gamma = 1.0$) (after [80Sch]).



Fig. 8.1.52. Measured output energy per pulse of Nd:YAG lasers as a function of the pump energy $E_{\rm p}$ and the output-coupling reflectance R [97Hod]. The lines represent interpolations.

Expression (8.1.89) can be used to calculate the output power of a stable resonator if the small-signal gain, the losses, and the fill factor of the resonator are known. The experimental determination of the gain and the loss is easy to accomplish (e.g. Findlay-Clay analysis). The determination of the fill factor, however, is more difficult if only a small number of transverse modes is considered. In multimode operation the fill factor can, to a good approximation, be calculated by using the cross-sectional area of the aperture. If the aperture with radius a is located at the active medium with radius b, the fill factor is given by

$$\gamma = \frac{\pi a^2}{\pi b^2} \,. \tag{8.1.98}$$

If the aperture is not close to the medium, the beam propagation inside the resonator can be used to determine the equivalent aperture radius inside the medium. In order to obtain the correct output power with (8.1.89), the diffraction loss per transit in multimode operation has to be incorporated into the loss factors $V_1 \ldots V_4$.

8.1.7.2.2 Inhomogeneous line broadening

Maximum output power:

$$P_{\rm out,max} = A_{\rm b} I_{\rm S} \; \frac{|\ln R|^2}{\alpha_0 \ell + \ln \sqrt{R}} \;. \tag{8.1.99}$$

Optimum output coupling:

$$\left[\alpha_0 \ell - \ln \sqrt{R_{\text{opt}}}\right]^3 = (g_0 \ell)^2 \left[\alpha_0 \ell + \ln \sqrt{R_{\text{opt}}}\right] \,. \tag{8.1.100}$$

Note that the output power (8.1.99) is always positive since (8.1.100) implies that $\alpha_0 \ell + \ln \sqrt{R}$ is greater than zero. The output power refers to the power of one single axial mode, and the small-signal gain is generated only by those inverted atoms whose resonance frequencies lie within the homogeneous linewidth around the axial-mode frequency. In order to obtain the total output power the contributions of the other axial modes have to be added. Again, these expressions are only valid for high mirror reflectances.

8.1.8 Thermal lensing in solid-state lasers

In solid-state lasers the active medium exhibits the properties of a thermally induced lens when the material is pumped. The refractive power is brought about by a combination of heat generation due to absorption of pump and laser radiation and the flow of heat to the outer periphery due to cooling [70Koe, 70Fos, 87Met]. For a laser rod this leads to a temperature distribution which is parabolic with respect to the radial position r if the thermal conductivity is constant and the medium is illuminated homogeneously by the excitation source. This results in a similar radial dependence of the index of refraction:

$$n(r) = n_0(1 - \gamma r^2) , \qquad (8.1.101)$$

where n_0 is the index of refraction at the center. To a first approximation, the refractive power D of a rod of length l is given by

$$D = 2\gamma n_0 \ell . \tag{8.1.102}$$

The refractive power is proportional to the pump power and inversely proportional to the pumped area A:

$$D = \frac{\alpha}{A} P_{\text{pump}} . \tag{8.1.103}$$

The proportionality constant α , referred to as the thermal lensing coefficient, is a figure of merit for the thermal properties of a laser material (Table 8.1.1).

Table 8.1.1. Thermal-lensing coefficients of solid-state laser rods and the corresponding refractive powers at 1 kW of pump power for a rod diameter of 10 mm. For flashlamp-pumped lasers, the pump power refers to the electrical power into the flashlamps. In the case of diode-pumping, the pump power is the optical pump power absorbed by the laser medium. The range of the thermal-lensing coefficient is generated by different pump conditions and variations in crystal quality and doping concentration.

Material	$\alpha \; [{\rm mm/kW}]$	$D \ [\mathrm{m}^{-1}]$
Flashlamp-pumped:		
Nd:YAG	$0.021 \dots 0.030$	$0.27 \dots 0.38$
Nd:glass	$0.15\ldots 0.19$	$1.91 \dots 2.42$
Nd:YAP	$0.037 \dots 0.069$	$0.47 \dots 0.88$
Nd:YLF	$-0.003^{\rm a}; 0.0014^{\rm b}$	-0.048; 0.018
Diode-pumped at 808 nm:		
Nd:YAG	$0.1 \dots 0.15$	$1.27 \ldots 1.91$
Nd:YVO ₄	$0.09\ldots 0.14$	$1.14 \dots 1.78$
Nd:YLF	$-0.015^{\rm a}; 0.008^{\rm b}$	-0.24; 0.09

^a σ -polarized at $\lambda = 1053$ nm.

^b π -polarized at $\lambda = 1047$ nm.

8.1.8.1 Fundamental-mode operation

In order to determine the propagation of the Gaussian beam (fundamental mode) inside the resonator, the concept of the equivalent g-parameters is applied [65Kog, 75Loe, 79Iff, 81Kor, 86Mag, 86Web, 87Mag, 93Hod4]. An optical resonator with an internal thermal lens (lens resonator) exhibits the same Gaussian beam radii at the mirrors as the equivalent empty resonator with the equivalent g-parameters g_i^* and the equivalent resonator length L^* (Fig. 8.1.53) with

$$g_i^* = g_i - Dd_i(1 - d_i/\rho_i), \quad i, j = 1, 2; \quad i \neq j,$$
(8.1.104)

$$g_i = 1 - (d_1 + d_2)/\rho_i , \qquad (8.1.105)$$

$$L^* = d_1 + d_2 - Dd_1 d_2 . ag{8.1.106}$$

Note that the term $d_1 + d_2$ is the effective resonator length L_{eff} of the resonator, which means that we could replace the thick lens by a thin lens with d_1 , d_2 being the distances to the two mirrors. The resulting ray-transfer matrix for a transit between the mirrors would be identical. By using the equivalent resonator parameters, the Gaussian beam propagation inside the resonator can be calculated by making use of the fact that the mirror surfaces are surfaces of constant phase for the beam.



Fig. 8.1.53. The lens resonator and its equivalent resonator. Both resonators exhibit identical Gaussian beam radii at the mirrors, but the beam propagation is different.



Fig. 8.1.54. Lens resonators with one internal lens move along straight lines through the equivalent g-diagram as the refractive power D is increased. The slope of the lines depends on the position of the medium and on the mirror curvatures.

The properties of a lens resonator can be visualized in the equivalent g-diagram (Fig. 8.1.54). The resonator starts at the point (g_1, g_2) and moves along a straight line through the diagram with increasing refractive power. In general, the lens resonator passes through stable zones and unstable zones. A lens resonator is called stable for $0 < g_1^* g_2^* < 1$ and unstable for $|g_1^* g_2^*| > 1$. Four characteristic refractive powers exist at which the resonator intersects stability limits (see Fig. 8.1.54):

(a)
$$g_1^* g_2^* = 1$$
 and $g_1^* > 0$: $D_{\rm I} = -\frac{1}{\rho_1 - d_1} - \frac{1}{\rho_2 - d_2}$, (8.1.107)

(b)
$$g_1^* = 0$$
: $D_{\text{II}} = -\frac{1}{\rho_1 - d_1} + \frac{1}{d_2}$, (8.1.108)

(c)
$$g_2^* = 0$$
: $D_{\text{III}} = \frac{1}{d_1} - \frac{1}{\rho_2 - d_2}$, (8.1.109)

(d)
$$g_1^* g_2^* = 1$$
 and $g_1^* < 0$: $D_{\rm IV} = \frac{1}{d_1} + \frac{1}{d_2}$. (8.1.110)

If $|D_{\text{II}}| < |D_{\text{III}}|$ holds, the lens resonator passes through the upper left unstable region in Fig. 8.1.54. Resonators with the property $D_{\text{II}} = D_{\text{III}}$ can reach the origin of the diagram (confocal resonator). These resonators are characterized by the condition

$$\frac{d_2}{d_1} = \sqrt{\frac{g_1}{g_2}} , \qquad (8.1.111)$$

which for a resonator with two identical mirrors means that the lens is located in the center. By using the equivalent resonator parameters, the beam propagation of the Gaussian beam with wavelength λ inside the lens resonator can be derived from the known beam radii at the resonator mirrors. The Gaussian beam exhibits two waists whose position and radius are a function of the refractive power. The following relations hold for the resonator shown in Fig. 8.1.55:



Fig. 8.1.56. Gaussian beam radius in the active medium $w_{\rm L}$, waist radii w_{0i} , and half angles of divergence θ_{0i} as a function of the refractive power of the thermal lens ($\rho_1 = \infty$, $\rho_2 = 3$ m, $d_1 = 0.3$ m, $d_2 = 0.5$ m).

 $w_{0i}^2 = \frac{\lambda L^*}{\pi} \frac{\sqrt{g_1^* g_2^* (1 - g_1^* g_2^*)}}{g_j^* (L^* / \rho_i)^2 + g_i^* (1 - g_1^* g_2^*)} \; .$

Beam radius at mirror i:

$$w_i^2 = \frac{\lambda L^*}{\pi} \sqrt{\frac{g_j^*}{g_i^* (1 - g_1^* g_2^*)}} . \tag{8.1.112}$$

Waist radii:

Beam radius at the principal planes:

Location of beam waist
$$i$$
:

Half angle of divergence:

$$w_{\rm L}^2 = w_1^2 \left[\left(1 - \frac{d_1}{\rho_1} \right)^2 + \left(\frac{d_1}{L^*} \right)^2 \frac{g_1^* (1 - g_1^* g_2^*)}{g_2^*} \right] \,. \tag{8.1.114}$$

$$L_{0i} = L^* \frac{g_j^* L^* / \rho_i}{g_j^* (L^* / \rho_i)^2 + g_i^* (1 - g_1^* g_2^*)} .$$
(8.1.115)

$$\theta_{0i} = \frac{\lambda}{\pi \, w_{0i}} \,. \tag{8.1.116}$$

The beam waist *i* is located to the left of mirror *i* for $L_{0i} < 0$ and to the right for $L_{0i} > 0$. In Fig. 8.1.55 L_{01} is positive and L_{02} is negative. Due to the thermal lens, the two angles of divergence are different because the two waist radii are not equal. The beam parameter product is the same on both sides of the resonator and is given by:

$$w_{01}\theta_{01} = w_{02}\theta_{02} = \frac{\lambda}{\pi} . \tag{8.1.117}$$

Unfortunately, the above shown expressions are too complicated to discuss in general; therefore, it is necessary to calculate the beam parameters to get a feeling for the beam characteristics. Figure 8.1.56 presents the beam radius in the medium $w_{\rm L}$ and the waist radii w_{0i} as a function of the refractive power. The stability limits are characterized by infinite or zero waist radii whereby

Landolt-Börnstein New Series VIII/1A2 (8.1.113)

the divergence exhibits a reciprocal behavior since the beam parameter product has to remain constant.

Typically, the Gaussian beam radius inside the medium exhibits a minimum which is approximately located in the middle of each stable zone. If ΔD denotes the range of the refractive power within which the resonator is stable, the minimum Gaussian beam radius in the medium reads:

$$w_{\rm L,min}^2 = \frac{4\lambda}{k\pi\,\Delta D}\,.\tag{8.1.118}$$

with k = 1 if the confocal point $g_1^* = g_2^* = 0$ is passed and k = 2 elsewhere. For the resonator in Fig. 8.1.56 and a wavelength of $\lambda = 1.064 \ \mu m$, we get a minimum beam radius of 0.531 mm in both stable zones. The Gaussian beam radius in the medium stays relatively constant over a range of the refractive power that covers about 90% of the stable zone. Within this range fundamental-mode operation can be attained if an aperture is placed in front of the medium whereby the aperture radius *a* needs to be adapted to the Gaussian beam radius:

$$a = 1.3 w_{\rm L,min}$$
 (8.1.119)

8.1.8.2 Transverse multimode operation

If no aperture is used inside the resonator, all transverse modes will oscillate whose beam radii are smaller than the radius b of the active medium [81Kor]. The number of oscillating modes is equivalent to the beam quality factor M^2 and can, to a good approximation, be calculated with

$$M^2 = \left(\frac{b}{w_{\rm L}}\right)^2,\tag{8.1.120}$$

where $w_{\rm L}$ is the Gaussian beam radius at the principal planes of the thermal lens. Since the Gaussian beam radius changes with increasing thermal lens power, the beam quality of multimode lasers is a function of the pump power. Insertion of (8.1.114) into (8.1.120) yields for the beam propagation factor:

$$M^{2} = \frac{\pi b^{2}}{\lambda L^{*}} \sqrt{\frac{g_{1}^{*}(1 - g_{1}^{*}g_{2}^{*})}{g_{2}^{*}}} \left[\left(1 - \frac{d_{1}}{\rho_{1}}\right)^{2} + \left(\frac{d_{1}}{L^{*}}\right)^{2} \frac{g_{1}^{*}(1 - g_{1}^{*}g_{2}^{*})}{g_{2}^{*}} \right]^{-1}.$$
(8.1.121)

This expression is only valid for $M^2 > 1$ since lower beam quality factors make no physical sense. At the stability limits the beam quality factor approaches $M^2 = 1$. For rectangular active media, the radius *b* has to be replaced with half the side length. In general, the beam quality factors in the *x*- and the *y*-direction are different. With the knowledge of the beam quality factor (8.1.121), the waist radii w_{mi} and the half angles of divergence θ_{mi} can be calculated using (8.1.113) and (8.1.116):

$$w_{mi} = \sqrt{M^2} w_{0i} , \quad \theta_{mi} = \sqrt{M^2} \theta_{0i}$$
 (8.1.122)

with the beam parameter product

$$w\theta = w_{m1}\theta_{m1} = w_{m2}\theta_{m2} = M^2 \frac{\lambda}{\pi} .$$
(8.1.123)

The locations of the waist radii are the same as for the Gaussian beam. It is interesting to note that the beam parameter product in multimode operation does not depend on the laser wavelength λ , since the beam quality factor scales with $1/\lambda$. We thus cannot improve the focusability by using



Fig. 8.1.57. Qualitative dependence of the beam parameter product $w\theta$ on the refractive power D for four resonators. In all resonators the thermal-lens power is varied from 0 to D_0 . Due to the different resonator geometries, the resonators pass through differently sized stable zones in the equivalent stability diagram [93Hod4].

a smaller wavelength. The beam parameter product of the Gaussian beam becomes lower with a smaller wavelength but the smaller Gaussian beam radius also leads to the oscillation of a higher number of transverse modes. Therefore, the beam parameter product remains constant. As previously discussed, the Gaussian beam radius in the medium exhibits a minimum near the center of a stable zone. Therefore, the multimode beam parameter product has a maximum at this point and decreases towards both stability limits. The maximum beam parameter product can be calculated by inserting (8.1.118) into (8.1.120):

$$(w\theta)_{\max} = k \; \frac{b^2 \Delta D}{4} \tag{8.1.124}$$

with

k = 1, if the origin of the equivalent stability diagram is passed,

k = 2, otherwise,

 ΔD : refractive power range within which the lens resonator is stable.

Figure 8.1.57 presents the qualitative dependence of the beam parameter product on the refractive power for four different resonators with different refractive power ranges ΔD and their different paths in the equivalent g-diagram. For stable resonators, the output power drops drastically as the resonators penetrate into unstable zones due to increasing diffraction losses at the active medium. The impetus of (8.1.124) is a relationship between maximum beam parameter product and range of pump power ΔP_{pump} within which the resonator is stable:

$$(w\theta)_{\max} = \frac{k}{4\pi} \,\alpha \,\Delta P_{\text{pump}} \,. \tag{8.1.125}$$

Since the output power is proportional to the pump power, the maximum beam parameter product is proportional to the maximum output power (provided that the resonator is stable at all pump powers). For diode-pumped Nd:YAG lasers, the maximum TEM_{00} -mode output power is about 30 W, higher output powers can only be achieved in multimode operation, with a typical increase in M^2 of 10 per 50 W of output power.

8.1.9 Ring resonators

8.1.9.1 General properties of ring resonators

In linear resonators, the interference of the two counterpropagating waves generates a modulation of the intensity along the optical axis. The intracavity intensity distribution of an axial mode of order q outside the active medium reads:

$$I(z) = I_0 \left[1 - \frac{2R}{1 + R^2} \cos(2\pi q z/L) \right] , \qquad (8.1.126)$$

where L is the resonator length, R is the reflectance of the output-coupling mirror, and z is the coordinate along the optical axis, starting at a mirror surface. The intensity minima exhibit a periodicity of half the wavelength of the axial mode $\lambda_q = 2L/q$. Due to gain saturation this results in a modulated axial gain distribution, referred to as *spatial hole burning*. The modulation frequency depends on the axial mode order.

In homogeneously broadened lasers, only one axial mode should theoretically be observed since the axial mode with the highest small-signal gain has access to the whole gain and will suppress the oscillation of adjacent modes. However, the spatial hole burning induces an axial mode competition because different modes are amplified in different areas of the gain medium. This may lead to multimode oscillation with random jumping between axial modes. Furthermore, the lower extraction efficiency at the intensity minima results in a decreased output power unless a high number of axial modes can coexist. In order to attain single-axial-mode operation with a high output power it is necessary to prevent spatial hole burning. In the twisted-mode resonator, this is accomplished by using orthogonal polarization states for the back- and the forth-traveling wave [65Evt]. Another way to generate a homogeneous intensity profile along the optical axis is unidirectional beam propagation in a ring resonator [63Tan, 65Her, 69Ana, 71Sch, 72Clo, 72Kog, 79Jar] (Fig. 8.1.58). By using an optical diode inside the resonator the axial modes are forced to propagate in one direction only. The optical diode can be realized with a retardation plate, a Faraday rotator, or a combination of both such that the losses generated at polarizers are higher for one propagation direction (Fig. 8.1.59). In asymmetric unstable ring resonators, the different propagation characteristics associated with the oppositely traveling waves can be used to generate unidirectional beam propagation [73Fre, 73Poz, 74Fre] (e.g. by inserting apertures). Unidirectionality can also be attained by means of an external feedback mirror which reflects the reverse oscillation [65Her]. However, with an external mirror, the ratio of the forward to the reverse power is only on the order of 1000:1 because a fraction of the power always has to propagate in the reverse direction to make this scheme work [74Fre]. A twisted-mode technique has to be applied in addition to suppress the residual hole burning.



Fig. 8.1.58. Concept of a stable ring resonator. Unidirectionality can be achieved with an intracavity optical diode or an external mirror that feeds back the wave having the wrong propagation direction.



Fig. 8.1.59. A cw diode-laser-pumped Nd:YAG laser with intracavity second harmonic generation. The Faraday rotator produces a polarization rotation by 7.5° at 1.064 µm. The same rotation is provided by the two half-wave plates. The two rotations add in one direction generating losses at the polarizer. In the intended propagation direction, the two rotations cancel. The laser provides an output power of 3.1 W at 532 nm in single-mode operation for a pump power of 14 W at 808 nm [96Mar] (\odot OSA 1996).

In general, ring resonators are set up with more than three mirrors to decrease the astigmatism induced by the mirror tilt. For an angle of incidence of θ , a curved mirror exhibits a focal length of $f \cos \theta$ in the tangential plane (paper plane in Fig. 8.1.58) and a focal length of $f/\cos \theta$ in the sagittal plane, where f is the focal length at normal incidence. By using more mirrors or a z-fold geometry (Fig. 8.1.59) the angles of incidence can be kept small. It is also possible to compensate the astigmatism by using special resonator designs [72Kog].

Similar to linear stable resonators, the Gaussian beam propagation in stable ring resonators can be evaluated by means of the Gaussian ABCD law. In order to include the astigmatism, four-dimensional ray-transfer matrices have to be used. Starting at an arbitrary plane, the ray-transfer matrix for a round trip is calculated. The Q^{-1} -matrix of the elliptical Gaussian beam is a solution of the generalized ABCD law:

$$Q^{-1} = (C + DQ^{-1})(A + BQ^{-1})^{-1}, \qquad (8.1.127)$$

where A, B, C, D are the 2×2 submatrices of the 4×4 ray-transfer matrix, and the superscript -1 denotes the inverse matrix. The matrix Q^{-1} contains the *q*-parameters of the Gaussian beam at the reference plane:

$$Q^{-1} = \begin{pmatrix} 1/q_{xx} & 1/q_{xy} \\ 1/q_{yx} & 1/q_{yy} \end{pmatrix} .$$
(8.1.128)

For planar ring resonators, the Gaussian beam is simple astigmatic, which means that the nondiagonal elements of Q^{-1} are zero. In this case, the beam propagation can be calculated separately for the sagittal plane and the tangential plane by using the common 2×2 round-trip matrices. The ring resonator is stable in one plane if the following relation holds for the elements of the corresponding ray-transfer matrix:

$$|\mathsf{A} + \mathsf{D}| < 2 \ . \tag{8.1.129}$$

In contrast to a linear resonator, the mirror surfaces of a ring resonator generally do not match the phasefronts of the Gaussian beam. The Gaussian beam waist is always located at that plane at which the round-trip matrix exhibits equal diagonal elements (A = D).



Fig. 8.1.60. Calculated maximum extraction efficiency of ring resonators with a homogeneously broadened active medium as a function of the optimum transmission of the output-coupling mirror. The curve parameters are the small-signal gain g_0l and the loss per transit α_0l .

The unidirectionality in ring resonators generates different beam properties as compared to those of linear resonators. First, the electric-field distribution is reproduced after one transit. Therefore, the frequency separation of the axial modes is twice that of linear resonators with length L:

$$\Delta \nu = \frac{c}{L} \,. \tag{8.1.130}$$

Secondly, light amplification takes place only once per round trip. The steady-state condition for ring oscillators, therefore, reads:

$$GRV_{\rm S} = 1$$
, (8.1.131)

where G is the gain factor, R is the reflectance of the output coupler, and $V_{\rm S}$ is the loss factor $(= 1 - \log s)$ per transit. For a linear resonator, the corresponding steady-state condition is given by $G\sqrt{R} V_{\rm S} = 1$. In order to reach laser threshold for the same output coupling, the pump power in a ring resonator has to be higher than in a linear resonator by a factor γ , with

$$\gamma = \frac{\ln(RV_{\rm S})}{\ln(\sqrt{R}V_{\rm S})} \,. \tag{8.1.132}$$

This implies that at identical pump powers, the output power of a ring oscillator is lower. However, this is only due to the fact that the ring resonator is overcoupled when the linear resonator exhibits optimum output coupling. If the output coupling is adjusted, the ring resonator provides the same maximum output power as the linear resonator. Maximum extraction efficiencies of ring resonators and the corresponding optimum output-coupling transmission are visualized in Fig. 8.1.60 for different values of the small-signal gain and the loss per transit. The maximum extraction efficiencies of ring resonators are identical to those attainable with linear resonators (see Fig. 8.1.50).

8.1.9.2 Unstable ring resonators

The ring geometry can also be applied to unstable resonator schemes as shown in Fig. 8.1.61 [69Ana, 73Poz, 73Fre, 74Fre]. Preferably, an asymmetric confocal resonator set-up is chosen which means that the transfer matrix for one round trip, starting at the output coupler, is different for the two propagation directions. The main advantage associated with the asymmetric design is a greater design flexibility. The flat mirrors in the collimated-beam path can be placed arbitrarily



Fig. 8.1.61. Asymmetric confocal unstable ring resonators. (a) confined output coupler, (b) scraper.

Fig. 8.1.62. Beam propagation of the diverging waves in a confocal unstable ring resonator with magnification M = 2. The resonator shown in Fig. 8.1.61a is represented as a lens waveguide. A mirror with radius of curvature ρ is replaced by a lens with focal length $f = \rho/2$. In the counterclockwise direction (ccw), the output beam (beam portion missing the biconcave lens) is collimated.

without changing the magnification of the resonator. This enables one to incorporate several active media or to increase the round-trip time. Unfortunately, there is no discrimination against one propagation direction since the forward and the reverse wave exhibit the same output-coupling losses. However, the different propagation characteristics of the two waves can be used to suppress one of the waves. The round trips in the resonator of Fig. 8.1.61a in the clockwise (cw) and the counterclockwise (ccw) directions are equivalent to the transits in the lens waveguides depicted in Fig. 8.1.62. In both directions, we can find two self-reproducing spherical wavefronts, referred to as the diverging wave and the converging wave. The converging wave in unstable resonators is generally not observed since it will transform itself into a diverging wave after a few round trips. Therefore, the asymmetric unstable ring resonator supports two diverging waves having opposite propagation directions. Although the asymmetry generates different propagation characteristics for these counterpropagating waves they exhibit the same magnification. The different propagation characteristics can be used to discriminate against one direction.

Unstable ring resonators have been realized in CO_2 lasers [69Ana, 73Fre, 73Poz, 74Fre, 88Cal], dye lasers [76Mar, 80Tes], and solid-state lasers [86Upp1]. The effect of thermal lensing on the performance of unstable solid-state ring lasers is discussed in [86Upp2]. A discrimination against the reverse-traveling wave can be achieved by locating the active medium where the forwardtraveling mode exhibits the larger mode volume [73Fre, 74Fre]. By using this technique, power ratios of 20:1 have been observed. Unidirectional operation, however, requires intracavity apertures to discriminate against the reverse oscillation. Alternatively, an optical feedback of the reversetraveling wave can be applied such that the radius of curvature and the propagation direction of the forward-traveling wave is matched. In [74Fre], the use of a feedback mirror increased the ratio of the forward to the reverse power of a CO_2 laser from 16:1 to 1,600:1.

The suppression of the reverse-traveling wave by means of intracavity apertures is very effective in negative-branch unstable ring resonators [78Pax, 83Oug, 84Skl, 86Upp1]. Both the forward and the reverse wave exhibit an intracavity focus at different locations in the resonator (Fig. 8.1.63a). By placing an aperture in the focus plane of the forward-traveling wave, a high loss is generated for the reverse-traveling wave. The aperture can also serve as a spatial filter to filter out phase distortions



Fig. 8.1.63. Beam propagation in negative-branch unstable ring resonators. (a) Fourier-transform resonator with the forward wave (solid line) and the reverse wave (broken line), (b) Forward-traveling wave in a self-imaging unstable resonator with $L = f_1 + f_2$. The reverse-traveling wave (not shown) is collimated in the upper section and it exhibits a focus at the scraper.

and smooth out the spatial intensity distribution. A special negative-branch ring resonator is the self-imaging unstable resonator as depicted in Fig. 8.1.63b [78Pax, 84Skl].

8.1.9.3 Nonplanar ring resonators

In planar ring resonators with the mirrors tilted around the same axis, s-polarized and p-polarized light represent the polarization eigenstates for a resonator round trip. The phase shifts induced at the mirror generally are different for the two polarizations, but no mixing between s- and p-polarization can occur. This is different in the nonplanar geometry in which the mirrors are tilted around different axes. The polarization properties of nonplanar rings can be evaluated by using the Jones matrix formalism [86Sie, 89Kin, 89Nil]. It can be shown that nonplanar ring resonators act as polarization rotators having left-circularly and right-circularly polarized light as the eigenstates. When combined with a Faraday rotator and a polarizing element, the rotation of the polarization plane at the mirrors can be used to discriminate against one propagation direction.

A widely used laser employing this technique is the diode-laser-pumped monolithic nonplanar Nd:YAG ring oscillator [85Kan, 87Kan, 88Nil, 88Sch, 89Nil, 91Che] (Fig. 8.1.64). A dc magnetic field is used to generate a polarization rotation via the Faraday effect (Verdet constant of YAG: $103^{\circ}/(\text{Tm})$). In the clockwise direction, the rotation of the polarization plane caused by the internal reflections at the surfaces B, C, and D and the Faraday rotation cancel, giving an s-polarized eigenmode at the front surface. In the opposite direction, the two rotations add, resulting in an elliptically polarized eigenmode. The reflectance R of the front surface (Fresnel-reflection) is slightly higher for the s-polarization than for the p-polarization. The wave in the ccw direction, therefore, has a higher laser threshold and is suppressed. The suppression also works if the two polarization rotations do not cancel in the clockwise direction. In [88Sch], the 17° rotation caused by the internal



Fig. 8.1.64. Beam propagation in a diode-pumped monolithic non-planar Nd:YAG ring oscillator. The size of the crystal is $11 \times 7 \times 3 \text{ mm}^3$. (a) side view, (b) top view [88Sch] (© AT Fachverlag 1988).

reflections was superimposed by a 0.6° Faraday rotation. At the coated front surface of the crystal, the resulting difference in polarization for the two propagation directions generated a loss difference of 0.6%. This small difference was sufficient to realize unidirectional single-mode operation with an output power of 100 mW (400 mW optical pump power).

Due to the mechanical stability and the highly stable pump source, the monolithic Nd:YAG ring resonator has the capability of providing single-mode operation with narrow linewidths and high output stability. Commercial lasers provide output powers in the 100 mW range with linewidths of less than 3 kHz and an rms stability of better than 0.05%. With active frequency stabilization, very narrow linewidths down to the sub-Hertz range have been achieved [88Nil, 89Day, 92Day].

8.1.10 Waveguide resonators

8.1.10.1 Motivation

In all high-power laser systems, the removal of the heat generated by the pump process is crucial for efficient operation. In sealed-off gas lasers, the gas is mainly cooled through heat conduction to the walls of the tube, a process which is not very efficient if cylindrical discharge tubes are used. The cooling can be approved by at least one order of magnitude if a discharge geometry is used that exhibits a larger ratio of the cooling surface to the discharge volume. Besides the annular geometry, the most promising concept pursued in the last decade is the slab waveguide laser, as depicted in Fig. 8.1.65. Heat removal in a slab laser scales inversely with the electrode separation d. Typical electrode distances for CO_2 slab waveguide lasers are between 1.5 mm and 2.5 mm. The discharge volume can be controlled via the cross-sectional area of the electrodes. This means that the output power of slab lasers is scaled by area rather than length as in conventional geometries.

Assuming a constant laser efficiency η_{tot} of 10%, the scaling law for the maximum achievable output power of a waveguide CO₂ laser per area of one electrode can be approximated by [91Now]:

$$P_{\rm out,max} = \frac{3}{d} \frac{W}{\rm cm^2} ,$$
 (8.1.133)

where the electrode distance d is given in mm. This expression holds only for distances d greater than 1.5 mm since smaller electrode gaps will decrease the laser efficiency due to an increase in diffraction losses and a decrease of the excitation efficiency. The small electrode separation puts a severe constraint on the resonator design. In the larger dimension an unstable resonator can

Cooling water Output coupler Rear mirror Electrodes Laser beam

Fig. 8.1.65. CO_2 slab waveguide laser with unstable waveguide resonator and a RF excited gas discharge (RF frequencies between 50 and 100 MHz) [91Now] (© AT Fachverlag 1991). Electrode separations between 1.5 and 2.5 mm are used. The width and the length of the discharge regions typically are 4 ... 7 cm and 60 ... 100 cm, respectively. Electrode materials commonly used are aluminum and copper.

be utilized to achieve near-diffraction-limited beam quality. Unfortunately, in the perpendicular direction the small gap width in combination with the length of the electrodes makes it virtually impossible to propagate a Gaussian beam through the discharge region without touching the electrodes. A stable resonator used in this direction, therefore, will exhibit modes that are guided by the electrodes rather than modes that follow the laws of free-space propagation as in conventional open resonators. This resonator concept using guided-mode propagation is referred to as a waveguide resonator [64Mar, 76Laa, 87Hal]. Guided-mode propagation occurs if the Fresnel number of the waveguide $a^2/(\lambda L)$ is much lower than 1, where L is the length of the waveguide and 2a its diameter. Waveguide resonators comprise two external mirrors and the mode properties result from a combination of the mode propagation between the waveguide's ends and the resonator mirrors.

8.1.10.2 Eigenmodes of hollow rectangular waveguides

In waveguide lasers, the cross section of the waveguide is very large compared to the wavelength λ . It is thus justified to neglect terms of first order or higher order in λ/a and λ/b . With this approximation, the fields in the walls and at the wall surface vanish and the internal fields become transverse. The following electric fields $\mathbf{E}(x, y, z, t) = (E_x, E_y, E_z)$ and magnetic fields $\mathbf{H}(x, y, z, t) = (H_x, H_y, H_z)$ are obtained for the rectangular waveguide shown in Fig. 8.1.66, provided that the waveguide length is large compared to its width and its height [76Laa]. The *y*-polarized hybrid modes $E^y H_{mn}^x$ read:

$$E_{x} = E_{z} = 0, \quad H_{y} = H_{z} = 0,$$

$$E_{y} = \frac{1}{\sqrt{ab}} \exp[i(\beta_{mn}z - \omega t)] \exp[-\alpha_{mn}z] \begin{bmatrix} \cos\left[\frac{m\pi x}{2a}\right] & \cos\left[\frac{n\pi y}{2b}\right] \\ \cos\left[\frac{m\pi x}{2a}\right] & \sin\left[\frac{n\pi y}{2b}\right] \\ \sin\left[\frac{m\pi x}{2a}\right] & \cos\left[\frac{n\pi y}{2b}\right] \\ \sin\left[\frac{m\pi x}{2a}\right] & \sin\left[\frac{n\pi y}{2b}\right] \end{bmatrix}, \quad m, n \text{ odd},$$

$$m \text{ odd}, n \text{ even},$$

$$H_{x} = -\sqrt{\frac{\varepsilon_{0}}{\mu_{0}}} E_{y} \qquad (8.1.134)$$

where $\omega = 2\pi c_0/\lambda$ is the angular light frequency, ε_0 is the permittivity of free space, μ_0 is the permeability of free space, and m, n are integer numbers with $m, n \ge 1$. For $\lambda \ll a$ and $\lambda \ll b$, the propagation constants are given by:



Fig. 8.1.66. A waveguide resonator with a hollow, rectangular waveguide.

$$\alpha_{mn} = \frac{\pi^2}{4k^2} \left[\frac{m^2}{a^3} \operatorname{Re}\left(\frac{1}{\sqrt{\varepsilon - 1}}\right) + \frac{n^2}{b^3} \operatorname{Re}\left(\frac{\varepsilon}{\sqrt{\varepsilon - 1}}\right) \right] , \qquad (8.1.135)$$

$$\beta_{mn} = k \left[1 - \frac{\pi^2}{8k^2} \frac{m^2}{a^2} \left(1 - \frac{2}{ak} \operatorname{Im}\left(\frac{1}{\sqrt{\varepsilon - 1}}\right) \right) - \frac{\pi^2}{8k^2} \frac{n^2}{b^2} \left(1 - \frac{2}{bk} \operatorname{Im}\left(\frac{\varepsilon}{\sqrt{\varepsilon - 1}}\right) \right) \right] , \quad (8.1.136)$$

where $k = 2\pi/\lambda$ is the wave number and Re, Im denote the real part and the imaginary part of the expression in the bracket, respectively. These field distributions represent the eigenmodes if the following relations hold:

$$\frac{m\lambda}{4a} \ll 1 , \quad \frac{n\lambda}{4b} \ll 1 , \tag{8.1.137}$$

$$\frac{m\lambda}{4a} \ll \left|\sqrt{\varepsilon - 1}\right| \ , \quad \frac{n\lambda}{4b} \ll \left|\frac{\sqrt{\varepsilon - 1}}{\varepsilon}\right| \ . \tag{8.1.138}$$

Similar expressions are obtained for x-polarized hybrid modes $E^{x}H_{mn}^{y}$ [76Laa].

Both sets of hybrid modes (x- and y-polarized) are orthonormal:

$$\int_{-b}^{+b} \int_{-a}^{a} E_{mn} E_{m'n'}^* \, \mathrm{d}x \, \mathrm{d}y = \delta_{mm'} \delta_{nn'} \exp[-2\alpha_{mn} z] \,, \tag{8.1.139}$$

where E_{mn} stands for E_y , E_x for y-polarized modes and x-polarized modes, respectively. The eigenmodes form a complete set and can thus be used as a basis set for a mode expansion. Any continuous field distribution E(x, y) launched into the waveguide can be expanded in a linear combination of the waveguide eigenmodes. Each eigenmode can be described as a superposition of two plane waves being reflected off the waveguide walls at an angle θ_m :

$$\theta_m = \frac{m\lambda}{4a} \,. \tag{8.1.140}$$

For all waveguide modes, the intensity maxima can be understood as a result of constructive interference of the two plane waves (Fig. 8.1.67). Waves propagating at any other angles than the ones given by (8.1.140) do not generate a constant intensity distribution in the z-direction due to a varying phase shift between the waves. The waveguide eigenmodes are determined by the fact that each of the plane waves has to interfere constructively with the other wave after being reflected by the wall. The propagation angles for the fundamental waveguide mode (m = 1) of CO₂ lasers typically are between 1 and 3 mrad. We are thus dealing with reflection at grazing incidence.



Fig. 8.1.67. The eigenmodes of a rectangular waveguide can be considered a superposition of two plane waves propagating at angles $\pm \theta_m$ to the z-axis. The horizontal, broken lines indicate the location of the intensity maxima. Note the phase difference of $\lambda/2$ between the two waves at the wall.



Fig. 8.1.68. Reflection of a plane wave at the interface of an absorbing medium. The power loss of p-polarized and s-polarized waves as a function of the angle θ between the wave vector and the surface is shown for different refractive indices n of the medium (calculation using Fresnel equations).

Table 8.1.2. Refractive indices and calculated absorption coefficients of the x-polarized (α_1^x) and the y-polarized (α_1^y) fundamental mode for different waveguide materials (one-dimensional waveguide in the x-direction with a = 1 mm and $\lambda = 10.6 \text{ µm}$, using (8.1.135) and the corresponding equation for the x-polarized mode). The refractive indices $n = n_1 + i n_2$ of the metals are from [95Paq] (extrapolated data). The refractive indices of the dielectrics are from [92Khe].

Material	n_1	n_2	$\alpha_1^x \ [1/\mathrm{mm}]$	$\alpha_1^y \; [1/\mathrm{mm}]$
Cu	15	63	1.054×10^{-4}	2.51×10^{-8}
Au	18	67	1.264×10^{-4}	2.63×10^{-8}
Al	28	95	1.966×10^{-4}	2.00×10^{-8}
Al_2O_3	0.67	0.136	2.238×10^{-6}	1.40×10^{-6}
BeO	0.64	0.59	4.688×10^{-6}	2.13×10^{-6}
SiO_2	1.9	0.73	1.506×10^{-5}	3.37×10^{-6}

An accurate treatment of the reflection at the interface of an absorbing medium reveals that at grazing incidence both p- and s-polarized waves experience a phase shift of π . Taking this phase shift into account, a zero phase shift between incoming wave 1 and the reflected wave 2 can only be realized at the angles θ_m . The absorption loss of waveguide modes can now be simply explained by applying the Fresnel equations for the reflection and absorption of light at the interface of an absorbing medium. For grazing incidence, the power loss due to absorption increases proportionally to the angle between the wave vector and the surface (Fig. 8.1.68).

In Table 8.1.2 refractive indices and calculated absorption coefficients of the x-polarized (α_1^x) and the y-polarized (α_1^y) fundamental mode for different waveguide materials are given.

Waveguide modes exhibit features that support high-power operation of a waveguide laser with diffraction-limited beam quality. The beam diameters of the transverse modes scale with the waveguide cross section. Therefore, all modes have access to the same gain region, generating a strong mode competition. In addition, due to the reflection losses, the active medium (waveguide) provides a transverse-mode discrimination proportional to m^2/a^3 . Since higher-order modes exhibit higher absorption losses, only a limited number of transverse waveguide modes (let us say not more than 10) will oscillate in a waveguide laser. Since waveguide modes are not eigensolutions of free-space propagation, the mode of the waveguide resonator is a combination of several



Fig. 8.1.69. Approximation of the field of the fundamental waveguide mode EH_{11} by a Gaussian beam with beam radius $w_0 = 0.703 \ a.$

waveguide modes. In order to attain a diffraction-limited laser beam, the lowest-loss mode of the waveguide resonator should closely resemble the fundamental waveguide mode. The goal of waveguide resonator design, of course, is to attain a near-diffraction-limited output beam, preferably with properties similar to those of a Gaussian beam. Since the field distributions of the waveguide modes form a complete orthogonal set, the one-dimensional field distribution $E_m(x)$ of the *m*-th resonator mode at the end of the waveguide can be written as a linear combination of waveguide mode field distributions $\Phi_n(x)$:

$$E_m(x) = \sum_{n=1}^{\infty} c_{mn} \Phi_n(x) , \qquad (8.1.141)$$

where the c_{mn} are complex constants. The propagation of this field distribution to the mirror and back to the waveguide will generate higher losses for the high-order waveguide modes since they propagate at higher angles of divergence. The free-space propagation thus provides an additional waveguide-mode discrimination. By using appropriate mirror curvatures and mirror distances, it should, therefore, be possible to generate a lowest-loss resonator mode exhibiting most of its power in the fundamental waveguide mode. The fundamental waveguide mode EH₁₁ can be approximated by a Gaussian beam with beam radius $w_0 = 0.703 \ a$ resulting in a power overlap of about 98% (Fig. 8.1.69). This Gaussian beam could be similar to the fundamental resonator mode, for instance, if the waveguide end was imaged onto itself by means of the resonator mirror. Higher-order waveguide modes would be suppressed because they are coupled back into the waveguide less efficiently. The same statement holds for circular, hollow waveguides, whose mode properties can be found in [61Sni, 64Mar, 74Mar, 76Deg, 79Abr]. The eigenmodes of coaxial waveguides can be found in [93Ehr].

8.1.10.3 Properties of waveguide resonators

With the knowledge of the waveguide eigenmodes and their propagation constants, it is possible to evaluate the mode properties of the waveguide resonator itself. The general waveguide resonator consists of two mirrors with radii of curvature ρ_1 and ρ_2 positioned at a distance d_1 and d_2 from the waveguide ends (Fig. 8.1.70). The free-space propagation from the waveguide to the mirrors



Fig. 8.1.70. Schematic of a waveguide resonator.

and back generates coupling losses for the waveguide modes. The goal of the waveguide resonator design is to find configurations that provide a fundamental resonator mode with low coupling losses and a good mode discrimination against higher-order modes.

8.1.10.3.1 Waveguide resonator configurations

Three different configurations can be distinguished: Case I, Case II, and Case III resonators.

8.1.10.3.1.1 Case I resonators

The resonator mirrors are placed close to the waveguide ends, at a distance of typically less than 10 times the waveguide dimension. The resonator modes can be closely approximated by the waveguide eigenmodes. All resonator modes experience a round-trip loss that is determined by the absorption loss $2\alpha L$. The mode discrimination is rather low and, depending on the small-signal gain, the output coupling, the type of line broadening, and the dimensions of the waveguide, the output beam will be a combination of several transverse waveguide modes. Assuming a refractive index of n = 1 of the gas, the resonance frequencies of the modes are determined by:

$$2\beta_{nm}L = q \ 2\pi \,, \tag{8.1.142}$$

where q is the axial mode order, L is the waveguide length, and β_{nm} is the propagation constant according to (8.1.136). In order to achieve fundamental-mode operation, it is necessary to choose the waveguide cross section a small so that higher-order modes cannot reach the laser threshold. Unfortunately, the coupling losses of the fundamental waveguide mode will increase too, resulting in a decrease in laser efficiency.

8.1.10.3.1.2 Case II resonators

The waveguide ends are imaged onto themselves by the resonator mirrors. Imaging requires a distance d_i of each mirror equal to its radius of curvature ρ_i . Each waveguide mode is imaged back onto itself and no coupling losses occur for infinite mirrors. However, the transverse mode discrimination is as poor as in true Case I resonators.

8.1.10.3.1.3 Case III resonators

In both, rectangular and cylindrical waveguides, the fundamental waveguide mode can be approximated by a Gaussian beam with 98% power overlap (Fig. 8.1.71). Propagation of the fundamental waveguide mode can be described by Gaussian beam propagation of a Gaussian beam with waist radius $w_0 = 0.703 \ a$ (rectangular symmetry) or $w_0 = 0.6345 \ a$ (circular waveguide). Given this assumption, the wavefront curvature of the mode at a distance z from the waveguide reads:



Fig. 8.1.71. Power coupling coefficients for the EH_{1m} modes of circular waveguides as a function of the ratio of the beam waist radius w_0 of the Gaussian beam to the guide radius a [92Jen] (© OSA 1992).

$$R(z) = 2z_0 \left[\frac{z}{z_0} + \frac{z_0}{z} \right] , \qquad (8.1.143)$$

where $z_0 = \pi w_0^2 / \lambda$ is the Rayleigh range. By choosing the mirror curvature ρ equal to R(z) at the mirror position d = z, the beam propagation is reversed and the EH₁₁ mode will be coupled back into the waveguide with low losses. This configuration combines low fundamental-mode loss and excellent transverse-mode discrimination since higher-order modes will not couple back efficiently. It is for this reason that most waveguide resonators investigated in the past were based on this principle. Numerical calculation of the coupling losses of the EH₁₁ mode indicated that positioning the mirrors at $d = z_0$ with $\rho = 2z_0$ yields the lowest loss.

8.1.10.3.2 Calculated round-trip losses of the lowest-loss resonator mode

Calculated round-trip losses of the lowest-loss resonator mode as a function of d/ρ are presented in Fig. 8.1.72 for symmetric resonators and circular waveguides. The parameter of the curves is $\alpha = 2\pi a^2/(\lambda \rho)$. The meaning of this parameter and others shown in the graph shall be explained in the next paragraph. Case I and Case II resonators are given by $d/\rho = 0$ and $d/\rho = 1$, respectively.



Fig. 8.1.72. Calculated roundtrip loss (lowest-loss mode) for symmetric waveguide resonators with a circular waveguide of radius a and effective length L' as a function of the mirror distance-toradius-of-curvature ratio d/ρ . The curve parameter is the inverse of the normalized mirror curvature $\rho/(ka^2)$; K is the wave number; the function X is explained in the text below. Both mirrors have a radius of curvature ρ and are located at a distance d from the adjacent waveguide end [90Alv] (© IEEE 1990).

Case III coupling with $d = z_0$ is equivalent to $\alpha = 2.41$ and $d/\rho = 0.5$. Although the curve with $\alpha = 2.41$ is not shown, it is obvious that Case III resonators provide similar low losses as the other two resonators. However, this graph indicates that low losses can also be realized with resonators that do not belong to one of the three resonator classes. These additional loss minima are a result of the interference between different waveguide modes. A realistic description of waveguide resonator properties generally requires the incorporation of multiple waveguide modes.

8.1.10.3.3 Mode properties of a general waveguide resonator

The mode properties of a general waveguide resonator are determined by seven independent parameters. For the free-space propagations, four independent parameters can be found in the phase term of the two Collins integrals. Those parameters are: ka^2/d_1 , ka^2/d_2 , d_1/ρ_1 , and d_2/ρ_2 . The independent parameters governing the waveguide propagation are contained in the propagation constant (8.1.135)/(8.1.136): L'/ka^2 , L/ka^2 , and Re $(1/\sqrt{\varepsilon-1})/ka$. A generalized presentation of waveguide resonator properties, therefore, is a seven-dimensional problem. However, in most cases the effective length L' differs from the geometrical length L by no more than a few percent and we may skip L as a parameter. In both circular waveguides and one-dimensional rectangular waveguides, we are dealing with the following six independent, dimensionless parameters:

-	the	effective	length	parameter:	
---	-----	-----------	--------	------------	--

- the effective length parameter:
$$\frac{L'}{ka^2}$$
,
- the distance parameters: $\frac{d_1}{ka^2}$, $\frac{d_2}{ka^2}$,

- the distance-curvature ratios:

- the waveguide loss parameter:

$$\frac{a_1}{\rho_1}, \frac{a_2}{\rho_2},$$

$$\frac{\operatorname{Re}(X)}{ka} \quad \text{with} \quad X = \begin{bmatrix} \varepsilon/\sqrt{\varepsilon - 1} \\ 1/\sqrt{\varepsilon - 1} \\ (\varepsilon + 1)/\sqrt{2(\varepsilon - 1)} \end{bmatrix} \quad \begin{array}{c} \operatorname{rect.}, x\text{-pol.}, \\ \operatorname{rect.}, y\text{-pol.}, \\ \operatorname{circular}. \end{array}$$

In order to simplify the discussion, we focus on resonators with two flat mirrors. This restriction leaves us with four independent parameters. The properties of waveguide resonators with curved mirrors can be found in [72Abr, 73Deg, 85Hil].

8.1.10.4 Waveguide resonator mode and loss calculations

Since the waveguide modes and the free-space modes (Gauss-Laguerre polynomials in circular symmetry and Gauss-Hermite polynomials in rectangular symmetry) both form complete sets, they can be used as basis sets for the resonator modes [84Ger, 88Hil, 89Jac2]. In general, a limited number of functions (on the order of 10 waveguide modes and 50 free-space modes) from both sets are taken into account. This is mathematically not correct, but the truncation errors can be kept in the sub-percent range if only the first three lowest-loss resonator modes are considered. For simplicity we restrict the discussion to one-dimensional rectangular waveguides [88Hil, 89Jac2]. A similar treatment for circular waveguides can be found in [84Ger]. The field distribution $E_n(x)$ of the p-th resonator mode at the end of the waveguide is written as a linear combination of N waveguide eigenmode field distributions $\Phi_n(x)$ and as a linear combination of M free-space modes $\Psi_m(x)$. For the free-space modes we use Gauss-Hermite polynomials with the beam waist with radius w_0 located at the waveguide end. The Gaussian beam waist radius w_0 is preferably chosen in the range $0.6a < w_0 < 0.8a$.



Fig. 8.1.73. Calculated resonator losses versus the effective-length parameter for the three lowest-loss modes of a symmetric waveguide resonator with flat mirrors (circular waveguide, a = 1 mm, $\lambda = 10.6 \mu\text{m}$, refractive index of the wall material: $n = 0.67 + i \ 0.136$ (alumina), mirror distances: $d_1 = d_2 = 20 \text{ mm}$, $d/(ka^2) = 0.0337$. Six hybrid modes EH_{1m} and 12 free-space modes TEM_{0m} were used as basis sets. An effective-length parameter L/ka^2 of 1.0 corresponds to a waveguide length of L = 593 mm. See text for an explanation of the vertical lines [94Pet] (© SPIE 1994).

Calculated round-trip losses of a symmetric, circular waveguide resonator with flat mirrors using the matrix method are presented in Fig. 8.1.73 for $\lambda = 10.6 \ \mu\text{m}$. The length of the 2 mm-crosssection aluminum waveguide is varied, but the distance from the waveguide ends to the mirrors remains fixed at $d = 20 \ \text{mm}$. Over several relatively wide ranges of the waveguide length, low losses of the fundamental mode and an excellent discrimination against the next-order mode can be simultaneously realized. In these length ranges the lowest-loss mode is predominantly EH_{11} with small amounts of higher-order waveguide modes. Operation near points of degeneracy, like at $L/ka^2 = 0.25$ or 0.5, has to be avoided since the output beam may switch between EH_{11} and EH_{12} -like intensity distributions. The most surprising behavior in this graph is the periodic oscillation of the loss curves. This structure can be attributed to interference among two or more waveguide eigenmodes.

At the loss minima, two or more waveguide modes emerge from the guide with relative amplitudes and phases such that the spot size of the beam coupled back into the waveguide assumes a minimum. The coupling loss at that point is lower than the coupling loss of the EH_{11} mode discussed earlier. These resonances were investigated in detail in [84Ger] (circular guide) and [88Hil] (rectangular guide) using two waveguide modes.

The periodicity of the resonances in $L/(ka^2)$ can be attributed to the periodic regeneration of the relative phase between the modes at the waveguide ends. Although the resonator modes in Fig. 8.1.73 represent combinations of more than two waveguide modes, most loss minima of the lowest-loss mode can be assigned to a resonance between two of the three waveguide modes EH_{11} , EH_{12} , and EH_{13} . The broken vertical lines indicate resonances that can be attributed to a mixing of EH_{11} and EH_{13} .

The waveguide resonator mode calculations rely on the exact knowledge of the refractive index of the guide material, a requirement that can be hardly met, especially for metals. Furthermore, the finish of the waveguide walls and heat-induced variations in the waveguide dimensions also have a strong influence on the mode properties. In addition, the mode calculations hold only for the passive resonator. The gain saturation and inhomogeneous gain profiles may considerably change the mode properties. Therefore, the numerical evaluation of the resonator properties can only be considered a starting point of the laser design, and the optimization of the laser properties have to be accomplished experimentally. Furthermore, in gas lasers with relatively large gain bandwidths (like CO_2 lasers), wavelengths other than the desired one may exhibit lower round trip losses resulting in the oscillation at the unwanted wavelength or a switching between wavelengths. This stems from the dependence of the refractive index on the wavelength, which is particularly strong for metallic waveguides.

The dimensions of the waveguide are usually determined by constraints on output power and heat removal which means that the waveguide cross section can only be varied within a small range. With the dimensions of the gain medium being more or less fixed, a first step in the resonator design would be to place the resonator mirrors as close as possible top the waveguide ends. Preferably, the resonator parameters should be chosen such that the resonator is working between two resonances of the EH_{12} and EH_{13} modes since the second-order resonator mode is most likely to exhibit a loss minimum near these resonances. For small distances d_1, d_2 between the mirrors and the rectangular waveguide, the absolute locations of the mode-mixing points between mode EH_{mn} and $EH_{m'n'}$, to a good approximation, are given by [88Hil]:

$$\frac{L+d_1+d_2}{ka^2} = \frac{16 \ p}{\pi \left|m^2 - m'^2 + n^2 - n'^2\right|}, \quad p = 1, 2, \dots$$
(8.1.144)

The experimental set-up can then be optimized experimentally by varying the mirror distances, the mirror curvatures, and, to some extent, the waveguide cross section. In general, waveguide lasers provide excellent beam properties, combining a high fill factor, good mode control, and beam qualities near the diffraction limit. In the first two decades since the first reported implementation of a waveguide laser in 1971 [71Smi], mostly circular waveguides were used. In the late 80s and early 90s the interest shifted to slab waveguides due to the area scaling of the output power and the realization of near-diffraction-limited beam quality using off-axis unstable resonator schemes [89Jac1, 90Nis, 90Now, 91Now, 92Col, 94Pet, 95Eis, 96Lap]. Several commercial CO₂ slab waveguide lasers are used in medical applications (100 W) and in material processing with output powers in the kW range [88Tul, 90Nis, 92Hob, 92Dal, 92Yar, 92Mef1, 92Mef2, 94Pet, 96Lap]. In order to achieve high output power with a compact laser head, folded waveguide designs are often used. Popular folding schemes are the U-fold, the V-fold, and the Z-fold, all three using plane or curved folding mirrors. The theoretical modeling of folded waveguide resonators is more complicated especially if curved mirrors are used [88Hil, 89Jac2, 90Hil].

8.1.10.5 Properties of slab waveguide lasers

One possible set-up of a slab waveguide resonator is depicted in Fig. 8.1.74. The positive-branch confocal unstable resonator in off-axis geometry with magnification M provides near-diffractionlimited beam quality along the wide dimension of the waveguide. The high-reflecting cylindrical mirrors form a one-dimensional Case I waveguide resonator with flat mirrors in the perpendicular direction. Typical dimensions for CO₂ lasers are: b = 30...50 mm; L = 60...120 cm; $d_1, d_2 =$ 15...30 mm; 2a = 1.5...2.5 mm. Mirror distances of less than 15 mm are not feasible due





to possible damage by the discharge plasma. In commercial systems, negative-branch confocal unstable resonators are preferred due to their lower misalignment sensitivity even though the efficiency is about 20% lower. In the negative branch, with both mirrors being concave in the unstable direction, spherical mirrors can also be used resulting in lower coupling losses for the waveguide modes.

The design of the unstable resonator can be performed geometrically since the equivalent Fresnel numbers are relatively high. The output-coupling loss is almost identical to the geometrical loss of 1 - 1/M for equivalent Fresnel numbers higher than 10. Typical equivalent Fresnel numbers are between 20 and 60, and magnifications in the range from 1.15 to 1.4 are used. The mirror curvatures needed to set up a confocal unstable resonator with magnification M are given by:

$$\rho_1 = \frac{2(L+d_1+d_2)}{1-M} , \qquad (8.1.145)$$

$$\rho_2 = \frac{-2M(L+d_1+d_2)}{1-M} , \qquad (8.1.146)$$

where $M = -\rho_2/\rho_1$ is positive for positive-branch resonators and negative for negative-branch resonators. The equivalent Fresnel number reads:

$$N_{\rm eq} = \frac{b^2}{2L\lambda} (M - 1) . \tag{8.1.147}$$

The electric field coupled out of the confocal unstable resonator exhibits a phase variation of less than $\lambda/5$ over the entire cross section of the beam (Fig. 8.1.75) resulting in a near-diffraction-limited beam quality. However, due to the different beam dimensions in the x- and the y-directions, the beam is highly astigmatic. The angle of divergence in the waveguide direction is about (|M| - 1) b/(2a) higher than in the perpendicular direction. In order to generate a near-circular focus, the astigmatism has to be compensated by means of cylindrical lenses. Typically the far-field intensity profiles of the transformed, stigmatic beams are close to Gaussian in the waveguide direction $(M^2 \sim 1.1)$ and exhibit small side lobes in the unstable direction with a 10 ... 20% higher beam propagation factor.



Fig. 8.1.75. Calculated intensity and phase distributions in the near field and in the far field for an off-axis confocal unstable resonator with equivalent Fresnel number $N_{\rm eq} = 20.8$ and magnification M = 1.15 ($\lambda = 10.6 \ \mu m$). The intracavity near-field intensity distribution at the output coupling mirror is shown in (a) and the corresponding phase distribution is given in (b). The shaded area indicates the portion being coupled out of the resonator. In the waveguide direction a homogeneous phase and amplitude profile of width 1.5 mm is assumed. Graph (\mathbf{c}) shows the intensity distribution in the far field. The beam is astigmatic due to a higher divergence in the waveguide direction [91Now] (© AT Fachverlag 1991).

References for 8.1

- 61Boy Boyd, G.D., Gordon, J.P.: Bell Syst. Tech. J. **40** (1961) 489.
- 61Fox Fox, A.G., Li, T.: Bell. Syst. Tech. J. 40 (1961) 453.
- 61Sni Snitzer, E.: J. Opt. Soc. Am. **51** (1961) 491.
- 62Boy Boyd, G.D., Kogelnik, H.: Bell. Syst. Tech. J. 41 (1962) 1347.
- 63Fox Fox, A.G., Li, T.: Proc. IEEE **51** (1963) 80.
- 63Rig Rigrod, W.: J. Appl. Phys. **34** (1963) 62.
- 63Tan Tang, C.L., Statz, H., deMars jr., G.: J. Appl. Phys. **34** (1963) 2289.
- 64Abr Abramowitz, M., Stegun, A.: Handbook of mathematical functions, New York: Dover Publ., 1964.
- 64Gor Gordon, J.P., Kogelnik, H.: Bell Syst. Tech. J. 43 (1964) 2873.
- 64Mar Marctili, E.A.J., Schmeltzer, R.A.: Bell. Syst. Tech. J. 43 (1964) 1783.
- 65Evt Evtuhov, E., Siegman, A.E.: Appl. Opt. 4 (1965) 142.
- 65Her Hercher, M., Young, M., Smoyer, C.B.: J. Appl. Phys. 36 (1965) 3351.
- 65Heu Heurtley, J.C., Streifer, W.: J. Opt. Soc. Am. 55 (1965) 1472.
- 65Kog Kogelnik, H.: Bell Syst. Tech. J. **44** (1965) 455.
- 65Li Li, T.: Bell Syst. Tech. J. **44** (1965) 917.
- 65Sie Siegman, A.E.: Proc. IEEE **53** (1965) 277.
- 65Str Streifer, W.: J. Opt. Soc. Am. **55** (1965) 10.
- 65Vak Vakhimov, N.G.: Radio Eng. Electron. Phys. (English Transl.) 10 (1965) 1439.
- 66Kah Kahn, K.W.: Appl. Opt. 5 (1966) 407.
- 66Kog1 Kogelnik, H., Li, T.: Appl. Opt. 5 (1966) 1550.
- 66Kog2 Kogelnik, H., Li, T.: Proc. IEEE **54** (1966) 1312.
- 66Mag Magnus, W., Oberhettinger, F., Soni, R.P.: Formulas and theories for special functions of mathematical physics, Berlin–Heidelberg–New York–London–Paris–Tokyo: Springer-Verlag, 1966.
- 67Bar Barone, S.R.: Appl. Opt. 6 (1967) 861.
- 67Sie Siegman, A.E., Arrathoon, R.: IEEE J. Quantum Electron. 3 (1967) 156.
- 68Ber Bergstein, L.: Appl. Opt. **7** (1968) 495.
- 68Str Streifer, W.: IEEE J. Quantum Electron. 4 (1968) 156.
- 69Ana Anan'ev, Y.A., Sventsitskaya, N.A., Sherstobitov, V.E.: Sov. Phys. JETP (English Transl.) 28 (1969) 69.
- 69Bau Baues, P.: Opto-electronics (London) 1 (1969) 37.
- 69Kru Krupke, W.F., Sooy, W.R.: IEEE J. Quantum Electron. 5 (1969) 575.
- 69San1 Sanderson, R.L., Streifer, W.: Appl. Opt. 8 (1969) 2129.
- 69San2 Sanderson, R.L., Streifer, W.: Appl. Opt. 8 (1969) 2241.
- 70Arn Arnaud, J.A.: Bell. Syst. Tech. J. **49** (1970) 2311.
- 70Col Collins, S.A.: J. Opt. Soc. Am. **60** (1970) 1168.
- 70Fos Foster, J.D., Osterink, L.M.: J. Appl. Phys. 41 (1970) 3656.
- 70Koe Koechner, W.: Appl. Opt. **9** (1970) 2548.
| 70Sie
70Zuc | Siegman, A.E., Miller, H.Y.: Appl. Opt. 9 (1970) 2729.
Zucker, H.: Bell Syst. Tech. J. 49 (1970) 2349. |
|---|--|
| 71Loc
71Sch
71Smi | Locke, E.V., Hella, R., Westra, L.: IEEE J. Quantum Electron. 7 (1971) 581.
Schäfer, F.P., Müller, H.: Opt. Commun. 2 (1971) 407.
Smith, P.W.: Appl. Phys. Lett. 19 (1971) 132. |
| 72Abr
72Ana
72Clo
72Fre
72Kog | Abrams, R.L.: IEEE J. Quantum Electron. 8 (1972) 838. Anan'ev, Y.A.: Sov. J. Quant. Electron. 1 (1972) 565. Clobes, A.R., Brienza, M.J.: Appl. Phys. Lett. 21 (1972) 265. Freiberg, R.J., Chenausky, P.P., Buczek, C.J.: IEEE J. Quantum Electron. 8 (1972) 372. Kogelnik, H.W., Ippen, E.P., Dienes, A., Shank, C.V.: IEEE J. Quantum Electron. 8 (1972) 373. |
| 73Deg
73Fre
73Hor
73Poz | Degnan, J.J., Hall, D.R.: IEEE J. Quantum Electron. 9 (1973) 901. Freiberg, R.J., Chenausky, P.P., Buczek, C.J.: Appl. Opt. 12 (1973) 1140. Horwitz, P.: J. Opt. Soc. Am. 63 (1973) 1528. Pozzo, P.D., Polloni, R., Svelto, O., Zaraga, F.: IEEE J. Quantum Electron. 9 (1973) 1061. |
| 73Wis | Wisner, G.R., Foster, M.C., Blaszuk, P.R.: Appl. Phys. Lett. 22 (1973) 14. |
| 74Fre | Freiberg, R.J., Chenausky, P.P., Buczek, C.J.: IEEE J. Quantum Electron. 10 (1974) |
| 74Mar
74Sie | Marcuse, D.: Theory of dielectric optical waveguides, in: Theory of dielectric optical waveguides, Pau, Y.H. (ed.), New York: Academic Press, 1974.
Siegman, A.E.: Appl. Opt. 13 (1974) 353. |
| 75Loe
75Szi | Lörtschner, J.P., Steffen, J., Herziger, G.: Opt. Quantum Electron. 7 (1975) 505.
Sziklas, E.A., Siegman, A.E.: Appl. Opt. 14 (1975) 1874. |
| 76Deg
76Ewa
76Laa
76Mar | Degnan, J.J.: Appl. Phys. 11 (1976) 1. Ewanitzky, T.F., Craig, J.M.: Appl. Opt. 15 (1976) 1465. Laakmann, K.D., Steier, W.H.: Appl. Opt. 15 (1976) 1334. Marowsky, G., Kaufman, K.: IEEE J. Quantum Electron. 12 (1976) 207. |
| 77Ber | Berger, N.K., Deryugin, N.A., Lukyanov, Y.N., Studenikin, Y.E.: Opt. Spectrosc. |
| 77Her
77Isa | Herbst, R.L., Komine, H., Byer, R.L.: Opt. Commun. 21 (1977) 5. Isaev, A.A., Kazaryan, M.A., Petrash, G.G., Rautian, S.G., Shalagin, A.M.: Sov. J. Quantum Electron. (English Transl.) 7 (1977) 746. |
| 78Pax
78Rig | Paxton, A.H., Salvi, T.C.: Opt. Commun. 26 (1978) 305.
Rigrod, W.: IEEE J. Quantum Electron. 14 (1978) 377. |
| 79Abr
79Iff
79Jar
79Nem | Abrams, R.W.: Waveguide gas lasers, in: Laser Handbook, Vol. 3, Amsterdam: North-Holland, 1979.
Iffländer, R., Kortz, H.P., Weber, H.: Opt. Commun. 29 (1979) 223.
Jarrett, S.M., Young, J.F.: Opt. Lett. 4 (1979) 176.
Nemoto, S., Makimoto, T.: J. Opt. Soc. Am. 69 (1979) 578. |

79Rea	Reading, S., Sze, R.C., Tallman, C.: Proc. Soc. Photo-Opt. Instrum. Eng. 190 (1979) 311.
79Sie	Siegman, A.E.: Opt. Commun. 31 (1979) 369.
80Car 80Eim 80Giu 80Hau 80Sch 80Tes	 Carter, W.H.: Appl. Opt. 19 (1980) 1027. Eimerl, D.: J. Appl. Phys. 51 (1980) 3008. Giuliani, G., Park, Y.K., Byer, R.L.: Opt. Lett. 5 (1980) 491. Hauck, R., Kortz, H.P., Weber, H.: Appl. Opt. 19 (1980) 598. Schindler, G.M.: IEEE J. Quantum Electron. 16 (1980) 546. Teschke, O., Teixeira, S.R.: Opt. Commun. 32 (1980) 287.
81Kor	Kortz, H.P., Iffländer, R., Weber, H.: Appl. Opt. 20 (1981) 4124.
82Sve	Svelto, O.: Principles of lasers, New York: Plenum Press, 1982.
83Oug	Oughstun, K.E., Slaymaker, P.A., Bush, K.A.: IEEE J. Quantum Electron. 19 (1983) 1558.
83Pic 83Rig	Piché, M., Lavigne, P., Martin, F., Belanger, P.A.: Appl. Opt. 22 (1983) 1999. Rigrod, W.W.: IEEE J. Quantum Electron. 19 (1983) 1679.
84Ger 84Her 84McC 84Moo 84Rip 84Skl	 Gerlach, R., Wei, D., Amer, N.M.: IEEE J. Quantum Electron. 20 (1984) 948. Herziger, G., Weber, H.: Appl. Opt. 23 (1984) 1450. McCarthy, N., Lavigne, P.: Appl. Opt. 23 (1984) 3845. Moore, G.T.: Proc. Soc. Photo-Opt. Instrum. Eng. 453 (1984) 255. Ripper, G., Herziger, G.: Feinwerktech. Messtech. 92 (1984) 301. Sklar, E.: J. Opt. Soc. Am. A 1 (1984) 537.
85Hil 85Iza 85Kan 85Lav 85McC 85McC	 Hill, C.A., Hall, D.R.: Appl. Opt. 24 (1985) 1283. Izawa, S., Suda, A., Obara, M.: J. Appl. Phys. 58 (1985) 3987. Kane, T.J., Byer, R.L.: Opt. Lett. 10 (1985) 65. Lavigne, P., McCarthy, N., Demers, J.G.: Appl. Opt. 24 (1985) 2581. McCarthy, N., Lavigne, P.: Opt. Lett. 10 (1985) 553. Taché, J.P.: Opt. Commun. 55 (1985) 419.
86DeS 86Har 86Mag 86Sie 86Skl 86Upp1 86Upp2 86Web	 DeSilvestri, S., Laporta, P., Magni, V.: Opt. Commun. 59 (1986) 43. Harter, D.J., Walling, J.C.: Opt. Lett. 11 (1986) 706. Magni, V.: Appl. Opt. 25 (1986) 2039. Siegman, A.E.: Lasers, Mill Valley: University Science Books, 1986. Sklar, E.: IEEE J. Quantum Electron. 22 (1986) 1088. Uppal, J.S., Monga, J.C., Bhawalkar, D.D.: Appl. Opt. 25 (1986) 97. Uppal, J.S., Monga, J.C., Bhawalkar, D.D.: Appl. Opt. 25 (1986) 389. Weber, H., Iffländer, R., Seiler, P.: Proc. Soc. Photo-Opt. Instrum. Eng. 650 (1986) 92.
87Apa	Apanasevich, P.A., Kvach, V.V., Koptev, V.G., Orlovich, V.A.: Sov. J. Quantum Electron. (English Transl.) 17 (1987) 160.
87Gro 87Hal	Gromov, A.N., Trashkeev, S.I.: Opt. Spectrosc. (USSR) 62 (1987) 369. Hall, D.R., Hill, C.A.: Radiofrequency-discharge-excited CO ₂ lasers, in: Handbook of molecular lasers. New York, Basel: Marcel Dekker, Inc. 1087, p. 165–108
87Kan 87Mag 87Met	Kane, T.J., Nilsson, A.C., Byer, R.L.: Opt. Lett. 12 (1987) 175. Magni, V.: J. Opt. Soc. Am. A 4 (1987) 1962. Metcalf, D., de Giovanni, P., Zachorowski, J., Leduc, M.: Appl. Opt. 26 (1987) 4508.

215

216	References for 8.1
87Oug	Oughstun, K.E.: Unstable resonator modes, in: Wolf, E. (ed.): Progress in optics XXIV, New York: Elsevier Science Publishers B.V., 1987, p. 165–387.
87Ru	Ru, P., Narducci, L.M., Tredicce, J.R., Bandy, D.K., Lugiato, L.A.: Opt. Commun. 63 (1987) 310.
87Sil 87Skl	Silichev, O.O.: Sov. J. Quantum Electron. (English Transl.) 17 (1987) 530. Sklar, E.: IEEE J. Quantum Electron. 23 (1987) 229.
88Bro 88Cal 88Coo 88DeS1 88DeS2	 Brorsen, S.D.: IEEE J. Quantum Electron. 23 (1988) 512. Calahan, K.R., Clayton, C.M., Paxton, A.H.: Appl. Opt. 27 (1988) 2694. Cooper, D., Tankersley, L.L., Reintjes, J.: Opt. Lett. 13 (1988) 568. DeSilvestri, S., Laporta, P., Magni, V., Svelto, O.: Opt. Commun. 67 (1988) 229. DeSilvestri, S., Laporta, P., Magni, V., Svelto, O., Majocchi, B.: Opt. Lett. 13 (1988)
88Hau1 88Hau2 88Hil 88Hod 88Lav 88Lit 88Mil	 201. Hauck, R., Hodgson, N., Weber, H.: J. Appl. Phys. 63 (1988) 628. Hauck, R., Hodgson, N., Weber, H.: J. Mod. Opt. 35 (1988) 165. Hill, C.A.: IEEE J. Quantum Electron. 24 (1988) 1936. Hodgson, N.: Proc. Soc. Photo-Opt. Instrum. Eng. 1021 (1988) 89. Lavigne, P., McCarthy, N., Parent, A., Snell, K.: Can. J. Phys. 66 (1988) 888. Litzenberger, L.N., Smith, M.J.: IEEE J. Quantum Electron. 24 (1988) 2270. Miloni, P.W., Eberly, J.H.: Lasers, New York, Chichester, Brisbane, Toronto, Singapore:
88Nil	John Wiley & Sons, 1988. Nilsson, A.C., Day, T., Farinas, A.D., Gustafson, E.K., Byer, R.L.: Narrow linewidth operation of diode-laser-pumped nonplanar ring oscillators, Topics in physics: Proceed- ings of the fourth symposium on frequency and metrology, Ancona, Italy, 1988, Berlin– Heidelberg–New York: Springer-Verlag, 1988.
88Ozy	Ozygus, B.: Losses and modestructures of optical resonators, Technical University Berlin, Germany: Internal Report, 1988.
88Sch 88Sne 88Tul	Schütz, I., Wiegand, S., Wallenstein, R.: Laser Optoelektron. 20 (1988) 39. Snell, K.H., McCarthy, N., Piché, M.: Opt. Commun. 65 (1988) 377. Tulip, J.: Carbon dioxide slab laser, United States Patent 4,719,639, January 1988.
89Day	Day, T., Nilsson, A.C., Fejer, M.M., Farinas, A.D., Gustafson, E.K., Nabors, C.D., Byer, R.L.: Electron. Lett. 25 (1989) 810.
89Eic 89Egg 89Emi	 Eicher, J., Hodgson, N., Weber, H.: J. Appl. Phys. 66 (1989) 4608. Eggleston, J.M., Frantz, L.M., Injeyan, H.: IEEE J. Quantum Electron. 25 (1989) 1855. Emiliani, G., Piegori, A., DeSilvestri, S., Laporta, P., Magni, V.: Appl. Opt. 28 (1989) 2832.
89Jac1 89Jac2 89Kin	Jackson, P.E., Baker, H.J., Hall, D.R.: Appl. Phys. Lett. 54 (1989) 1950. Jackson, P.E., Hall, D.R., Hill, C.A.: Appl. Opt. 28 (1989) 935. King, T.A.: Ring laser resonators, in: Hall, D.R., Jackson, P.E. (eds.): Physics and technology of laser resonators. London: Adam Hilger, 1989, p. 62–79.
89Nil 89Par 89Ruf	Nilsson, A.C., Gustafson, E.K., Byer, R.L.: IEEE J. Quantum Electron. 25 (1989) 767. Parent, A., Lavigne, P.: Opt. Lett. 14 (1989) 399. Ruff, J.A., Siegman, A.E.: Technical Digest CLEO 89, 296. Washington, D.C.: Optical Society of America, 1989, p. 296.
90Alv 90DeS	Alvarez-Chust, R., Jimenez-Lidon, J.J.: IEEE Photon. Technol. Lett. 2 (1990) 24. DeSilvestri, S., Magni, V., Svelto, O., Valentini, G.L.: IEEE J. Quantum Electron. 26 (1990) 1500.
90Hil 90Hod1	Hill, C.A., Jackson, P.E., Hall, D.R.: Appl. Opt. 29 (1990) 2240. Hodgson, N., Haase, T., Weber, H.: Proc. Soc. Photo-Opt. Instrum. Eng. 1277 (1990) 70.

90Hod2 90Hod3 90Nis 90Now	Hodgson, N., Weber, H.: IEEE J. Quantum Electron. 26 (1990) 731. Hodgson, N., Weber, H.: Opt. Quantum Electron. 22 (1990) 39. Nishima, J., Yoshizawa, K.: Proc. Soc. Photo-Opt. Instrum. Eng. 1226 (1990) 340. Nowack, R., Opower, H., Schaefer, U., Wessel, K., Hall, T., Krüger, H., Weber, H.: Proc. Soc. Photo-Opt. Instrum. Eng. 1276 (1990) 18.
90Pie	Piegari, A., Tirabassi, A., Emiliani, G.: Proc. SPIE (Int. Soc. Opt. Eng.) 1125 (1990)
90Yas	Yasui, K., Yagi, S., Tanaka, M.: Appl. Opt. 29 (1990) 1277.
91Che 91Mag	Cheng, E.A.P., Kane, T.J.: Opt. Lett. 16 (1991) 478. Magni, V., DeSilvestri, S., Svelto, O., Valentini, G.L.: Hundred-watts diffraction lim- ited Nd: YAG laser with super-Gaussian resonator, Technical Digest CLEO 91, Optical Society of America, Washington, D.C., 1991.
91Now	Nowack, R., Opower, H., Wessel, K., Krüger, H., Haas, W., Wenzel, N.: Laser und Optoelektronik 3 (1991) 68.
92Bow	Bowers, M.S.: Opt. Lett. 17 (1992) 1319.
92Col 92Dal	Colley, A.D., Baker, H.J., Hall, D.R.: Appl. Phys. Lett. 61 (1992) 136. Dallarosa, J., Gardner, P.: RF excited CO ₂ slab waveguide laser, United States Patent 5.131.004. July 1992.
92Day 92Hob	Day, T., Gustafson, E.K., Byer, R.L.: IEEE J. Quantum Electron. 28 (1992) 1106. Hobart, J.L., Yarborough, J.M., Dallarosa, J., Gardner, P.: RF excited CO ₂ slab wave- guide laser, United States Patent 5,123,028, June 1992.
92Hod1	Hodgson, N., Haase, T.: Opt. Quantum Electron. 24 (1992) 903.
92Hod2	Hodgson, N., Weber, H.: J. Mod. Opt. 39 (1992) 1873.
92Jen 92Kho	Jenkins, R.M., Devereux, R.J.: Appl. Opt. 31 (1992) 5080. Khelkal M. Herlemont $\mathbf{F} \cdot \mathbf{A}$ ppl. Opt. 31 (1992) 4175
92Mag	Magni, V., DeSilvestri, S., Qian, LJ., Svelto, O.: Opt. Commun. 94 (1992) 87.
92Mef1	Mefferd, W.S.: RF excited CO_2 slab waveguide laser, United States Patent 5,131,003, July 1992.
92Mef2	Mefferd, W.S.: RF excited CO ₂ slab waveguide laser, United States Patent 5,155,739, October 1992.
92Oug	Oughstun, K.E. (ed.): Selected papers on scalar wave diffraction, SPIE Milestone Series Vol. MS51, Bellingham, WA: The Society of Photo-Optical Instrumentation Engineers 1992.
92Ren	Reng, N., Eppich, B.: Opt. Quantum Electron. 24 (1992) 973.
92Web 92Yar	Weber, H.: Opt. Quantum Electron. 24 (1992) 1027. Yarborough, J.M., Hobart, J.L., Dallarosa, J.: RF excited CO_2 slab waveguide laser, United States Patent 5,140,606, August 1992.
93Doc	Document ISO/TC 172/SC 9/WG 1 N 56, ISO/CD 11 146, Optics and optical instru- ments – Lasers and laser related equipment – Test methods for laser beam parameters: Beam width divergence angle and beam propagation factor. Nov. 26, 1993
93Dup	Duplain, G., Verlyu, P.G., Dobrowolski, J.A., Waldorf, A., Buissiere, S.: Appl. Opt. 32 (1993) 1145
93Ehr	Ehrlichmann, D., Habich, U., Plum, HD.: IEEE J. Quantum Electron. 29 (1993) 2211.
93Hod1	Hodgson, N., Bostanjoglo, G., Weber, H.: Appl. Opt. 32 (1993) 5902.
93Hod2	Hodgson, N., Bostanjoglo, G., Weber, H.: Opt. Commun. 99 (1993) 75.
93Hod3	Hodgson, N., Dong, S., Lü, Q.: Opt. Lett. 18 (1993) 1727.
93Hod4 93Pie	Hodgson, N., Weber, H.: IEEE J. Quantum Electron. 29 (1993) 2497. Piegari, A., Emiliani, G.: Appl. Opt. 32 (1993) 5454.

218	References for 8.1
94Bos1	Bostanjoglo, G., Bernhardt, A.: Proc. Soc. Photo-Opt. Instrum. Eng. 2253 (1994) 791.
94Bos2	Bostanjoglo, G., Hodgson, N.: Proc. Soc. Photo-Opt. Instrum. Eng. 2206 (1994) 459.
94Bos3	Bostanjoglo, G., Hodgson, N., Weber, H.: Pure Appl. Opt. 3 (1994) 497.
94Bur	Burnham, R., Witt, G., DiBiase, D., Koechner, W.: Proc. Soc. Photo-Opt. Instrum. Eng. 2206 (1994) 489.
94Pet	Petersen, C.L., Eisel, D., Brzezinski, J.J., Gross, H.: Proc. Soc. Photo-Opt. Instrum. Eng. 2206 (1994) 91.
95Paq	Paquin, R.A.: Properties of Metals; in: Bass, M., Stryland, E.W. Van, Williams, D.R., Wolfe, W.L. (eds.): Handbook of Optics, Vol. II, Washington: Optical Society of America, 1995, Chap. 35.
95Cha	Chandra, S., Allik, T.H., Hutchinson, J.A.: Opt. Lett. 20 (1995) 2387.
95Eis	Eisel, D., Gross, H., Petersen, C.L.: Slab-waveguide CO_2 laser, United States Patent 5,412,681, May 1995.
96Lap	Lapucci, A., Labate, A., Rossetti, F., Mascalchi, S.: Appl. Opt. 35 (1996) 3185.
96Mar	Martin, K.I., Clarkson, W.A., Hanna, D.C.: Opt. Lett. 21 (1996) 875.
97Hod	Hodgson, N., Weber, H.: Optical resonators, Berlin–Heidelberg–London–New York: Springer-Verlag, 1997.

Landolt-Börnstein New Series VIII/1A2

218

9.1 Interferometry

H.J. TIZIANI, N. KERWIEN, G. PEDRINI

9.1.1 Introduction

Interferometry is a powerful tool for many applications in optics, spectroscopy, and optical metrology. Interference phenomena can be observed in daily live, for example when colors in a soap bubble or in oil slicks on a wet road appear. Colored fringes, the so-called "Newton rings", are seen in a thin air film enclosed between two glass plates brought in contact. Interference results from the superposition of two or more at least partially coherent electromagnetic waves. The characteristic feature of the interference phenomenon is a spatially modulated intensity distribution in the region of superposition, that differs in magnitude from the mean intensity values of the superimposing beams.

The development of optical interferometry is closely linked with the history of wave optics. In 1690 Huygens put forward the basis for the theory we use today. In 1801, Thomas Young performed a fundamental experiment for demonstrating interference and the wave nature of light, see Fig. 9.1.26. He used monochromatic light passing a single pinhole first and subsequently traveling to two separated slits or pinholes. The light diffracted from these pinholes illuminates a screen at a distance much larger than the pinhole separation. Then interference fringes can be seen in the overlapping area. The simple demonstration is the well-known Young experiment. In 1816 Fresnel and Arago perfected these investigations. Since two orthogonal polarized light waves could not interfere Young concluded out of these experiments that light waves are transverse waves.

Applications of interferometry in optical metrology followed in succession. In 1896 Michelson carried out the first interferometric measurements to determine the length of the Pt-Ir bar which was the international prototype of the meter and defined its length in terms of the wavelength of the red cadmium radiation. In 1960 followed the redefinition of the meter with the wavelength of the orange radiation of 86 Kr.

Interferometry plays also an important role in optical testing and high-precision metrology. Twyman used 1916 a modified Michelson-interferometer for testing optical components. This interferometer was adapted by Linnik 1933 for the examination of microscopic structures of reflecting surfaces. Furthermore, interferometry has important applications in spectroscopy as well as for studies of fluid flow and combustion, but these will not be discussed here.

Different developments led to new applications of interferometry. To these belongs the progress in the development of lasers as coherent light source as well as the automated computer-assisted fringe analysis together with the development of high-resolution detector arrays based on CCD (Charge Coupled Device) or CMOS (Complementary Metal Oxide Semiconductor) technology.

For micro- and nanotechnology and especially for structures with lateral dimensions near or smaller than the wavelength of the light used, coherence properties and polarization need to be considered, as will be discussed, see Sect. 9.1.2.4 and Sect. 9.1.4.4.

Speckle pattern interferometry (see Sect. 9.1.7) as well as holographic interferometry (see Sect. 9.1.8) plays an important role in non-destructive testing in our days. The test object is, however, not a highly polished surface but rather an optically rough surface.

9.1.2 Basic principles of interference

Before we start with a detailed discussion of different interferometric methods in the following sections, we first give a short summary of the basic principles of interferometry. Starting point in Sect. 9.1.2.1 is the very short discussion of coherence as a necessary prerequisite for the occurrence of interference phenomena, see also Part 6 in this volume. In Sect. 9.1.2.2 we describe the mathematical basics of two-beam interference, in Sect. 9.1.2.3 the interference in a plane-parallel plate, and in Sect. 9.1.2.4 we account for the vectorial character of light interference.

9.1.2.1 Coherence

Two overlaid beams of light produce only an interference pattern, if their relative phases are stable long enough to be sensed by a detecting system (i.e. the eye). In this case we say the two beams are coherent. The coherence of light will be treated extensively in Born and Wolf [87Bor] as well as in Part 6 of this volume, therefore a brief introduction will be given here only.

The first quantitative concepts of coherence were formulated by Laue 1907. The basis for modern coherence theory was derived by van Cittert [34Van, 39Van] and Zernike [38Zer]. The van Cittert–Zernike theorem is well-known in spatial coherence descriptions. The theory was further developed by Hopkins [51Hop, 53Hop, 55Hop] and Wolf [54Wol, 55Wol] including higher-order coherence theory described by Mandel and Wolf [65Man]. For simplicity and to point out the main aspects we will discuss in the following only two special cases of the general term of coherence: the pure temporal as well as the pure spatial coherence.

9.1.2.1.1 Temporal coherence

The Fourier Transform (FT) \mathcal{F} of the spectral energy distribution of the light source leads to the coherence time $T_{\rm coh}$. Multiplication with the velocity of the light c gives the coherence length $L = c T_{\rm coh}$. With $T_{\rm coh} = 1/\Delta \nu$ and $\nu = c/\lambda$ we get:

$$L \approx \frac{\lambda^2}{|\Delta\lambda|} \,. \tag{9.1.1}$$

These relationships are illustrated in Fig. 9.1.1.



Fig. 9.1.1. Relationship between spectral energy density $E(\nu)$ of a light spectrum and its temporal behavior of the corresponding light train with coherence time $T_{\rm coh}$.

9.1.2.1.2 Spatial coherence

It can be shown that the complex degree of coherence is equivalent to the calculation of the complex amplitude in a diffraction pattern. The well-known van Cittert–Zernike theorem was developed by van Cittert [34Van, 39Van] and presented in a simpler way by Zernike. For simplicity let us consider the source to be replaced by a screen with randomly distributed grains with random phase illuminated by a spherical wave. In the focusing plane the complex degree of spatial coherence is the Fourier transform of the source area [51Hop, 87Bor].

9.1.2.2 Two-beam interference

The light wave is propagating as an electromagnetic wave in vacuum. For simplicity we assume at first a scalar plane wave propagating in vacuum in z-direction. In Sect. 9.1.2.4 polarization effects will be discussed. The electric field E at any point can be represented by a sinusoidal function of distance z and time t:

$$E = a \cos\left[2\pi\nu\left(t - \frac{z}{c}\right)\right] , \qquad (9.1.2)$$

where a is the amplitude, ν the frequency, and c the speed of propagation of the wave in vacuum. The period of vibration T is

$$T = \frac{1}{\nu} = \frac{2\pi}{\omega} \; ,$$

where ω is the circular frequency. The wavelength λ is given by the relation

$$\lambda = c T = \frac{c}{\nu} \; .$$

k is the propagation constant of the wave,

$$k = \frac{2\pi}{\lambda} \; .$$

The speed of the light propagation in the medium with the refractive index n is v = c/n. The wavelength in the medium is now

$$\lambda_n = \frac{\lambda}{n} \; .$$

It is convenient to use the complex representation of light waves. Equation (9.1.2) can then be written as

$$E = \operatorname{Re}\left\{a \exp\left[i 2\pi\nu \left(t - \frac{zn}{c}\right)\right]\right\} , \qquad (9.1.3)$$

where Re {} represents the real part of the expression within the braces and $i = \sqrt{-1}$. We can write (9.1.3) as a product of the time- and space-varying functions

$$E = \operatorname{Re}\left\{a \exp\left(-\mathrm{i}\,\phi\right) \exp\left(\mathrm{i}\,2\,\pi\,\nu\,t\right)\right\} = A\,\exp\left(2\,\mathrm{i}\,\pi\,\nu\,t\right)\,,\tag{9.1.4}$$

where $\phi = 2\pi n z / \lambda$ is the phase and nz is the optical path difference from the origin of z.

It should be noticed that the frequency of the visible light is in the order of 6×10^{14} Hz. Therefore a direct observation of the electric field is not possible. The only measurable quantity is the time average of the amount of energy, which in a unit of time crosses a unit area normal to the

direction of the energy flow. This is proportional to the time average of the square of the electric field:

$$\langle E^2 \rangle = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} E^2 \, \mathrm{d}t = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} a^2 \, \cos^2(\omega t - \phi) \, \mathrm{d}t = \frac{a^2}{2} \,. \tag{9.1.5}$$

We are, however, not interested in the absolute intensity, but only in the relative values over the specified region. We can ignore the factor $\frac{1}{2}$, as well as any other factors of proportionality, and define the optical intensity as

$$I = a^2 = |A|^2 . (9.1.6)$$

Let us now consider two monochromatic waves propagating in the same direction superposed in a point P. The total electric field at this point is

$$E = E_1 + E_2 , (9.1.7)$$

where E_1 and E_2 are the electric fields due to the two waves. Assuming that the two waves have the same frequency, the intensity at this point is

$$I = |A_1 + A_2|^2 {.} {(9.1.8)}$$

Considering $A_1 = a_1 \exp(-i\phi_1)$ and $A_2 = a_2 \exp(-i\phi_2)$ to be the complex amplitudes of the two waves of the same wavelength, we obtain:

$$I = A_1^2 + A_2^2 + A_1 A_2^* + A_1^* A_2 = I_1 + I_2 + 2\gamma_{12}\sqrt{I_1 I_2} \cos(\phi_1 - \phi_2) , \qquad (9.1.9)$$

where I_1 and I_2 are the intensities at P for the two waves acting separately (incoherent superposition) and γ_{12} is the coherence factor. Equation (9.1.9) is often called the general interference law for stationary optical fields [87Bor]. For incoherent light $\gamma_{12} = 0$ and for perfectly coherent light $\gamma_{12} = 1$. For partial coherence γ_{12} is between 0 and 1. We consider at first $\gamma_{12} = 1$, i.e. perfect coherence.

The intensity has its maximum I_{max} when $\phi_1 - \phi_2 = 2 m \pi$, where *m* is an integer. The minima occur for $\phi_1 - \phi_2 = (2m + 1)\pi$. The intensity can be written according to (9.1.9) as

$$I = I_0 \{ 1 + \gamma \cos(\phi_1 - \phi_2) \} , \qquad (9.1.10)$$

where $I_0 = I_1 + I_2$ and the contrast γ or visibility (sometimes called modulation m) is conveniently defined according to Michelson

$$\gamma = \frac{(I_{\max} - I_{\min})}{(I_{\max} + I_{\min})} = \frac{2\sqrt{I_1 I_2}}{I_1 + I_2} .$$
(9.1.11)

Maximum visibility is obtained for $I_1 = I_2$.

9.1.2.3 Interference in a plane-parallel plate

Up to now we have treated interference from a quite abstract point of view. In the following we will give a simple example where interference occurs in a plane-parallel plate. The effects demonstrated here will also be the basis for more sophisticated methods described especially in Sect. 9.1.3. The principal arrangement and the notations used are depicted in Fig. 9.1.2. The refractive index of the glass plate is n_1 and the surrounding medium is considered to be air (n = 1.0). The angle of incidence is ϑ and the angle after refraction is ϑ'_1 . The optical path



Fig. 9.1.2. Optical path difference between the front and back surface of a plane-parallel plate.

Fig. 9.1.3. Plane-parallel plate illuminated with (\mathbf{a}) a small light source (parallel incident light) with the corresponding fringe pattern and (\mathbf{b}) a large light source together with the appropriate fringe pattern at infinity or in the focus plane of a lens. In order to see fringes a small wedge angle was introduced in the plate for (\mathbf{a}) .

difference between the waves reflected from the front and back surface for a plane-parallel plate is by neglecting a phase jump at the interface

$$\Delta = [ABC] - [AD] = 2n_1 d \cos(\vartheta'_1) , \qquad (9.1.12)$$

where [ABC] means the optical path corresponding to the geometrical path following the points A, B, and C multiplied with the refractive index n_1 . It should be noted that there is a phase jump of π at the interface, which can be derived from the Fresnel equations at the boundary. The optical path difference is therefore

$$\Delta = 2n_1 d \cos\left(\vartheta_1'\right) \pm \frac{\lambda}{2} \,. \tag{9.1.13}$$

In Fig. 9.1.3 fringe patterns for a small and a large monochromatic light source, respectively, are seen. The source is placed in the focal plane of a collimating lens. An ideal point source therefore generates a parallel beam, whereas an extended light source results in a divergence of the light beam. The fringes in Fig. 9.1.3a are called Fizeau fringes (fringes of equal thickness) and in Fig. 9.1.3b Haidinger fringes (fringes of equal inclination). It should be noted that in Fig. 9.1.3a there is a small wedge angle introduced into the plate in order to see parallel and equidistant fringes (tilt fringes). For a perfect plane-parallel plate no Fizeau fringes would be observed (uniform field). The existence of fringes of equal inclination depends on the direction of the incident light being reflected by the two parallel surfaces. The Optical Path Difference (OPD) depends on the angle of incidence of the waves. The fringe patterns are therefore different, for the same plane-parallel plate, depending on whether a point or extended source is used. By contrast for the same incidence direction the OPD is the same.

9.1.2.4 Vector effects of interference

Up to now we have treated light as a scalar quantity and therefore didn't account for the real nature of light, especially its vectorial character. Initially assuming that light performs longitudinal vibrations, Fresnel and Arago did systematic investigations on this topic in 1816. Their set-up resembled Young's earlier experiments despite the fact, that different configurations of polarizers in front of the light source, the slits, and the screen are involved in the experiments. Their work was the basis of understanding the transverse nature of light and led to empirical statements, called the interference laws of Fresnel and Arago [1819Ara]. They can be formulated in a slightly modernized form [92Bro, 93Col] as (see also Fig. 9.1.4):

- 1. Two waves linearly polarized in the same plane can interfere (Fig. 9.1.4a).
- 2. Two waves linearly polarized with perpendicular polarizations cannot interfere (Fig. 9.1.4b).
- 3. Two waves linearly polarized with perpendicular polarizations, if derived from perpendicular components of unpolarized light and subsequently brought into the same plane of polarization, cannot interfere (Fig. 9.1.4c).
- 4. Two waves linearly polarized with perpendicular polarizations, if derived from the same linearly polarized wave and subsequently brought into the same plane of polarization, can interfere (Fig. 9.1.4d).

Stating these laws as experimental facts, Fresnel and Arago didn't provide the theoretical and mathematical basis. This step was preserved to Stokes, who picked up the investigations of Fresnel and Arago and formulated his famous paper in 1852 [1852Sto]. A deeper discussion would lead to the partial coherence theory of polarized light, which is not the scope of this text. We will just restate some fundamental results and definitions here, which are important for some experiments to follow. For further reading we refer to excellent literature that covers this subject [62Shu, 92Bro, 93Col]. Following the same considerations as in Sect. 9.1.2.2 for two quasi-monochromatic plane waves in







Fig. 9.1.4. Interference laws of Fresnel and Arago.



Fig. 9.1.5. Vector effects of interference. The contrast of the interference fringes depends on the mutual propagation direction and the polarization of the interfering plane waves.

vector formulation

$$\boldsymbol{E}_{1/2}(t) = \boldsymbol{a}_{1/2}(t) \exp\left\{ i \left(\boldsymbol{k}_{1/2} \, \boldsymbol{r} - 2 \, \pi \bar{\nu} \, t + \phi_{1/2}(t) \right) \right\} \tag{9.1.14}$$

we get a quite similar interference equation as in (9.1.9) nevertheless revealing the characteristic polarization properties of interference:

$$I = \langle \boldsymbol{A}_{1}^{2} \rangle + \langle \boldsymbol{A}_{2}^{2} \rangle + 2 \operatorname{Re} \left(\langle \boldsymbol{A}_{1} \boldsymbol{A}_{2}^{*} \rangle \exp \left\{ i \left(\boldsymbol{k}_{1} - \boldsymbol{k}_{2} \right) \boldsymbol{r} \right\} \right) , \qquad (9.1.15)$$

where $\overline{\nu}$ denotes the mean frequency, $A_{1/2} = a_{1/2} \exp(-i\phi_{1/2})$ the in general complex vector amplitudes, $\phi_{1/2}$ the phases, and $k_{1/2}$ the wave vectors of the corresponding plane waves, respectively. As in Sect. 9.1.2.2 Re stands for the real part and $\langle \rangle$ for the time-averaging operation defined in (9.1.5).

The chief difference to (9.1.9) consists in the replacement of the conventional product by the vectorial inner product both in the amplitude and the phase part of the interference term. This fact has quite important consequences, which lead to interference effects not explainable by scalar theory. They are subject of the following discussion.

We start with the two special cases depicted in Fig. 9.1.5. The plane spanned by the two wave vectors serves as reference plane. If the waves are both linearly polarized perpendicular to the reference plane (often referred to as TE-polarization) the field vectors are always parallel to each other regardless of the mutual angle of incidence θ and therefore cause the same interference pattern.

The situation is quite different for polarization vectors parallel to the reference plane. In this case of so-called TM-polarization the interference term depends on the aperture angle and is proportional to $\cos(\theta)$ as a consequence of the inner vector product. This means that starting with parallel propagating waves the contrast of the interference pattern decreases with increasing aperture angle. It even vanishes for $\theta = 90^{\circ}$, i.e. perpendicular mutual incidence, before increasing again up to plane waves propagating in opposite direction.

As learned from the discussion above, waves with orthogonal polarization states don't interfere. This is also true for two orthogonal¹ polarized plane waves, which propagate in the same direction. Nevertheless, the question arises if there exists a relationship between the two vector components of the resulting wave, that can be described in a corresponding formalism to scalar coherence theory and that causes, under certain circumstances, similar interference effects.

To discuss this situation in more detail the propagation direction of the resulting plane wave is chosen along the z-axis, whereas the two partial waves are linearly polarized in x- und y-directions, respectively:

$$E_{x/y}(t) = a_{x/y}(t) \exp\left\{ i \left(\phi_{x/y}(t) - 2\pi \overline{\nu} t \right) \right\} .$$
(9.1.16)

In order to force the two orthogonal vector components to interfere, we have to project them onto the same plane of vibration (i.e. the same plane of polarization). Experimentally this is done by a polarizer. Let the polarizer be aligned in direction α to the x-axis, then we get:

¹ Up to now we have used the word "orthogonal" in a geometrical sense. Entering the scope of polarimetry the meaning is broadened to its mathematical definition common in theory of linear algebra.

 $E(t,\alpha) = E_x(t)\cos(\alpha) + E_y(t)\sin(\alpha) . \qquad (9.1.17)$

This results in the intensity:

$$I(\alpha) = \langle E(t,\alpha) E^*(t,\alpha) \rangle$$

= $J_{xx} \cos^2(\alpha) + J_{yy} \cos^2(\alpha) + 2\sqrt{J_{xx}} \sqrt{J_{yy}} \cos(\alpha) \sin(\alpha) \operatorname{Re}(j_{xy}) ,$ (9.1.18)

where we fall back on the definition of the coherency matrix:

$$\mathsf{J} = \begin{bmatrix} \langle E_x E_x^* \rangle & \langle E_x E_y^* \rangle \\ \langle E_y E_x^* \rangle & \langle E_y E_y^* \rangle \end{bmatrix} = \begin{bmatrix} \langle a_x^2 \rangle & \langle a_x a_y e^{\mathrm{i}(\phi_x - \phi_y)} \rangle \\ \langle a_x a_y e^{-\mathrm{i}(\phi_x - \phi_y)} \rangle & \langle a_y^2 \rangle \end{bmatrix}$$
(9.1.19)

and the complex correlation factor

$$j_{xy} = \frac{J_{xy}}{\sqrt{J_{xx}}\sqrt{J_{yy}}} = |j_{xy}| e^{iA\phi} .$$
(9.1.20)

While the absolute value of j_{xy} describes the degree of coherence of the two polarization components the phase of j_{xy} represents their effective phase difference. The introduced definitions and the structure of (9.1.15) resemble a natural generalization of the scalar interference law for stationary optical fields (9.1.9) to its vectorial form.

We close this section by defining a further essential quantity, the degree of polarization [98Mov, 99Piq, 00Mov]:

$$P = \frac{I_{\text{polarized}}}{I_{\text{total}}} = \sqrt{1 - \frac{4 \det(\mathsf{J})}{(J_{xx} + J_{yy})^2}} \,. \tag{9.1.21}$$

In contrast to the coherency matrix and the correlation factor, the degree of polarization is independent of the choice of the coordinate system and therefore provides a fundamental measure of the partially coherent light field.

Now we have arrived at the basis of vectorial coherence theory. In Sect. 9.1.4.4 we will give some measurement examples that demonstrate the implementation of this formalism for metrological purposes. It is strongly related to the neighboring fields of polarimetry and ellipsometry.

9.1.3 Interferometry for optical testing

There are two general methods to produce mutually coherent waves for interference, namely wavefront division and amplitude division. For wavefront division different points on a wavefront are sampled to produce two new wavefronts. For amplitude division the wavefront is divided into two beams by some sort of beam splitter. Amplitude division is mainly used in optical testing. Interferometry is frequently used for length measurement in optical testing [66Fra]. In the following Fizeau fringes are considered.

9.1.3.1 Basic interferometer types

9.1.3.1.1 Michelson interferometer

Interferometry can be used to study phase variation across an optical wavefront. A frequently used interference arrangement is based on the Michelson principle, see Fig. 9.1.6. The beam splitter BS



Fig. 9.1.6. Michelson arrangement.

separates reference and object waves and recombines them after reflection at the mirrors M_1 , M_2 . A compensating plate K of identical glass and the same thickness as the beam splitter compensates for the dispersion when a broad-band source is used. The compensating plate is not needed for optical testing with monochromatic light [21Mic, 66Fra, 72Wya, 87Bor, 90Har, 92Mal, 98Mov, 99Piq, 00Mov, 03Har].

9.1.3.1.2 Twyman–Green interferometer

The Twyman–Green arrangement is frequently used for optical testing with monochromatic light. A laser serves mainly as light source. Due to the small bandwidth no compensating plate is needed. The interferometer is appropriate for testing optical components and systems as shown schematically in Fig. 9.1.7 for testing a spherical surface (S) or alternatively the objective (Obj.). The laser beam is expanded by the lenses L_1 and L_2 . The reference is reflected by the reference mirror R and the beam splitter BS. The object beam is reflected by BS and the spherical mirror S with its center of curvature in the focal point of the objective under test. The test object is passed twice in turn.

For testing the prism quality in transmission alternatively a prism mirror system can be put into the parallel test beam. The fringe pattern obtained is stored on a CCD for instance and analyzed by fringe analyzing procedures, which are described in Sect. 9.1.3.2. The most frequent first-order aberration types observed in optical systems are also shown schematically in Fig. 9.1.7, namely (a) for a perfect wavefront, no fringes can be seen, (b) shows a wavefront with defocus, (c) with astigmatisms, (d) with tilt, (e) with spherical aberration and defocus, (f) with coma or centering error.



Fig. 9.1.7. Twyman–Green interferometer for testing optical systems with typical fringe patterns.

9.1.3.1.3 Fizeau interferometer

In the Fizeau interferometer one optical reference surface is compared with a test surface by placing them in close proximity and is therefore nearly a common path interferometer and very robust with respect to turbulence and vibration. Fringes of equal thickness between the reference and test surface are observed and analyzed. The Fizeau and the Twyman–Green interferometer are standard interferometers for optical testing. In Fig. 9.1.8 a Fizeau interferometer is shown.



Fig. 9.1.8. Fizeau interferometer set-up (a) for testing a flat surface and (b) for testing a spherical surface.

A collimated plane wave is partly reflected back by the reference and the object surface, respectively. The optical path from the reference surface to the detector is nearly the same as for the corresponding object wave (after the beam splitter). For testing spherical surfaces the measuring wave is focused to the center of curvature of the spherical surface under test. The reference wave can either be reflected by a plane reference surface or alternatively by the last surface of the collimating lens system, an aplanatic surface, see Fig. 9.1.8b.

9.1.3.1.4 Mach–Zehnder interferometer

In the Mach–Zehnder interferometer reference and object waves are separated. The illuminating beam is divided by BS₁ into an object and reference beam and recombined again by BS₂ as shown in Fig. 9.1.9, where a configuration for an interference microscope system is shown. BS₁ and BS₂ are beam splitters, M_1 and M_2 are mirrors, and R is a compensating plate to compensate for the optical path of the substrate plate in the object path, O_1 and O'_1 are identical microscope objectives. The Mach–Zehnder interferometer can be very useful to measure the refractive index



Fig. 9.1.9. Mach–Zehnder interference arrangement.

of air for instance, where the path in air is compared with the same geometrical path in a vacuum cell.

9.1.3.1.5 Shearing interferometry

Shearing interferometry has been used as a tool for many decades. Lasers, optical components, and elements can be tested quickly. In shearing interferometry, both wavefronts are derived from the system under test. The interference pattern occurs by shearing one wavefront with respect to the other; no reference is needed. It is a very robust technique with respect to mechanical disturbances like vibrations, since both interfering waves follow nearly the same optical path [64Mur, 92Mal, 03Har]. Radial shear is frequently used for testing rotationally symmetric wavefronts. Lateral shearing interferometers have found the most applications in optical testing. In lateral shearing diffractive gratings can be very attractive for separating and combining the beams [96Lei, 97Sch, 03Har]. Rotational and axial shearing can be used for particular applications.

9.1.3.1.6 Fabry–Perot interferometer

We have assumed in the discussion so far that only two beams interfere. There occur, however, situations where multiple beams are involved such as in a diffraction grating or plane-parallel plate. Considering the division of amplitudes in a plane-parallel plate, where the surfaces are coated to give higher reflectivity, the intensity for a specific wavelength decreases for multiple reflections. This effect used by a Fabry–Perot interferometer is an important example of a system which makes use of multiple beam interference. The interferometer can serve as a high-resolution spectrometer and as an optical resonator which is an essential component of a laser. The optical resonator consists of two highly reflective parallel surfaces separated by a distance d. The laser cavity is a Fabry–Perot consisting of the two separated plates or a single plane-parallel glass plate with highly reflecting coatings.

The performance of a Fabry–Perot cavity is described by the finesse. It is a measure of the number of resolvable spectral lines. It is defined as the ratio of the phase difference between adjacent fringes to the Full Width-Half Maximum (FWHM) of a single fringe. The finesse, F, depends strongly on the reflectivity R and can be written [87Bor]

$$F = \frac{\pi\sqrt{R}}{1-R} \,. \tag{9.1.22}$$

The finesse of a cavity with flat mirrors is limited by the flatness and parallelism of the mirrors and is about 30. For etalons consisting of two curved mirrors it can be up to 10000.

9.1.3.2 Quantitative electronic phase evaluation techniques

Increasing use of electronic phase measuring techniques for direct measurements of phase differences in interferometry with very high accuracy has led to a high acceptance of interferometry in industry. Traditional methods of recording and analyzing interference patterns have largely been replaced by electronic techniques. Photodetector arrays based on CCD (Charge Coupled Device) or CMOS (Complementary Metal Oxide Semiconductor) technology together with computer analysis lead to a fast interpretation of interference patterns [93Rob].

There are many techniques using computer-aided fringe analysis, only the most used are discussed here, namely: - Fourier-transform technique,

- phase-shifting technique.

In Fig. 9.1.7 some interference patters of lens systems in a Twyman–Green set-up are shown. Fringe interpretation is required for a quantitative analysis.

For the Fourier-transform technique the fringes should not be closed. Sometimes they can be opened by introducing an additional tilt to be subtracted afterwards. For closed fringes the phaseshifting technique is most appropriate.

9.1.3.2.1 The Fourier-transform technique

For open fringes or fringe patterns that can be opened by the introduction of a tilt in order to generate an optical carrier the Fourier-transform technique as shown by Takeda [82Tak] is appropriate. No additional interaction with the interferometer is necessary. This spatial carrier is modulated by the local variation of the phase difference, $\phi_1 - \phi_2 = \Delta \phi$, between the two beams. The intensity of the interference pattern can be written, according to (9.1.9):

$$I(x,y) = a(x,y) + b(x,y)\cos(2\pi\beta_0 x + \Delta\phi(x,y))$$
(9.1.23)

which can be expressed as

$$I(x,y) = a(x,y) + c(x,y) \exp(i 2\pi\beta_0 \kappa) + c^*(x,y) \exp(-i 2\pi\beta_0 \kappa) , \qquad (9.1.24)$$

where $c(x,y) = \frac{1}{2} b(x,y) \exp(i\Delta\phi)$. The Fourier transform of the intensity with respect to x is

$$I(\beta, y) = A(\beta, y) + C(\beta - \beta_0, y) + C^*(\beta + \beta_0, y) .$$
(9.1.25)

 β is the spatial frequency in the x-direction and β_0 the carrier frequency which should be high in order to be able to separate the first from the zeroth order. One first order will be isolated and shifted back by β_0 to lead to $C(\beta, y)$. Its Fourier transform is c(x, y) from which the phase difference at any point can be calculated, namely

$$\Delta\phi(x,y) = \arctan\left[\frac{\operatorname{Im}\left\{c\left(x,y\right)\right\}}{\operatorname{Re}\left\{c\left(x,y\right)\right\}}\right]$$
(9.1.26)

The Fourier-transform method for fringe analysis is fast and needs no mechanical movements. However, closed fringes should be avoided in order to avoid ambiguity by the analysis. Fringe patterns can be opened by introducing additional tilts in the reference beam for instance. Ge et al., showed [01Ge] that the Fourier-transform technique can be extended to analyze closed fringes.

9.1.3.2.2 Fringe analysis by phase shifting

Phase-shifting techniques are mostly used for fringe analysis. For phase shifting the optical path difference is changed in one beam either continuously or in steps [74Bru, 82Doe, 91Cre, 92Gre, 93Sch, 03Har]. In the phase-shifting method the optical path difference between the reference and measuring beam can be changed by means of a mirror mounted on a piezoelectric element (PZT). Remember that in (9.1.10) we have three unknown, namely the mean intensity I_0 , the contrast γ as well as the phase difference $\Delta \phi = \phi_1 - \phi_2$. At least 3 independent equations are needed [82Doe]. However, due to possible errors by the phase shifting [91Cre] more phase-shifted interference patterns are used. Algorithms using an averaging technique have been introduced successfully by Schwider [83Sch, 93Sch] for instance. The introduced phase shifts are typically $\pi/2$ each, therefore



Fig. 9.1.10. Five phase-shifted interferograms with 4×90 degrees phase shifts introduced.



Fig. 9.1.11. Phase-shifting methods based on (a) a mirror on PZT, (b) a rotating plane-parallel plate, (c) a laterally shifted grating, (d) a rotating $\lambda/2$ polarizing plate. Additional methods for phase shifting are based on the change of wavelength. There are different methods to change the wavelength of a laser, e.g. changing the current of a diode laser or using a diode laser with an external cavity.

$$I_m = I_0 \left(1 + \gamma \cos \left\{ \Delta \Phi + m \, \frac{\pi}{2} \right\} \right) \,, \tag{9.1.27}$$

where $m = 0, 1, 2, 3, 4, 5, \dots$

The most commonly used algorithms for the calculation of the phases are the four- and fiveframe intensity data recordings with phase shifts of $\pi/2$ between.

For 4 frames (3 phase shifts between) the phase difference $\Delta \phi$ is given by

$$\Delta \Phi = \arctan\left(\frac{I_3 - I_1}{I_0 - I_2}\right) . \tag{9.1.28}$$

Considering 5 frames (4 phase shifts between). The phase difference $\Delta \phi$ is given by

$$\Delta \phi = \arctan\left(\frac{2(I_1 - I_3)}{I_4 + I_0 - 2I_2}\right) . \tag{9.1.29}$$

In Fig. 9.1.10 five phase-shifted interferograms with 4×90 degrees phase shifts introduced are shown. In Fig. 9.1.11 several phase-shifting methods are depicted.

A problem in phase-shifting interferometry for optical testing is the sensitivity to the environment such as air turbulence and vibrations, leading to a reduced measuring accuracy. Simultaneous phase shifting is a solution to reduce the influence of the environment.



Fig. 9.1.12. Phase cam set-up.

In Fig. 9.1.12 a holographic optical element separates the same wavefront to be measured into 4 spatially separated parts [03Wya]. Each wavefront will be phase-shifted by phase plates with a multiple of 90 degrees phase shift (9.1.27). The phase-shifted interference patterns will be detected and analyzed with a single detector (CCD).

9.1.3.2.3 Phase-locked interferometry

In phase-locked interferometry the phase of a reference beam is modulated sinusoidally with a small amplitude. Different techniques can be used such as controlling the injection current of a laser diode or moving a mirror. The time-varying signal detected by a point detector is filtered to obtain a signal that contains the original signal and the introduced modulated phase signal of the reference [03Har]. The value of this signal is held at zero using a feedback loop.

9.1.4 Interferometry for surface metrology

9.1.4.1 Interferometry with extended range and reduced sensitivity: Oblique incidence interferometry

Oblique incidence on the object reduces the sensitivity by $1/(\cos \vartheta)$, but with extended measuring range, as shown in (9.1.13) and (9.1.30). There are different possibilities and arrangements to be considered such as prismatic or diffractive configurations. In the diffractive method in addition to the reduction of the sensitivity the object shape can be adapted by the holographic optical element designed for oblique incidence and shape adaptation.

9.1.4.1.1 Prismatic interferometer with oblique incidence

It was shown in Sect. 9.1.2.3 that the optical path difference of a plane incident wave reflected at the front and back surface, respectively, is proportional to the thickness d and $\cos \vartheta$, where ϑ is the incidence angle on the reflecting surface. The basic principle of oblique-incidence interferometry with a prism is shown in Fig. 9.1.13a, where the light of a plane wave is coupled into a glass



Fig. 9.1.13. Prism interferometer with oblique incidence. (a) Basic principle. (b) Typical arrangement.

prism with refractive index n. An air gap (refractive index is considered to be 1.0) is separating the perfect reference surface from the test surface; the separation is h. The optical path difference OPD between the reference and object beam is $OPD = ABC - nAD = 2h \cos \vartheta$ and its phase is

$$\phi = 2\pi \frac{2h\cos\vartheta}{\lambda} = 2\pi \frac{2h}{\lambda_{\text{eff}}} , \qquad (9.1.30)$$

where $\lambda_{\text{eff}} = \lambda/(\cos \vartheta)$. Therefore, fringes of equal thickness with a reduced sensitivity are obtained. λ_{eff} can typically vary from 0.98 µm to 10 µm when ϑ varies from 50 to 86.4 degrees for

$$\lambda = 632.8$$
 nm.

For perfect reference and object surfaces a uniform fringe would be seen. Therefore, fringes obtained indicate defects $\Delta h(x)$ of the surface under test. The fringes will be analyzed as described in Sect. 9.1.3.2.

A typical prism interferometric arrangement we used is shown in Fig. 9.1.13b, where the reference mirror is not part of the prism. The reference mirror is moved for phase shifting by a piezo-element attached to it. In addition, it is slightly tilted to reduce unwanted reflections introducing some wedge or tilt fringes. By the arrangement the distortion in the interferogram can be reduced. A rotating ground glass reduces the spatial coherence in the time average and leads to a speckle reduction as will be discussed together with the speckle properties (Sect. 9.1.7). It should be noted that in addition to extend the measuring range for interferometric measurement, optically rougher surfaces can be measured as compared to interferometry with normal incidence. To extend the measuring range further synthetic wavelengths can be implemented by using two or more neighbor wavelengths of diode lasers [98Fra], see Sect. 9.1.4.2.

9.1.4.1.2 Grating interferometer with oblique incidence

For comparison a grating interferometer for oblique incidence is shown in Fig. 9.1.14. The CGH 1 (Computer-Generated Hologram) is splitting the incidence plane wave into the non-diffracted reference wave and the diffracted measuring wave. CGH 2 combines the reference wave and the wave coming from the object by diffracting the object wave in order to be parallel with the undiffracted reference wave in Fig. 9.1.14. Alternatively, the diffracted reference wave at CGH 2 could be combined with the undiffracted object wave. CGH 1 can, in addition, be used for shape adaptation. The phase variation due to the height variation is equivalent to that of the prism interferometer in (9.1.30) [86Sch].



Fig. 9.1.14. Grating interferometer with oblique incidence.

9.1.4.2 Multiwavelength interferometry

To reduce the sensitivity of surface measurements an IR wavelength with $\lambda = 10.6 \,\mu\text{m}$ can be used. Although the laser is available, high-resolution detectors with a comparable performance as obtained with CCDs in the visible are not yet available or too expensive. Two or more neighboring wavelengths can be used leading to beat wavelengths as will be seen by the two-wavelength heterodyne interferometry (Sect. 9.1.5). The beat wavelength for two wavelengths λ_1 and λ_2 is

$$\Lambda = \frac{\lambda_1 \lambda_2}{|\lambda_1 - \lambda_2|} \,. \tag{9.1.31}$$

As an example the interferograms of an aspheric surface are shown in Fig. 9.1.15a for $\lambda_1 = 822$ nm and in Fig. 9.1.15b for $\lambda_2 = 812$ nm. The synthetic wavelength is $\Lambda = 66.7$ µm. Fringes for the single wavelengths cannot be analyzed. By using both wavelengths, producing the synthetic wavelength, 3 fringes are observed as seen in Fig. 9.1.15c. The wavefront can be analyzed pixel by pixel, as shown in Fig. 9.1.15d from where the shape of the surface is obtained. For accurate measurements the interference pattern of a single wavelength can be used by taking into account (pixel-wise) the results obtained with the synthetic wavelength.



Fig. 9.1.15. Interferometry with 2 wavelengths, (a) $\lambda_1 = 822$ nm and (b) $\lambda_2 = 812$ nm; (c) is the interferogram obtained with the synthetic wavelength; (d) is the analyzed wavefront.

9.1.4.3 White-light interferometry

In white-light interferometry a short-coherence source is used. The 2π phase ambiguity discussed earlier can be overcome. However, the object needs to be scanned in depth. Starting at a point along the z-axis where the optical path is the same for the reference and test beam and moving the test surface along the z-axis in steps, the intensity by superposing the two interfering beams can be written:

$$I(z) = I_1 + I_2 + 2\sqrt{I_1 I_2} \ V(p) \cos\left(2\frac{\pi p}{\bar{\lambda}}\right) .$$
(9.1.32)

 I_1 and I_2 are the independently acting intensities of the two beams, V(p) is the visibility or coherence function (envelope of the fringes in Fig. 9.1.16b), p is the optical path difference. Interference can only be seen when the optical path difference between the reference path and the measuring path is smaller than the coherence length, which may be a few micrometers. Sources based on thermal white light, LEDs, multimode laser diodes, or super-luminescence diodes can be used. One way to measure the topography is to focus the microscope objective in order to obtain maximum visibility (zero optical path difference) for different object points by using a piezo-electric element. There are many algorithms for finding the coherence peaks (maximum fringe contrast). Combining the coherence-peak sensing and phase-shifting interferometry may provide the advantage of both methods, i.e. high precision and large dynamic range [90Kin, 92Dre].

There are different arrangements to be chosen, we concentrate on the Mirau arrangement only. The Mirau arrangement shown in Fig. 9.1.16a is a very robust set-up (nearly common path).



Fig. 9.1.16. (a) White-light interferometer arrangement according to Mirau and (b) the corresponding fringe pattern.



Fig. 9.1.17. (a) Interference fringes obtained with the white-light interference arrangement using a Mirau objective. The object was a coated Si waver. (b) Topography of the coated Si waver.

Figure 9.1.16b shows a typical fringe pattern by defocusing. For very high numerical aperture and extended working distance a Linnik arrangement is preferable. It is based on the Twyman–Green interferometer (Fig. 9.1.7), where two identical microscope objectives are used to focus the beams on the reference mirror as well as on the object, respectively.

In the Mirau set-up the illuminating wave is separated by the partly reflecting mirror into the reference and object beams and recombined after reflection at the mirror or test object, respectively, by the partly reflecting mirror.

In Fig. 9.1.17a as an example, results obtained by measuring a Si waver are shown. Interference fringes can only be seen when the Optical Path Difference (OPD) between reference and object wave is smaller than the coherence length. The maximum contrast occurs when the OPD is zero, hence the topography can be obtained; the topography of the Si waver is shown in Fig. 9.1.17b.

9.1.4.4 Polarization interferometry

The basic principle of the interferometric methods presented until now is the comparison of the object wave with an external reference. Typically this external reference is provided by the wavefront originating from a well characterized reference object. If such an external reference is not available or not preferable for technical reasons one can also compare the object wavefront with itself by dividing the wavefront and shifting one part by a certain amount. In dependence on the magnitude of this separation we talk about shearing interferometry with a differential or total shear. Picking up this idea of self-reference the vectorial character of the light field provides a further possibility of interferometry: namely the interferometric comparison of two orthogonal polarization states of the same wavefront originating from the object. This measurement principle is especially useful for linear microstructures with a size in the order of a wavelength. In this case we have a strong polarization dependence of the sample-light interaction and the polarization state divides in two eigenpolarizations, imprinted by the structure on the light field: One parallel to the line structures (TE-polarization) and one perpendicular (TM-polarization). A straightforward way to perform such interferometric measurements is using a liquid-crystal phase shifter with a following analyzer. The details of the method and its application to metrology of subwavelength structures can be found in [00Tot, 04Tiz].

In the following we will describe a related measurement technique in more detail that allows determining the complete partially coherent polarization state of the light field encoded in the coherence matrix. It falls back to classical principles of polarimetry.

The experimental set-up is based on a polarization microscope, operating in Deep UltraViolet (DUV) at 248 nm. Figure 9.1.18 shows the schematic set-up. As light source serves a xenon-mercury lamp, which is followed by a monochromator ($\lambda = 248$ nm, $\Delta \lambda = 5$ nm). The freely accessible Koehler illumination set-up incorporates besides an adjustable aperture stop a polarization-state generator, consisting of a polarizer with a following rotating quarter-wave plate. By this means the illuminating polarization state can be chosen freely. For imaging we used a cemented DUV-objective $150 \times /0.9$. The post magnification is stepwise eligible between $1 \times$ and $6 \times$. The polarization state emerging from the incoherently illuminated object is measured by the Polarizer-Sample-Compensator-Analyzer (PSCA) principle in combination with a Fourier-analytic evaluation procedure [93Col]. For this purpose a sequence of typically 8 images is captured for different uniformly distributed orientations of a quarter-wave plate followed by a fixed orientated analyzer. The pixel-wise evaluation of the image stack leads to a linear system of equations.



Fig. 9.1.18. Set-up for high-resolution polarimetry.



Fig. 9.1.19. Diagonal elements of the measured coherence matrix. The dependence of the contrast and the position of the intensity minima indicating the edge position with respect to the illuminating polarization state are clearly visible.

provides the 2D coherency matrix, which describes in the limit of high magnification a partially polarized light field completely.

The coherency matrix is the basis for determination of different quantities, valuable for precise quantitative image interpretation and evaluation. To these belong degree of coherence, degree of polarization, relative phase retardation etc. On the basis of this formalism polarization-induced interaction processes can be measured within a single measurement cycle. In the following, we give a measurement example, which demonstrates the validity of the method. The examined object consists of a beam-like structure of 160 nm height separated by a 330 nm wide groove. The illumination was linearly polarized under 45° relative to the structure's edges with the full numerical aperture of NA = 0.9. Figure 9.1.19 shows the diagonal elements of J as well as the overall intensity $I = J_{xx} + J_{yy}$ (divided by the factor of two for better comparison) along a cross section of the structure.

The direct comparison shows, that the different interaction behavior for both eigenpolarizations of the structure causes a conspicuous deviation in the corresponding intensity distributions. So, TM-polarization shows a stronger contrast in the intensity profile and more pronounced intensity minima in comparison to TE-polarization. This is caused by more efficient scattering at the structure edges for TM as in this case the electrical-field vector matches the physical boundary condition to be preferably perpendicular on the Si edges. The physical boundary condition also forces the tangential component of the electric field (i.e. TE-polarization) to be continuous, while the normal component (corresponding to TM-polarization) performs a jump at the surface border. This behavior results in a pushing effect of TE-polarized light away from the structure edges. This thesis is confirmed by the fact, that due to these effects the penetration of TE-polarized light in the groove (with dimensions in order of the wavelength) is strongly reduced and results in a further reduction of the contrast in comparison to singular edges. In TM-polarization, however, this effect is much less pronounced. The field displacement described above is also the reason for different positions of the intensity minima in TE- and TM-polarization. These minima are typically used for edge detection of pure phase objects. In the present case the positions are different in the order of about 50 nm. This means that a difference in linewidth measurement of about 100 nm occurs.

Besides the pure polarization-dependent intensity profile, represented by the diagonal elements of J, the nondiagonal elements of the coherency matrix can also be used for quantitative image interpretation. For this purpose we use the complex correlation factor of (9.1.20). It is a measure of the mutual correlation of the polarization states in x- and y-direction. It represents the different interaction behavior of the structure's eigenpolarizations and is localized in the optical image at structure details such as edges. Figure 9.1.20 (left) shows the relative phase retardation induced by the measured structure. It has the same physical origin as the different intensity profiles described above: While TM-polarized light propagates along the Si edge, TE-polarized light is pushed away from the edge and therefore is ahead in phase. Here also the effect is strengthened in the groove region. TE-polarization is more ahead on phase than at the singular edges due to the reduced pen-



Fig. 9.1.20. Nondiagonal elements J_{xy} of the coherency matrix. Left: effective relative phase difference, right: degree of coherence.

etration capabilities into the groove. The right-hand side of Fig. 9.1.20 reveals a further important detail of polarization-dependent image formation. In the incoherently illuminating light field the polarization-dependent scattering process causes a strong reduction of mutual coherence, which is equal to the degree of polarization in the present case of 45° -orientated illumination (in this case J_{xx} and J_{yy} are equal). This means, that the minima in the intensity images are strongly linked to depolarization effects located at the structure edges.

9.1.5 Heterodyne interferometry for velocity and distance measurement

9.1.5.1 Principle of heterodyne interferometry

In heterodyne interferometry, a small frequency shift is introduced in one of the interferometer beams. It is a very powerful tool for the measurement of movements and velocities and distances by integrating the velocity. It is related to the well-known Doppler-shift measurement [71Dur]. The basic principle of heterodyne interferometry follows from the superposition of two waves A_1 and A_2 , where one is frequency-shifted. It should be noted that perfect coherence is assumed. Describing the two waves as

 $A_1 = a_1 \cos(\omega_1 t + \Phi_1)$ and $A_2 = a_2 \cos(\omega_2 t + \Phi_2(t))$,

where $\Phi_2(t) = \Phi_2 \pm \frac{2\pi}{\lambda} 2v(t)t$, we obtain by superposing the two waves

$$I = |A_1 + A_2|^2 = |a_1|^2 + |a_2|^2 + 2a_1a_2 \cos\left\{\Delta\omega t \pm \frac{2\pi}{\lambda} 2v(t)t + \Phi_1 - \Phi_2\right\}, \qquad (9.1.33)$$

where Φ_1 and Φ_2 are constant, not time-dependent, phases of the amplitudes a_1 and a_2 of the two interfering beams. v(t) is the velocity of the moving object, $\Delta \omega = \omega_2 - \omega_1 = 2\pi (f_2 - f_1)$.

There are different methods to introduce the frequency shift. Frequently an Acousto-Optical Modulator (AOM) is used, working at 40 or 80 MHz for instance [88Dae, 91Sod]. Because of its high working frequency it can be more convenient to carry out the analysis at a lower frequency. Therefore 2 acousto-optic modulators can be used with slightly different frequency shifts; hence for



Fig. 9.1.21. Principle of the heterodyne interferometer with polarizing beam splitter.

Landolt-Börnstein New Series VIII/1A2

the analysis the interferometer is working on the difference frequency, typically 100 kHz. Alternative phase-shifting methods are a rotating grating or a rotating $\lambda/4$ plate or the Zeeman splitting of a laser where a frequency shift is introduced in the laser cavity between two polarization directions. In Fig. 9.1.21 the principle of an arrangement used for distance measurement is shown. It is a Twyman–Green arrangement with retro reflectors instead of the plane mirrors for the reference and object beams in order to avoid the influence of small vibrations and tilts. The moving object leads to a frequency shift due to the time-varying phase $\Phi(z,t) = \Phi_0 + \Phi(t)$. The heterodyne frequency introduced by the interferometer is $f_2 - f_1$, the frequency shift introduced by the moving object is

$$\Delta f = \frac{2}{\lambda} \frac{\mathrm{d} z(t)}{\mathrm{d} t} = \pm \frac{2v}{\lambda} , \qquad (9.1.34)$$

where z(t) is the displacement variation projected onto the line of sight. Subtracting the known introduced frequency difference $f_2 - f_1$ from the measured frequency leads to Δf and hence to the speed of the movement. By integrating over the measuring time we deduce the distance z:

$$z = \frac{\lambda}{2} \int_{T_1}^{T_2} \Delta f \, \mathrm{d} t \;. \tag{9.1.35}$$

It is a very accurate incremental measuring principle used first in a commercial instrument by Hewlett Packard (HP). In the basic principle of the interferometer (Fig. 9.1.21) the two waves coming from the laser are polarized in two directions perpendicular to each other and have a slightly different frequency $f_2 - f_1$ (Zeeman splitting in the cavity). The polarizing beam splitter separates and recombines the two waves. A 90 degree rotation is required for the two waves to travel to the detector after reflection at the corner cubes (retro reflectors). Moving the retro reflector attached to the object leads to the frequency shift, from where the velocity respectively distance can be obtained. The method is frequently used for distance measurement in industrial environment. It is an accurate and robust but incremental method. An uncertainty of 10⁶ can be achieved when a stabilized laser is used as light source. The measuring accuracy is limited by atmospheric disturbances. For accurate measurements, the atmospheric condition needs to be measured and its influence compensated.

9.1.5.2 Absolute heterodyne interferometry: Double heterodyne interferometry (DHI)

It has been pointed out that the heterodyne interferometer is an incremental interferometer, there is ambiguity whenever the optical path difference is equal or larger than the wavelength of the laser light used. To overcome some of the drawbacks of classical heterodyne interferometry, such as ambiguity but also its limitation to the use of highly polished surfaces only, a Double-Heterodyne-Interferometry (DHI) principle can be chosen [88Dae, 91Sod].

A Double Heterodyne Interferometer (DHI) can be considered to consist of two superposed heterodyne interferometers where each works on a slightly different wavelength. It can be shown that the two heterodyne interferometers lead to a beat frequency of the two wavelengths to be used for the distance measurement as discussed in [91Sod]. A developed double heterodyne interferometer will be discussed briefly.

The basic principle of an absolute heterodyne interferometer is shown in Fig. 9.1.22, where a high-frequency AOM is illuminated by two waves from a HeNe laser for instance with wavelengths λ_1 and λ_2 , corresponding to the frequencies ν_1 and ν_2 , where $\nu = c/\lambda$. The non-diffracted beams are focused onto the object under test by the lenses L_1 and L_2 after passing a Polarizing Beam Splitter (PBS) and a Quarter-Wave Plate (QWP) and will be combined with the reference after passing the PBS and the lens L_3 as well as a Half-Wave Plate (HWP). The heterodyne difference frequency is introduced by a Rotating Grating (RG) in the reference beam. The AOM is operated at a driver frequency $f_{\rm d}$ leading to two frequencies in the first-order diffracted beam, namely for λ_1 of $\nu_1 + f_d$ and for λ_2 of $\nu_2 + f_d$. The diffracted reference beams from the AOM traverse two adapted diffraction gratings G and RG. G has a high spatial frequency (600 $\rm lp/mm)$ and splits the two HeNe wavelengths in α_1 and α_2 (first diffraction order) leading to the difference $\Delta \alpha$. The second, a Rotating Grating (RG), compensates the diffraction angle difference $\alpha_2 - \alpha_1$ of the two reference beams with frequencies $\nu_1 + f_d$ and $\nu_2 + f_2$ in such a way that the zeroth order of one is parallel to the diffracted first order, frequency-shifted by f_m . The first-order diffracted beams are frequency-shifted by the rotation of RG by f_m whereas the zeroth order is not. After RG the zeroth-order beam having a frequency $\nu_2 + f_d$ and the first-order with a frequency $\nu_1 + f_d + f_m$ are parallel, as well as the zeroth order with the frequency $\nu_1 + f_d$ and the minus-first order with the frequency $\nu_2 + f_d - f_m$, one pair is selected only. After superposition with the object wave the signal can be directly applied to a low-frequency phase meter (for example a two-channel lockin amplifier). The reference for the lock-in amplifier was directly taken from the angle-encoder detector (square wave signal) output. The beat of the two heterodyne signals can be detected after demodulation and band-pass filtering (to remove the carrier) and neglecting a stochastic phase term in the carrier. The detected signal is:



Fig. 9.1.22. Double heterodyne interferometer with rotational matched grating.

9.1.6 Interferometry with adaptive optics

$$i(t) \propto I_0 \cos\left[2\pi f_m t - \frac{2\pi\Delta z}{\Lambda}\right],$$
(9.1.36)

where Δz is the optical path difference between object and reference paths,

$$\Lambda = \frac{\lambda_1 \lambda_2}{|\lambda_1 - \lambda_2|} = \frac{c}{|\nu_1 - \nu_2|} . \tag{9.1.37}$$

Two-wavelength double heterodyne interferometry has proven to be a powerful tool for accurate interferometric measurements on smooth as well as on optically rough surfaces, providing that the roughness is slightly smaller than the depth resolution. For example by using a HeNe laser with the two wavelengths $\lambda_1 = 632.8$ nm and $\lambda_2 = 640.1$ nm, leading to $\Lambda = 55.5$ µm, and for a target distance of 6.5 mm, an accuracy of 0.3 µm was obtained. For high accuracy the singlewavelength heterodyne technique could be used with the same set-up. It should be noted that there are different set-ups for two-wavelength double heterodyne interferometry [88Dae, 91Sod, 96Tiz1, 96Tiz2, 97Tiz1].

9.1.6 Interferometry with adaptive optics

9.1.6.1 Interferometry with a null corrector

For testing polished optical plane or spherical surfaces to a shape accuracy of some nm RMS, interferometry is mainly applied. To measure or compare complicated wavefronts such as the shape of polished aspheric surfaces, holographic interferometry could be extremely useful. If a master optical component or optical system is available a hologram can be made of the wavefront produced by the component or system. The stored wavefront can be used to perform null tests of similar optical components or systems. If a master optical system is not available for making the hologram, a synthetic or Computer-Generated Hologram (CGH) can be produced to provide the reference wavefront. The CGH can be thought of as a binary representation of the ideal interferogram that would be produced by interference of the reference wavefront with the wavefront from a perfect aspheric surface for instance. In making the CGH the ray tracing needs to be done for the whole interferometer to determine the so-called aspheric wavefront at the hologram plane. Alternatively the compensator, also called null lens, can be designed as a refractive lens system, but in our days more frequently it is a diffractive element [72Wya, 85Doe].

Different interference arrangements can be used for testing polished optical surfaces, such as Fizeau, Twyman–Green, or Mach–Zehnder arrangement as discussed in Sect. 9.1.3. A Twyman– Green arrangement is shown in Fig. 9.1.23 with a CGH as null corrector for testing optical components and systems in transmission (Fig. 9.1.23a). The reflecting CGH is compensating the errors introduced by a refractive null lens or an element of an optical system to be tested. The use of the CGH for testing an aspheric surface is shown schematically in Fig. 9.1.23b. Frequently a lens system as used for testing spherical surfaces (Fig. 9.1.7 and Fig. 9.1.8) is added. It is useful in order to reduce the structure density of the CGH as shown in Fig. 9.1.23c. It should be noted that the null corrector in the test arm is designed in such a way that the rays hit the perfect aspheric surface under test perpendicular. For a perfect surface under test the reflected rays pass the null corrector again and are directed parallel to the incoming rays. The wave under test is combined with the reference wave, leading to interference patterns. If the measuring system is perfectly adapted only deviation from the perfectly designed aspheric shape leads to interference fringes to be analyzed in order to measure the shape deviation. However, care needs to be taken with respect to the positioning and alignment of the components of the test system and null corrector. A Fizeau set-up could equally well be used; the difference is mainly in the path of the reference beam.



Fig. 9.1.23. (a) Interferometer based on the Twyman–Green principle with a CGH as null corrector for testing a system in transmission. Set-ups for testing a surface in reflection with a CGH (b) without and (c) with an additional collimating objective.



Fig. 9.1.24. (a) Multiple-function CGH with a ring-shaped reflection hologram (amplitude hologram) for adjustment and centering. The test hologram is a phase hologram. (b) Interference pattern of a not yet perfectly adjusted CGH.

For testing with null correctors, the null lens needs to be designed and manufactured with very high precision and tested. For a CGH as null corrector it can be an amplitude or phase element. To increase the diffraction efficiency phase holograms are used.

CGH's can be generated by using a photo plotter or by E-beam technology. As pointed out before, care needs to be taken in order to accurately adjust the CGH together with the components of the interferometer. Recently additional holograms for adjustments and centering were used [03Rei, 04Tiz]. In our case additional adjustment holograms are generated on the same substrate practically at the same time in order to reduce systematic errors in the testing and calibration procedure [03Rei]. In Fig. 9.1.24 the adjustment hologram is a reflecting amplitude hologram generated on the same substrate and made practically together with the CGH null system. The adjustment hologram shown schematically in Fig. 9.1.24a is placed at the periphery. In Fig. 9.1.24b an interference pattern of a not yet perfectly adjusted CGH is shown. The reference CGH makes the adjustment like positioning and centering of the hologram much simpler.

It should be noted that there are more parameters to describe an aspherical surface than a spherical. The degrees of freedom during alignment cause more types of corresponding wavefront aberrations than for spheres and centering tolerances for instance are more severe.



Fig. 9.1.25. Comparison of the results by measuring an aspheric surface with 3 different CGH's. (a) DI-CGH, (b) ITO-CGH, (c) Jenoptic-CGH.

After the interference patterns are obtained, the interference fringes are analyzed using one of the well-known techniques (Sect. 9.1.3.2.2) and the deviation from the perfect required shape of the aspheric surface is determined. Adjustment and calibration are very important for high-precision measurements [03Rei]. In Fig. 9.1.25 the results of an aspheric surface tested with three different CGH's produced from different laboratories with different techniques are shown. The manufacturers were (a) Diffraction International (DI), (b) ITO, University of Stuttgart, and (c) Jenoptic.

9.1.6.2 Adaptive optics with optical light modulator

Conventional adaptive optical systems for laser beam propagation or imaging are used in connection with astronomical telescopes, for instance. They consist of a wavefront sensor to detect the optical disturbance, a deformable mirror to correct the optical disturbance, and a computer to decode the sensor information for the active mirror. Interesting elements like Spatial Light Modulators (SLM's) as well as membrane mirrors are available today. The development of SLM's was driven by the data projectors, a consumer product. SLM's typically contain a matrix of individually addressable pixels of liquid-crystal cells or flip mirrors and more recently piston mirrors. Pixel sizes are on the order of tens of microns and therefore appropriate as active diffractive elements or flexible miniature masks. In consumer products the SLM's are mainly optimized for intensity modulation in transmission or reflection and not for phase modulation as it would be useful for some applications in wavefront shaping and analysis to be discussed, but progress has been made recently. For the analysis of the wavefront, Zernike polynomials are frequently used.

Wave aberrations limit the performance of optical systems. The aberration can be compensated using an appropriate hologram. At first, however, the wavefront needs to be determined. Interferometry or a Hartmann or Shack–Hartmann sensor could be appropriate. A method related to Young's double-slit experiment has been adapted together with an LCD as shown schematically in Fig. 9.1.26. The two "pinholes" generated by the LCD are moved across the pupil of the system under test. Interference between the two point sources occurs as in Young's experiment. The maximum of intensity appears, when the optical path difference, including the phase due to aberrations is a multiple of 2π . The fringe separation is inversely proportional to the separation of the pinholes.

The wavefront measurement obtained by analyzing the phase difference of the waves from the two pinholes carried out over the whole pupil is the base to produce the compensating hologram. The experimental set-up for wavefront measurement and wavefront correction is shown in Fig. 9.1.27a. The hologram generated is shown in Fig. 9.1.27b. The results obtained without and



Fig. 9.1.27. (a) Experimental set-up for wavefront measurement and wavefront correction, based on the application of an LCD. (b) Phase distribution of the compensating hologram of the system with aberrations. (c) Point spread functions without and (d) with wavefront correction.

with the hologram introduced into the pupil plane are shown in Fig. 9.1.27c and Fig. 9.1.27d, respectively. It should be noted that the holograms were written in the LCD.

9.1.6.3 Adaptive optics with deformable membrane mirror

To correct wavefronts or adapt the wavefront to an aspheric surface a deformable membrane mirror can be useful as long as the local slope is not too high.

A typical membrane mirror consists of a thin, typically 1 μ m thick, aluminum-coated silicon nitride membrane stretched over an array of electrodes. Due to the electrostatic forces, the membrane deforms by applying an electrostatic field between membrane and electrodes. Since the membrane deformation is smooth, the wavefront deformation is smooth, too.

The wavefront deformation depends on:

- the material constants and stress of the membrane,
- the gap between electrodes and membrane, and
- the voltage applied.





Fig. 9.1.28. (a) Principle of a membrane mirror. (b) Electrode configuration for hexagonal electrode structure and (c) for ring-shaped structure.

Fig. 9.1.29. Arrangement for testing an aspheric surface by the use of a membrane mirror to reduce the fringe density.

Material constants and stress depend on the fabrication process. The gap defines the maximum deflection. The deflection should not exceed 70% of the gap width [95Vdo]. For small deflections, the deflection is proportional to the square of the applied voltage. The basic principle of a membrane mirror is shown in Fig. 9.1.28a, two electrode configurations are shown in Fig. 9.1.28b and Fig. 9.1.28c with hexagonal and circular electrode arrangement, respectively.

In Fig. 9.1.29 a set-up of the dynamic "null lens" for testing an aspheric surface is integrated into a Twyman–Green interferometer set-up [04Pru]. The polarized beam in the test arm passes the polarizing beam splitter and the quarter-wave plate, is reflected back from the membrane mirror surface, and passes the retardation plate again. The quarter-wave plate in double pass rotates the polarization by 90 degree. The returning beam from the aspheric test surface passes the membrane



Fig. 9.1.30. Fringe pattern (a) without and (b) with partial wavefront compensation.

mirror again, which leads to a wavefront deformation that is four times the deformation of the membrane. For calibration of the membrane mirror the quarter-wave plate is rotated in order to have the reflected wavefront from the membrane mirror to be compared with the reference beam.

Membrane mirrors with a diameter up to 50 mm were used [95Vdo, 04Pru]; the experimental results shown in Fig. 9.1.30 were obtained with a diameter of 25 mm and a gap width of 70 μ m. For driving the 37 channels of the mirror with a voltage up to 700 V a HV amplifier electronics was developed. Deflections up to 40 μ m PV were obtained with a shape reproducibility of 50 nm. Figure 9.1.30a shows an interference pattern without and Fig. 9.1.30b with wave compensation.

9.1.6.4 Adaptive optics for optical stitching using dynamically tilted reference wave

For resolving high-density interference patterns a stitching technique could be used in order to select a small portion of the interference pattern at the time. By combining the different overlapping portions of the interferograms the whole wave field is obtained.

A very promising technique for a flexible wavefront measuring of otherwise not resolvable interference patterns is a flexible tilted reference or target wave [04Lie]. The technique is in the process of being developed for testing aspheric surfaces. The principle shown in Fig. 9.1.31 is based on a Twyman–Green interferometer. The main difference is a dynamically and flexibly tilted reference wave. The key elements are the addressable LCD together with the pinholes, defining a perfect spherical wave with a well defined selectable tilt angle. Details of the key element are shown in Fig. 9.1.32.

To explain the working principle a plane wave is falling obliquely on the LCD and is focused by micro lenses in the pinhole plane, but missing the pinholes all together. By generating a grating in the electrically addressed LC elements the first diffraction order deviates the selected beam (or beams) to pass the appropriate pinhole. The point source is now active and leads to a tilted reference wave after the collimating lens. Furthermore, a phase shift is introduced by laterally shifting the displayed grating proportional to the shift/grating period. The amplitude of the reference wave can be adapted by the diffraction efficiencies of the generated phase grating independently for the different point reference sources.

The LCD used in the described set-up was from CRL OPTO and had 1024×768 pixels with a pixel pitch of 36 µm. From the micro-lens array 70×70 LCD pixels were used in front of each micro lens. Gray-tone structures with 8 levels were produced by direct writing into photo resist with the photo plotter and transferred into fused silica substrate by dry etching: Reactive Ion Etching (RIE). The array of pinholes was etched into a black chromium layer as the absorbing layer with a separate mask. The micro-lens array works in its first diffraction order (Fig. 9.1.32).

On the back side of the 9 mm silica substrate were the pinholes generated by a separate mask aligned with respect to the existing micro-lens array. For the alignment of the pinholes alignment marks [01Gru] were printed onto the mask for the pinholes and the lenses simultaneously (self-imaging). 70×70 LCD pixels were used in front of each micro lens.



Fig. 9.1.31. Principles of dynamic reference beam.



Pinhole

array

Fig. 9.1.32. Detail for generating the tilted reference wave and the phase shifting.



Fig. 9.1.33. Single-phase measurements of a defocused reference surface with 5×5 tilts (left), the result after unwrapping of 5×5 interferograms is shown (right).

Figure 9.1.33 shows first experimental results of the new stitching procedure. On the left are single-phase measurements of a test surface where different tilts were introduced into the reference beams, on the right is the unwrapped wavefront. In a comparison with simulated results after some calibration of the experimental set-up, the RMS value agreed by $\lambda/14$: The accuracy can be further improved. Alternatively, the object wave can be tilted (stitching) in the same way as demonstrated with the reference wave, which could even be more appropriate for practical applications.

9.1.7 Speckle pattern interferometry

The speckle phenomenon has long been familiar, but only the introduction of the laser, 1960, has brought attention and a deeper understanding. Laser speckles appear whenever an optically rough surface, with a roughness on the order of the wavelength or coarser, is illuminated with highly coherent light. The interference of the dephased, but coherent, secondary spherical wavelets originating from the surface results in granular patterns of intensity, commonly called speckle patterns. The path difference of these wavelets may be different by many wavelengths. Speckle

patterns appear everywhere, also in free space. Speckles can easily be seen when the wave field is intersected by a screen or a two-dimensional detector. Furthermore, the waves scattered from an optically rough surface have not only random phases but also random amplitudes. Analogous statistical phenomena are found in radio-wave propagation.

The speckle phenomenon is often a nuisance but there exist several techniques to reduce the graininess. Reducing the spatial coherence of the illumination [84McK] is one way, introducing a phase modulation or small vibration during the exposure is another way to reduce the speckle contrast in the time average. Multiple apertures can reduce the speckle contrast in the image plane as well.

Speckle techniques can be a powerful tool for studying surface properties of objects that are optically rough. In recent years, faster cameras with high resolution along with high memory and faster interfaces have revitalized the use of speckle techniques from its early stages of infancy to a tool very reliable for industrial tasks. The attractiveness of speckle methods lies in its ability to vary the sensitivity and range based both on speckle photography and speckle interferometry. Speckle techniques provide a large measurement range for in-plane and out-of-plane displacement, displacement derivative, and shape measurement. Another aspect of speckles is that they are generated due to the roughness of the object and therefore have also been used to measure roughness parameters.

Speckles and speckle photography have been described in many articles and books [70Arc, 72Tiz1, 72Tiz2, 75Ste, 78Erf, 79Fra, 80Tiz, 84McK, 89Jon, 93Sir, 93Yam, 94Oka, 94Sre, 01Leh, 01Ras1]. Speckle photography is based on the analysis of the speckle pairs resulting from the speckles before and after lateral displacement of the object or part of it.

By contrast, in speckle interferometry the phase variation of the speckles is analyzed; they should not move between the two exposures. Speckle interferometry is therefore used to detect small movements as is the case in holographic interferometry. Before concentrating on speckle applications some properties of speckles will be discussed.

9.1.7.1 Some properties of speckles

Speckle statistic is described in different books and papers [65Gol, 70Hop, 75Goo, 78Erf, 84Dai, 84Per, 89Jon, 01Leh]. Therefore some relevant results for its application in speckle interferometry will be summarized only. Scattered light has a grainy appearance when viewed in space. These speckles are termed objective speckle pattern. This graininess-like appearance can be observed on the image of the object as well. The speckles in the image are called subjective speckles (Fig. 9.1.34). The formation of speckles is random. The first-order statistics of a monochromatic polarized speckle



Fig. 9.1.34. Speckle formation and speckle size (for rectangular pupil).
deals with the probability density function of irradiance at a single point for an ensemble average of scatterers. The amplitude of the light wave at a given point of observation is considered to have contributions from different scattering regions of the surface. This pattern is identical to the classical pattern of random walk in a complex plane.

It is assumed that the scattering regions are statistically independent of each other and the phases are uniformly distributed between π and $-\pi$. It can be shown that the irradiance at a point for linearly polarized Gaussian speckles follows the negative exponential statistics which is given by

$$P_{I}(I) = \frac{1}{\langle I \rangle} \exp\left(-\frac{I}{\langle I \rangle}\right), \quad I \ge 0,$$

$$P_{\phi}(\phi) = \frac{1}{2\pi}, \quad \phi \in [0, 2\pi],$$
(9.1.38)

where the mean intensity value of $\langle I \rangle$ is $2\sigma^2$, given by the variance [75Goo, 84Dai].

Consequently, we find that the most probable brightness of the speckle in a fully developed polarized speckle pattern is zero. The contrast of the speckle pattern is a ratio of the variance to the average irradiance. Using this definition, the contrast of a polarized speckle is found to be unity. Speckle patterns can be added both on an amplitude and intensity basis. In amplitude addition, there is no alteration in the statistics, except for a scaling factor. Addition on the intensity basis changes the statistics and the most probable brightness of the speckle is not zero anymore. The density function as a result of this is a sum of the simple decaying exponentials and the contrast of the speckle pattern assumes a minimum value of $1/\sqrt{2}$. The contrast of a sum of N speckle patterns on an intensity basis turns out to be $1/\sqrt{N}$.

For the description of other properties such as coarseness of the spatial structure of the speckle pattern or the spatial variation of amplitude and phase the second-order statistics needs to be considered as discussed by Goodman and Dainty [75Goo, 84Dai].

To describe the speckle size in an image-forming system let us consider the coherent image formation as a Fourier transform from the object to the entrance pupil, EP, and a second Fourier transform from the exit pupil, AP, to the image plane (Fig. 9.1.34) neglecting aberrations. α , α' are aperture angles, E, E' are the center of the pupils. If we put a screen or intensity recording detector in the entrance pupil, a speckle pattern as seen in the Fraunhofer plane would be observed (objective speckle), the mean speckle size and shape depends on the illuminated object area and shape and is basically given by the interference of the waves coming from the edge of the illuminated area. The lateral speckle size for an illuminated rectangular object area would be $\approx \lambda/\theta$, where θ is the total angle of the illuminated area in x-direction as seen from the middle of the entrance pupil (Fig. 9.1.34). For the image plane we consider the Fourier transformation of the random phase variation superposed on the object information on the entrance pupil and transferred to the exit pupil (neglecting aberrations) (Fig. 9.1.34). The information of the image is superposed to the speckle field (subjective speckles). The smallest subjective speckle size in the image plane (rectangular aperture) is $\Delta x'_{\min} \approx \lambda/\theta_0$, where θ_0 is the total aperture angle in the image space (Fig. 9.1.34). It turns out that it is also the dominant speckle size.

Assuming that we put a screen in the entrance pupil we would obtain the speckle size in the entrance pupil EP for a circular area of illumination

$$\Delta x_{\min} = 1.22 \lambda \frac{z}{D} \tag{9.1.39a}$$

and in the image space

$$\Delta x'_{\min} = 1.22 \,\lambda \,(1+M)F \,, \tag{9.1.39b}$$

where z is the distance from the illuminated object to the observation screen (the entrance pupil in Fig. 9.1.34). D is the diameter of the illumination area. For the image formation a circular aperture is considered, where F is the F-number of the image-forming system, M is the magnification, and

 λ is the wavelength of light used. It should be noted that the diameter of the speckle in the image is inversely proportional to the aperture of the imaging system whereas the intensity of the speckle is directly proportional to the square of the diameter. The speckle in 3D space has the shape of a cigar; the length of the longitudinal speckle size for a circular aperture can be expressed as

$$\Delta z' = \frac{8\lambda}{\theta_0^2} \,. \tag{9.1.40}$$

For dark spots in a speckle pattern the intensity is zero as well as the phase in these regions, however, walking around the zero intensity regions one observes that the phase goes through a full 2π circle. These regions of phase jumps are called dislocations or phase singularities. It should also be noted that light scattered from most common objects does not necessarily retain the incident polarization state of the incident optical wave (Goodman [75Goo]).

9.1.7.2 Speckle applications

There are basically two important speckle application techniques for displacement-, deformation-, stress-, and vibration measurements, namely speckle photography and speckle pattern interferometry. Speckle photography is based on the displacement of the speckles due to an object displacement between two exposures [70Arc, 72Tiz1, 72Tiz2, 75Ste, 78Erf, 78Gre, 80Tiz, 93Yam, 94Oka, 94Sre, 99Sir]. Therefore two speckle patterns are recorded one before and one after deformation. There are different techniques such as the use of the principle of Young's experiment where the pairwise separated speckles stand for the pinholes in the Young experiment where the fringe spacing is inversely proportional to the spacing of the pinholes. To simplify let us consider one individual speckle before and after the object movement. Let us consider that the speckle moves with the object, which agrees for small movements. From Young's fringe separation the object displacement can be evaluated. Therefore the fringe spacing is inversely proportional to the speckle movement. Speckles are usually stored on photographic emulsion, before and after displacement. Instead of storing the speckles on photographic emulsions, photothermoplastic or photorefractive storage material or a CCD can be used [80Tiz, 94Oka]. Great progress has been made by using CCD storage devices and fast analysis. By contrast to speckle photography in speckle pattern interferometry the phase variation inside each speckle leads to the required information. We will concentrate on the application of speckle pattern interferometry and on digital holographic interferometry.

9.1.7.3 Speckle pattern interferometry for deformation measurements

In the previous Sects. 9.1.1 to 9.1.6 interferometry on optically polished surfaces was discussed. Speckle pattern interferometry, Electronic Speckle Pattern Interferometry (ESPI), and digital holography are whole-field interferometric techniques appropriate to study small deformations of objects with optically rough surfaces. The methods are based on interference of scattered light from the object superposed with a reference wave. The result is again a speckle pattern with randomly distributed speckles (with a random phase and amplitude) recorded on a CCD or CMOS chip, for instance. The changes in the object state due to displacement, deformation, or stress lead to phase changes within each speckle. A comparison with the phases of the speckles of the original object (without change) is the basis for ESPI [71But, 74Hun, 75Ste, 78But, 78Enn, 84Lok, 85Cre, 86Joe, 90Joe, 90Ker, 91Sir, 92Rod, 93Dav, 93Hun, 93Kau, 93Lok, 93Ped, 93Pry, 94Joe, 94Zou, 96Mei, 97Her, 97Hun, 97Ras1, 97Ras2, 98Ste, 01Kum, 01Ras2, 04Joe] and digital holography. In the ESPI method the speckle patterns before and after deformation of the object are recorded on a photoelectric detector and added or subtracted electronically with the appropriate software.



Fig. 9.1.35. Set-up for electronic speckle pattern interferometry.

A typical ESPI arrangement is presented in Fig. 9.1.35. A beam splitter separates the laser beam into an object and reference beam. The reference beam is reflected by a mirror mounted on a piezo-electric component (PZT) used for phase shifting, which is very useful for the analysis, as has been described in Sect. 9.1.3.2.

The resulting intensity in the recording plane (on a CCD or CMOS chip) of the overlapping waves at a given transverse point (x, y) for a given longitudinal position z and time t is given by

$$I(x, y, t) = |U_{\rm R}(x, y, t) + U_{\rm O}(x, y, t)|^{2}$$

= $|U_{\rm R}(x, y, t)|^{2} + |U_{\rm O}(x, y, t)|^{2}$
 $+ U_{\rm R}(x, y, t) \cdot U_{\rm O}^{*}(x, y, t) + U_{\rm R}^{*}(x, y, t) \cdot U_{\rm O}(x, y, t)$ (9.1.41)

and

$$I(x, y, t) = I_{\rm R}(x, y, t) + I_{\rm O}(x, y, t) + 2\sqrt{I_{\rm R}(x, y, t)I_{\rm O}(x, y, t)} \cos\left[\phi_{\rm O}(x, y, t) - \phi_{\rm R}(x, y, t)\right] ,$$
(9.1.42)

where $U_{\rm R}(x, y, t)$, $U_{\rm O}(x, y, t)$, $\phi_{\rm R}(x, y, t)$, and $\phi_{\rm O}(x, y, t)$ are the complex amplitudes and phases of the reference and object beams, respectively, and $I_{\rm R}$, $I_{\rm O}$ are the intensities of the reference and object beams. The term $\cos [\phi_{\rm O}(x, y, t) - \phi_{\rm R}(x, y, t)]$ denotes the modulation of the intensity of the combined speckle pattern. This term produces the resulting overall intensity variation in a transverse plane (x, y) at a given time for a time-varying signal. Equation (9.1.41) can be modified similarly to classical interferometry (Sect. 9.1.3, see (9.1.27)):

$$I(x, y, t) = I_{\rm m}(x, y, t) \left\{ 1 + V \cos \left[\Delta \phi(x, y, t) \right] \right\} , \qquad (9.1.43)$$

where $I_{\rm m}(x, y, t)$ is the bias of the interference pattern and is equal to $I_{\rm R}(x, y, t) + I_{\rm O}(x, y, t)$, $\Delta \phi(x, y, t) = \phi_{\rm O}(x, y, t) - \phi_{\rm R}(x, y, t)$, and the so-called visibility is

$$V = \frac{2\sqrt{I_{\rm R}(x, y, t)I_{\rm O}(x, y, t)}}{I_{\rm R}(x, y, t) + I_{\rm O}(x, y, t)}$$

Temperature-, mechanical-, or chemical perturbation will result in optical path length differences introduced between the two beams. These induced changes lead to a new speckle field to be generated whose characteristic statistical property is the same except for changes introduced by the perturbation. It was shown that in conventional interferometry an interferogram is observed, however, in speckle interferometry, the information is not directly revealed in one single state of the superposing waves. To extract useful information from speckle interferometers we need (as in holographic interferometry or digital holographic interferometry, as will be shown in the next section) waves from a second state of the object. Comparing the two states of the superposing waves the information corresponding to the phase is revealed in form of fringes, called correlation fringes. Consequently, two states of the object can either be added or subtracted based on how the speckle patterns are recorded. Since the added intensities lead to fringes with low visibility, they are rarely used. Therefore the subtracted intensity of the two states will be given here only, namely:

$$I_2(x, y, t) - I_1(x, y, t) = I_M \left| \cos\left(\Delta\phi_2(x, y, t)\right) - \cos\left(\Delta\phi_1(x, y, t)\right) \right| , \qquad (9.1.44)$$

where $I_{\rm M}$ is the modulation of the combined speckle field. Here we have made the important assumption that only the phase has varied and that the individual speckle has not changed in position or shape. The above equation can be rewritten as

$$|I_{2}(x, y, t) - I_{1}(x, y, t)| = I_{M} \left| \sin \left(\frac{\Delta \phi_{2}(x, y, t) + \Delta \phi_{1}(x, y, t)}{2} \right) \right| \left| \sin \left(\frac{\Delta \phi_{2}(x, y, t) - \Delta \phi_{1}(x, y, t)}{2} \right) \right| .$$
(9.1.45)

Correlation fringes are obtained as a square of the difference and can be shown to be equal to

$$\left\langle \left| I_2(x,y,t) - I_1(x,y,t) \right|^2 \right\rangle = 2 \left\langle I_M^2 \right\rangle \sin^2 \left(\frac{\Delta \phi_{12}(x,y,t)}{2} \right) ,$$
 (9.1.46)

where $\Delta \phi_{12}(x, y, t) = \Delta \phi_2(x, y, t) - \Delta \phi_1(x, y, t)$. Correlation fringes are obtained as a variation in intensity in the (x, y) plane and are related to the phase difference between the initial and the final states of the object. The fringes are modulated by speckles and even in the bright region one can observe dark regions caused by low modulation on the detector.

It should be noted that the speckle size discussed above needs to be adapted to the pixel size of the detector (e.g. a CCD), which can be achieved easily by choosing the aperture of the image-forming lens system as has been discussed in Sect. 9.1.7.1. Larger speckles reduce the spatial resolution; smaller speckles reduce the visibility because there would be more than one speckle in the pixel.

Variation of the incidence angle of the illumination or the wavelength or in special cases the refractive index of the object-surrounding medium between the two recordings leads to contour fringes for the object shape measurement [90Joe, 92Rod, 94Zou, 97Her, 01Ras2]. To obtain the derivative of surface movements, speckle shearing interferometry is a powerful tool for strain analysis [74Hun, 86Joe, 93Pry, 97Ras1, 98Ste].

Speckle interferometry has also been applied for vibration analysis, either in the time average or by stroboscopic illumination. Alternatively, illumination with short pulses, with pulse widths of 20 ns and pulse separations down to μ s, is very useful for industrial applications. Two or multiple pulses can be applied which reduces the analysis of vibrating objects to two or multiple recordings of the speckle patterns [84Lok, 87Tyr, 93Lok, 93Ped].

Furthermore, the direction of illumination and observation determines the measurement sensitivity as well as the type of displacement the system will be sensitive to. Several novel techniques have been devised to visualize correlation fringes in speckle interferometry.

Figure 9.1.36 shows a result where the double-pulse speckle technique has been used for the measurement of vibrations.



Fig. 9.1.36. The intensity pattern obtained by subtraction of two specklegrams of a vibrating cylindrical object. The frequency of vibration was 778 Hz and the pulse separation 200 μ s.

9.1.7.4 Phase analysis in speckle interferometry

9.1.7.4.1 Phase analysis by phase stepping

The phase analysis occurs in the same way as in phase-stepping interferometry, a number of identical phase-stepped interference fringes are analyzed (see Sect. 9.1.3.2). A minimum of three measurements is required to solve for three unknowns in the interference equation. Typical phase steps introduced are 90 degrees, occasionally 120 degrees; in addition the analysis of more than three interferograms reduces possible errors occurring by the phase shifts introduced [85Cre, 90Ker, 93Kau] as in classical interferometry discussed in Sect. 9.1.3. Frequently the speckle fringes have to be processed in order to increase the visibility to achieve good phase plots. The accuracy of the method is limited by speckle distribution and speckle decorrelation. For the phase analysis in speckle interferometry we assume that the phase does not change appreciably over the time the three or four specklegrams are recorded. Other techniques like temporal phase steps or differentiating two interferograms that are phase-shifted by 90 degrees or the use of look-up tables are possible but not frequently used.

Several authors have discussed the phase-stepping technique for speckle pattern interferometry [85Cre, 90Ker, 93Kau]. To overcome the effect of random fluctuations in the speckle amplitude usually spatial and local averaging techniques are used to smooth the data. In this smoothing process a weighted averaging kernel K(m, n) is used before the phase at a given pixel is evaluated. Wavelet transforms have been used to do signal processing on speckle interferometric fringes to reduce speckle noise by several authors [01Kum]. Alternative to this method is to perform temporal averaging of the speckles by collecting an independent series of identical interferograms except for the speckles, hence they are averaged by the summation.

9.1.7.4.2 Phase analysis by spatial phase shifting

In spatial phase shifting a known tilt is introduced by tilting the reference plane wave, for instance, in such way that the adjacent pixels on the CCD have an inherent phase difference of $\pi/2$. By analyzing the information on three or four adjacent pixels the phase at a given pixel can be evaluated. This procedure requires only one specklegram or one frame of the speckle pattern to extract the phase. The price to be paid is the loss in spatial resolution because the speckles have to extend over at least three CCD pixels if three phase-shifted specklegrams are used [97Bot].

9.1.7.4.3 Analysis of vibrating objects

Vibrating objects can be analyzed in a similar way as discussed when double-pulse technique is applied. In the case of vibrating objects where time-averaging techniques are used the fringes obtained are not cosinusoidal anymore, but lead to a zeroth-order Bessel function. Therefore, the analysis is not straightforward anymore. However, it is possible to make quantitative analysis using temporal phase stepping with time-averaged fringes. The speckle modulation is obtained by using the four-phase-step method [98Gre, 99Moo, 01Moo].

9.1.7.5 Temporal speckle pattern interferometry (TSPI)

For the analysis of dynamic phenomena the processing of the interference signals can be carried out in the temporal domain, instead of the spatial. In temporal phase evaluation procedures a series of phase maps are obtained as the object is gradually deformed developing complete information of the object deformation. In Temporal Speckle Pattern Interferometry (TSPI) the interference signal is analyzed in the temporal domain instead of the spatial [93Hun, 97Tiz2, 98Joe1, 98Joe2, 99Joe, 01Tiz]. If the temporal development of the interference signal at each pixel is investigated, the signal noise can be reduced, the calculation procedure is simplified, the accuracy improved, and the measuring range increased.

In TSPI, a point on the object is observed over time as the object is being continuously deformed or rotated or the wavelength changed continuously (Fig. 9.1.37). A bright speckle turns dark by a phase change of π .

The intensity modulation of the speckles thus obtained provides the temporal evolution related to the movement. By analyzing the time-dependent signal, object deformations ranging from few microns to few hundreds of microns or the shape surface from few hundred micrometers up to few centimeters can be measured.

A typical signal obtained is shown in Fig. 9.1.38. By Fourier transformation of the temporal



Fig. 9.1.37. Temporal evolution of a speckle on a CCD pixel.



Fig. 9.1.38. Temporal intensity variation in a speckle.



Fig. 9.1.39. Fourier spectrum of a temporal-varying speckle on a pixel.





Fig. 9.1.40. Reconstruction of a side band of the spectrum in Fig. 9.1.39.

Fig. 9.1.41. Arrangement for recording of temporal speckle patterns for measuring shape and out-ofplane movements.

signal recorded in the series of frames, two bands are obtained on either side of a central peak in the Fourier spectrum as shown in Fig. 9.1.39 and one side band of the spectrum is filtered using a band-pass filter and shifted back to the origin. An inverse Fourier transform of the filtered spectrum portion leads to the raw phase that is then unwrapped to get the total phase at a given point (Fig. 9.1.40). The carrier frequency generated due to the object displacement or other changes should be large enough to separate the side bands from the central peak.

The object movement or deformation can be extracted using the Fourier-transformation method. This basic principle of TSPI can be extended to measure the object shape using variable wavelength or dual beam illumination, for in-plane motion detection [98Joe2]. Furthermore, the slope of the object deformation can be obtained by Temporal Fourier-Transform Speckle Pattern Shearing Interferometry (TSPSI) [97Ras1].

A TSPI arrangement is shown schematically in Fig. 9.1.41. The arrangement shown is basically a Twyman–Green set-up. One beam, after the beam splitter, illuminates a diffuse object and the other is reflected from a mirror forming the reference beam. The intensity of the reference beam is adjusted by introducing a polarizing system in order to adapt its brightness with respect to the object beam. The reference arm and the object arm are set to have arbitrary path lengths. A telecentric system is used to focus the object onto the sensor of a high-speed Charge-Coupled-Device (CCD) camera. In addition, the telecentric system insures the propagation axis of the beam from the mirror and the object to be collinear. It can be used for measuring the shape by varying the wavelength during the recording. With the same set-up out-of-plane deformation measurements can be carried out where the laser wavelength is kept constant and the time-varying optical path change due to object movement or time-varying out-of-plane deformation is studied. For deformation measurements a He-Ne laser beam is expanded with a spatial filtering set-up and then collimated. The intensity of the interference signal observed at the sensor plane can be expressed as (see (9.1.43))

$$I(x, y, t) = I_0(x, y) \left\{ 1 + V(x, y) \cos \left[\phi_0(x, y) + \phi(x, y, t) \right] \right\},$$
(9.1.47)

where I_0 is the mean intensity, V the visibility, $\phi_0(x, y)$ is the random phase of a speckle, and $\phi(x, y, t)$ is the position and time-dependent phase change introduced when an out-of-plane arrangement is used, see (9.1.50). For the shape measurement the height variation is measured as will be discussed in Sect. 9.1.7.5.1.

9.1.7.5.1 Shape measurement with TSPI by using time-varying wavelength change

For shape measurements either the wavelength is changed during exposure or the object is rotated in order to obtain a time-dependent phase variation leading to the shape information [98Joe1]. The wavelength change during exposure will be discussed only. Because the phase variation is analyzed, no mode hopping should occur during the wavelength change. The time-varying intensity due to the wavelength variation can be written in the same way as in (9.1.47), namely:

$$I(x, y, \lambda) = I_0(x, y) \{ 1 + V(x, y) \cos [\phi_0(x, y) + \phi(x, y, \lambda)] \}, \qquad (9.1.48)$$

where $\phi(x, y, \lambda)$ is the phase variation due to the optical path difference in air, by assuming n = 1.0. For normal incidence of the illumination onto the object

$$\phi(x,y,\lambda) = \frac{4\pi h(x,y)}{\lambda(t)} , \qquad (9.1.49)$$

where h(x, y) is the height to be measured. The time-dependent wavelength variation $\lambda(t)$ will be discussed below.

9.1.7.5.2 Laser diode with external cavity for wavelength change

For the time-varying wavelength the variation of current or temperature in a semiconductor laser leads to a very small wavelength change without mode hopping ($\Delta \lambda \leq 0.3$ nm). To extend the range a laser diode with external cavity was chosen. The arrangements according to Littman/Metcalf and Littrow were investigated [01Tiz]. Both use a reflecting mirror to select the wavelength. The Littman/Metcalf set-up was chosen as shown in Fig. 9.1.42.

The grating diffracts the collimated wave, the -1st diffractive order is returned to the laser cavity by the wavelength-selecting mirror arranged to be in auto collimation, whereas the zeroth order is directed to the speckle interferometer. An anti-reflex coating is needed for the laser-diode exit mirror in order to activate the external cavity system. To avoid mode hopping care needs to be taken with respect to the selection of the wavelength and corresponding cavity length variation by choosing the exact position of the rotation point.

The method can be applied for shape measurements of polished as well as rough surfaces. As an example to demonstrate the principle a staircase made of steel sheets of 0.2 mm thickness was analyzed. The result is shown in Fig. 9.1.43.



Fig. 9.1.42. Laser diode (LD) with external cavity according to Littman/Metcalf.

Fig. 9.1.43. TSPI-shape measurement of a staircase with continuous wavelength variation.

9.1.7.5.3 Application of TSPI for deformation measurement

Similar expressions can be derived for in-plane movement, object deformation- or the shape measurement of the object. The relationship for $\phi(x, y, t)$ depends on the set-up used [97Tiz2, 98Joe1, 98Joe2, 99Joe, 01Tiz]. For the out-of-plane arrangement we obtain for the phase:

$$\phi(x,y,t) = \frac{2\pi W(x,y,t)(1+\cos\theta)}{\lambda} , \qquad (9.1.50)$$

where θ is the angle of incidence of the object illumination, λ the wavelength, and W(x, y, t) the deformation to be measured. The instantaneous temporal frequency f_i of the signal observed at a given point of the image of the object during t seconds of recording is

$$f_{\rm i}(x,y,t) = \frac{1}{2\pi\lambda} \frac{\partial\phi(x,y,t)}{\partial t} . \tag{9.1.51}$$

Thus the temporal frequency of the signal observed is dependent on the object deformation or its slope or the object height. A large number of frames of the speckle pattern is recorded sequentially as the object is being deformed. For each pixel we observe the speckle intensity variation over time, providing the pixel history. From the angle of the complex values of the resultant transform, the phase information is extracted. The initial phase is eliminated if the difference between the first frame and the last frame is taken. This procedure of obtaining the phase does not demand the displacement to be linear. For non-linear deformations the side bands broaden, demanding that the width of the band-pass filter used must be changed appropriately. This phase is then unwrapped as in phase-shifting interferometry thereby giving the total phase at that point. Obtaining the phase at all the points of the object provides the map of the time-dependent phase of the object deformation. The instantaneous angular frequency or the linear velocity generated during deformation of the object point is also determined. From these extrapolations the 3D plot of the object deformation is extracted. For in-plane measurement a two-beam illumination is used [98Joe2].

9.1.7.5.4 Deformation measurements by TSPI and digital holography, a comparison

For an experimental comparison an arrangement for digital holographic interferometry and TSPI was realized. The set-up is shown in Fig. 9.1.44. Object and reference beams were recorded on a high-speed CCD camera. For a direct comparison the exposures were the same but the analysis of the information was different. For the reason of simple data handling 200×200 image points



Fig. 9.1.45. A comparison of the results of the deformation measurement with TSPI and digital holography. (a) shows the result of TSPI and (c) a cross-section whereas (b) shows the result obtained by digital holography and (d) a cross-section.

were analyzed only and 1024 images were recorded for TSPI. For digital holography only every 25th image was analyzed leading to a deformation of 2 μ m in the center. The results are shown in Fig. 9.1.45, where Fig. 9.1.45a shows the result of the TSPI analysis. It should be noted that in addition to the deformation the movement of the whole object during the measurements is observed with TSPI which is not the case in classical digital holographic interferometry as shown in Fig. 9.1.45b. Digital holographic interferometry will be discussed in Sect. 9.1.8. In addition a

cross-section of the deformations is shown for each method in turn in Fig. 9.1.45c, d. The body movement of 52 μ m is seen clearly in Fig. 9.1.45c, measured with TSPI, which cannot be measured with a classical double-exposure digital holographic interferometry method. To obtain the sign information preliminary information can be used, or a more elaborate heterodyne technique can be used [01Tiz]. To measure deformation with temporal speckle patterns the temporal Hilbert-Transform (HT) method can also be applied for the phase analysis [03Mad].

9.1.7.5.5 Vibration measurement with TSPI

The evaluation of the time-dependent speckle modulation due to optical path-difference changes can be extended to vibration measurement [01Tiz]. For a sinusoidal signal evaluation with varying velocity: from zero at its reversal point, to maximum speed. For both cases we assume a harmonic vibration. For evaluation, the distance of two adjacent zero-velocity points is determined, giving half the mechanical time base.

Let us consider an experimental set-up like a Fizeau interferometer, see Fig. 9.1.46, where the object is harmonically vibrating. The periodical movement is expressed by:

$$z(t) = \hat{z} \cdot \sin(2\pi f_{\text{mec}} t - \varphi_0) , \qquad (9.1.52)$$

where f_{mec} is the frequency of mechanical oscillation and \hat{z} its amplitude. The initial phase is represented by φ_0 and can be neglected. The instantaneous velocity v(t) is given by the derivative of the mechanical path z(t):

$$v(t) = \frac{dz}{dt} = 2\pi f_{\rm mec} \,\hat{z} \cdot \cos\left(2\pi f_{\rm mec} \,t\right) \,. \tag{9.1.53}$$

Using (9.1.48), the intensity at the detector can be written as

$$I(t) = I_0 \left(1 + V \cos\left\{ \left(\frac{4\pi \hat{z}}{\lambda}\right) \sin\left(2\pi f_{\text{mec}} t\right) - \varphi_0 \right\} \right) \,. \tag{9.1.54}$$

The intensity at the detector is phase-modulated, the phase is

$$\varphi(t) = \frac{4\pi}{\lambda} z(t) , \qquad (9.1.55)$$

written as frequency modulation of v(t):

$$f(t) = \frac{2v(t)}{\lambda} = \frac{4\pi f_{\text{mec}}}{\lambda} \hat{z} \cdot \cos\left(2\pi f_{\text{mec}}t\right) , \qquad (9.1.56)$$

where f(t) is the instantaneous modulation frequency. In the analysis a frequency demodulation or a phase demodulation can be used in order to obtain the velocity or the optical path variation. FM-demodulation is done by FM/AM conversion for instance. Therefore the signal is differentiated and then low-pass-filtered in order to determine the envelope. An alternative demodulation method is the Fourier-based phase demodulation.

The experimental set-up is shown in Fig. 9.1.46. A fiber-optic arrangement is shown where the reference is reflected from the not coated fiber tip close to the object. For the experiment a loudspeaker was used as object with a sinusoidal frequency synthesizer at a frequency of 200 Hz. The fiber tip was put close to the vibrating membrane of the loudspeaker. The laser power of an InGaAs laser diode was 20 μ W. The measuring time was 25 ms corresponding to 5000 samples.

Figure 9.1.47 shows the measurement of a vibration sequence recorded with the Fizeau set-up. The detector signal represents the frequency-encoded movement of the loudspeaker. The Fourierbased phase demodulation reconstructs the mechanical velocity which is shown in Fig. 9.1.48. The comparison to a high-precision vibrometer based on Doppler frequency measurement shows a good conformity including also higher harmonics. The noise is due to the small laser power and can be reduced when applying a band-pass or a special low-noise amplifier.



Fig. 9.1.47. Detector signal at an object point of the loudspeaker at 200 Hz.

Fig. 9.1.48. Reconstructed mechanical velocity of the vibration analysis of the loudspeaker working at 200 Hz obtained with the Fizeau arrangement (upper part). For comparison a measurement with a highperformance vibrometer based on heterodyne technology is seen on the lower part (dashed line).

9.1.8 Holographic interferometry

9.1.8.1 Principle of holography

Holography is an interferometric technique for recording and reconstructing wavefronts. It was invented by Gabor in 1948 as an attempt to increase the resolution in electron microscopy [48Gab]. Gabor demonstrated that if a suitable coherent reference wave is present simultaneously with the light scattered from an object, the information about both the amplitude and phase of the object wave could be recorded, although the recording media only register the light intensity. The concept of holography was developed first as in-line arrangement. Gabor's holography was limited to film transparencies using a mercury arc lamp as the light source. His holograms contained distortions and an extraneous twin image.

For the next ten years, holography techniques and applications did not develop further until the invention of the laser 1960, whose coherent light was ideal for making holograms. By using a laser source, Leith and Upatnieks developed in 1962 the off-axis reference beam method which is most often used today [64Lei]. It permitted the making of holograms of solid objects by reflected light. Basically in holography, two coherent wavefronts, a reference and an object wave are recorded as an interference pattern as shown in Fig. 9.1.49. In classical holography the interference patterns are stored on high-resolution photographic emulsions, occasionally on photothermoplastic or photorefractive material. For the reconstruction a coherent wave, usually a laser, is used. Illuminating the hologram with a laser the physical reconstruction of the recorded wave field is performed as shown schematically in Fig. 9.1.50. By reconstructing the hologram the stored phase information



Fig. 9.1.50. Reconstruction obtained by illuminating the hologram with the original reference wave. The real (a) and the virtual (b) images are reconstructed. H is the hologram.

contains the information on the third dimension, the depth. Therefore a quasi three-dimensional reconstruction appears.

Leith's and Upatniek's off-axis arrangement [64Lei, 71Col, 74Cat, 79Sch, 79Ves, 89Jon, 96Kre] leads to a spatial separation of different reconstructions because the reconstructed waves and the zeroth order propagate in different directions. Two images are reconstructed to obtain a so-called virtual image ($U_{\rm O}$) and a real image ($U_{\rm O}^*$) as indicated schematically in Fig. 9.1.50. The virtual image is usually considered for holographic interferometry. Its position and size will be the same as for the original object when reference and reconstructing waves are the same with respect to geometry and wavelength.

There are different configurations for recording and reconstruction of the holograms [71Col, 79Sch, 79Ves, 96Kre], namely:

- Fresnel hologram,
- Fourier hologram,
- image-plane hologram.

A Fresnel hologram is the more general case where the recording is neither in the image nor Fourier plane of the object. Fourier holograms are recorded in the Fourier plane (focal plane of a lens). If the origin of the spherical reference wave is located in the plane of the object we call the hologram recorded a quasi-Fourier hologram. In image-plane holography, the image of the object is formed practically in the recording plane. Image-plane holograms are very convenient for digital holographic interferometry and speckle interferometry. An application of holography for optical engineering is holographic interferometry. The principles of holographic interferometry (Sect. 9.1.8.2) will be discussed mainly as introduction to the digital holographic interferometry (Sect. 9.1.8.4).

9.1.8.2 Principle of holographic interferometry

Holographic interferometry [79Sch, 79Ves, 96Kre] allows a comparison of different states of the object at different times. This technique uses the ability of the hologram to store two or more complex wavefronts by performing multiple exposures. The subsequent interference of these wave fields when they are reconstructed together produce a fringe pattern, from where the information of the deformations or change of refractive index of the object between the exposures can be obtained. There are different techniques to be used, namely:

- double-exposure technique,
- real-time technique,
- time-average technique.

In double-exposure holography two exposures are recorded (before and after deformation) on the storage material, a high-resolution photographic emulsion, for instance, and reconstructed together by illuminating with a laser.

In real-time holography the first hologram is recorded, the holographic plate is developed in its position or returned after development in exactly the same position. Because the precise repositioning is difficult, the holograms are frequently recorded on photothermoplastic or on photorefractive material. When the stored hologram is compared with the wavefront coming from the deformed object interference fringes, corresponding to the object deformation can be seen.

In the time-average technique [79Sch, 79Ves, 88Ste, 96Kre] the harmonic vibrations are recorded in the time average, i.e. during many vibration cycles.

We will concentrate on the double-exposure technique which is the most used for industrial applications. Two exposures are recorded on the hologram; one before and one after the deformation, for instance. The two recorded holograms are reconstructed together by illuminating the recorded holograms with a laser. In addition to the two reconstructed images of the object an interference pattern is superposed containing the information of the variation of the object between the two exposures. It should be noted that the deformation needs to be very small, in the range of a few micrometers. Hence the two reconstructed images are practically superposed. Furthermore, the virtual reconstructed image is usually analyzed as seen schematically in Fig. 9.1.51, but $U_{\rm O}$ and $\overline{U}_{\rm O}$ have a phase difference due to the deformation between the two exposures.

For a short analysis may $U_{\rm R}$ be the reference wave and $U_{\rm O}$ the object wave before and $\overline{U}_{\rm O}$ after deformation. To simplify the writing the coordinate dependences of $U_{\rm R}$, $U_{\rm O}$ and $\overline{U}_{\rm O}$ are omitted.

Furthermore, may $U_{\rm O} = |U_{\rm O}| \exp(-i\varphi)$ be the complex object wave before and $\overline{U}_{\rm O} = |\overline{U}_{\rm O}| \exp(-i\overline{\varphi})$ after deformation. We assume $|\overline{U}_{\rm O}| \approx |U_{\rm O}|$.

The recording of a single hologram can be written in the form

$$I_{1} = |U_{\rm R} + U_{\rm O}|^{2} = |U_{\rm R}|^{2} + |U_{\rm O}|^{2} + U_{\rm R}^{*}U_{\rm O} + U_{\rm R}U_{\rm O}^{*}$$
(9.1.57)

(where * means complex-conjugate), and the recording of a double-exposed hologram is



Reconstruction wave

Landolt-Börnstein New Series VIII/1A2

j

Fig. 9.1.51. Principle of a reconstruction of a doubleexposed hologram. After the hologram development and reconstruction by illuminating the hologram having a transmittance T with $U_{\rm R}$ we get

$$U_{\rm R}T = \text{const.} + \beta \tau \left[\left| U_{\rm R} \right|^2 \left(U_{\rm O} + \overline{U}_{\rm O} \right) + U_{\rm R}U_{\rm R} \left(U_{\rm O}^* + \overline{U}_{\rm O}^* \right) \right] \,. \tag{9.1.59}$$

 β and τ are parameters depending on the recording and developing process. τ is the exposure time and β the parameter of the photographic emulsion for instance. Due to the oblique incidence of the reference wave the different terms in (9.1.58) are spatially separated in the reconstructions. The intensity of the reconstructed wave of interest (virtual reconstruction) is given by

$$I_{\rm V} = \left| \beta \tau \left| U_{\rm R} \right|^2 \left| U_{\rm O} \right| \left(e^{-i\varphi} + e^{-i\overline{\varphi}} \right) \right|^2 = 2 \left(\beta \tau \right)^2 \left| U_{\rm R} \right|^4 \left| U_{\rm O} \right|^2 \left[1 + \cos \left(\varphi - \overline{\varphi} \right) \right] \,.$$
(9.1.60)

The additional phase difference $\varphi - \overline{\varphi}$ accounts for the deformation between the two exposures, leading to an interference pattern on the reconstructed images. From the analysis of the interference fringes and by knowing the geometrical data of the set-up and the object shape the 3D object deformation can be computed [71Col, 74Cat, 79Sch, 79Ves, 96Kre].

9.1.8.3 Digital holography

9.1.8.3.1 Principle of digital holography

In classical holography the interference patterns of the reference and object waves are stored on high-resolution photographic emulsions: for holographic interferometry occasionally on photothermoplastic or photorefractive material. For the reconstruction a coherent wave, usually a laser, is used.

The holographic process is time-consuming, particularly with respect to the development of the photographic film, making the method unsuited for industrial environments. In recent years optoelectronic recording devices like CCD or CMOS cameras have become an attractive alternative to photographic film, photothermoplastic or photorefractive crystals as recording media in holography. The time-consuming wet processing of photographic film is avoided and the physical reconstruction of holograms can be replaced by numerical reconstruction using a computer, normally an ordinary PC. One disadvantage of using CCD sensors for hologram recording is, however, the relatively low spatial resolution, compared to that of photographic material. The limitation in spatial resolution is a problem but the technology is improving further. The enormous progress in the performance of computers and increase in the computer capacity is very favorable for digital holography.

The problem of recording holograms on low-resolution media has been discussed in many books and publications since people first begun to record holograms [71Col, 74Cat, 94Sch, 95Ped, 96Sch]. Particle size and position measurements were based on digital reconstruction [97Ada]. Different holographic recording arrangement as in classical holography can be applied. The off-axis hologram arrangement using a point source (P) as a reference can be used (Fig. 9.1.49). The intensity of the interference between the reference ($U_{\rm R}$) and the object ($U_{\rm O}$) waves is given by (9.1.57). The maximum spatial frequency that can be recorded is limited by the resolution of the recording medium. To record a hologram of the entire object, the resolution must be sufficient to resolve the fringes formed by the reference wave and the wave from the object point farthest from the reference point. If $\theta_{\rm max}$ is the maximal angle between the object and the reference beam (Fig. 9.1.49), the maximal spatial frequency $f_{\rm max}$ in the hologram is:

$$f_{\max} = \frac{2}{\lambda} \sin\left[\frac{\theta_{\max}}{2}\right] , \qquad (9.1.61)$$

where λ is the wavelength of the laser light used for the recording. Consider that the intensity described by (9.1.57) is recorded on a two-dimensional array of sensors ($M \times N$ cells) of cell dimension $\Delta x \times \Delta y$. For the recording, in most cases a CCD sensor is used, but technology is progressing and a CMOS sensor may be used in future. The discretized intensity recorded by the sensor can thus be written in the form $I_{\rm H}(m\Delta x, n\Delta y)$, where n and m are integer numbers. For the most common case of equidistant cells, we have $\Delta x = \Delta y = \Delta$. From the sampling theorem (see e.g. [96Goo]) it is necessary to have at least two sampled points for each fringe, therefore the maximum spatial frequency which can be recorded is $f_{\rm max} = 1/(2\Delta)$. The interference fringes obtained between the light coming from the object and the reference should have a period greater than 2Δ .

There are different types of digital holograms as in classical holography [96Kre], namely Fresnel, Fourier, and image-plane holograms. Furthermore, there are in-line and off-axis holograms, as in classical holography. For off-axis holograms the separation of the reconstructed images is simpler even though the reconstruction occurs by computer and not physically as described above. We will concentrate now on off-axis digital holography. A Fresnel hologram is the more general case where the recording is neither in the image nor Fourier plane of the object. If the origin of a spherical reference wave is located in the plane of the object we call the hologram a quasi-Fourier hologram. Image-plane holograms are very convenient for digital holographic interferometry. The recording is therefore very similar to ESPI recording (Sect. 9.1.7).

The electronic recording device is considered as a two-dimensional (thin) hologram (Fig. 9.1.49). It is possible that electronic digital devices could be developed in the future where even volume holograms can be recorded and reconstructed digitally. We will restrict our discussion to holograms that can be digitized with currently available devices.

9.1.8.3.2 Configurations for recording and reconstruction of digital holograms

Holograms may be classified according to the geometry used for the recording and reconstruction. Quasi-Fourier, Fresnel, image-plane, and in-line and off-axis holograms will be described together with the digital reconstruction. All these different configurations can be used for recording on an electronic device.

In classical holography the physical reconstruction process is performed by illumination of the developed hologram with the reference wave. In digital holography this process is replaced by computer simulation. Before considering the different recording and reconstruction geometries, we will introduce some mathematical definitions that are helpful for the description of digital reconstruction. From the diffraction theory we know that if the amplitude and phase of a wavefront are known at a plane (x_1, y_1) , it is possible to calculate the propagation of such a wavefront in the plane (x_2, y_2) , see Fig. 9.1.52. The recorded intensity is multiplied with a numerical representation of the reference wave followed by numerical calculation of the resulting diffraction field at a certain distance d in z-direction from the hologram plane where the reconstructions, which may be deduced by solving the wave equation. If we consider the Rayleigh–Sommerfeld diffraction integral



Fig. 9.1.52. Geometry for Rayleigh–Sommerfeld diffraction.

[96Goo], between two planes the wave field at a distance d is given by:

$$U_2(x_2, y_2, d) = \frac{\mathrm{i}\,A}{\lambda} \iint U_1(x_1, y_1) \frac{\exp(\mathrm{i}\,kr_{12})}{r_{12}} \,\cos\left(\theta\right) \,\mathrm{d}x_1 \,\mathrm{d}y_1 \,, \tag{9.1.62}$$

where $r_{12} = \sqrt{d^2 + (x_2 - x_1)^2 + (y_2 - y_1)^2}$ and $k = 2\pi/\lambda$; A is the amplitude of the incident wave and $\cos(\theta)$ is the obliquity factor. It should be mentioned that a slightly different assumption in the boundary conditions lead to different obliquity factors. If we consider e.g. the Kirchhoff–Fresnel approximation the obliquity factor is given by $(1 + \cos \theta)/2$ [96Goo]. For small θ there is only a small difference between the two. In the following the obliquity factor will be neglected.

If the distance d between the object and hologram plane or the hologram and the image plane is large compared to $(x_1 - x_2)$ and $(y_1 - y_2)$ the binominal expansion of the square root up to the third term yields:

$$r_{12} \approx d \left[1 + \frac{1}{2} \left(\frac{x_1 - x_2}{d} \right)^2 + \frac{1}{2} \left(\frac{y_1 - y_2}{d} \right)^2 \right] \,. \tag{9.1.63}$$

Equations (9.1.62) and (9.1.63) give the Fresnel approximation of the diffraction integral:

$$U_2(x_2, y_2, d) = \frac{A e^{i k d}}{i \lambda d} e^{i \frac{k}{2d} \left(x_2^2 + y_2^2\right)} \iint U_1(x_1, y_1) e^{i \frac{k}{2d} \left(x_1^2 + y_1^2\right)} e^{-i \frac{k}{d} \left(x_1 x_2 + y_1 y_2\right)} dx_1 dy_1 .$$
(9.1.64)

Equation (9.1.64) has the form of a convolution integral, and we know from the Fourier theory that a convolution of two functions in the space domain is equivalent to multiplying their individual Fourier transforms. By inverse Fourier transforming we go back to the space domain. This property is very useful since it allows the computation of the wavefront propagation in a short time. If we consider a digitized wave field from which we would like to compute the propagation (e.g. formed by 1000×1000 points) it would take a very long time to compute the convolution integral described by (9.1.64), convolutions are in fact time-consuming. The Fourier transform of a discrete function can be calculated in a short time by using the FFT (Fast-Fourier-Transform) algorithm and this allows a dramatically reduced computation time of the discretized wave fields.

In order to simplify the notation we introduce a function approximating a spherical wavefront diverging from a point located at distance d from the origin of the coordinate system (Fresnel approximation)

$$h(x_1, y_1, d) = \exp\left[\frac{\mathrm{i}\,k}{2d}\left(x_1^2 + y_1^2\right)\right] \,. \tag{9.1.65}$$

The Fresnel transform described by (9.1.64) can be seen as a Fourier transform of the object wave field multiplied by a quadratic factor and can be written in the more compact form

$$U_2 = \frac{A e^{i k d}}{i \lambda d} U_1 \otimes h , \qquad (9.1.66)$$

where \otimes denotes convolution. The multiplicative phase term $(A e^{i kd}) / (i \lambda d)$ will be omitted in the following calculations. To simplify the writing the coordinate dependence is omitted.

The wavefront in the hologram plane is related to the distribution in the object plane by the convolution of U_1 with the impulse response h. If we write U_0 as the object wave instead of U_1 (see (9.1.57)), the intensity recorded by the sensor may be written in the form

$$I_{\rm H} = |U_{\rm R}|^2 + |U_{\rm O} \otimes h|^2 + U_{\rm R} [U_{\rm O} \otimes h]^* + U_{\rm R}^* [U_{\rm O} \otimes h] .$$
(9.1.67)

This expression will be used in the calculations describing the digital reconstruction of the wavefronts. After this mathematical review of the formulae of wave propagation we may now consider some recording geometries and the corresponding reconstruction method.

We will begin our description with the lensless Fourier hologram, which is relatively simple because the digital reconstruction requires only the calculation of the Fourier transform of the recorded discrete intensity.

9.1.8.3.2.1 Lensless Fourier hologram

The arrangement used to record a lensless Fourier (or quasi-Fourier) hologram (these are the two names used in the literature) is shown in Fig. 9.1.53. Instead of recording the holograms on films as discussed before, they are recorded on a low-resolution electronic device. For the hologram recording a point source located close to the object is used as a reference. The spherical reference $U_{\rm R}$ used for the hologram recording can be approximated by the function h(x + a, y, d) (see (9.1.65)), where a denotes the offset of the reference point source, P = (-a, 0). By taking the Fourier transform of $I_{\rm H}$ we will get the reconstruction of four wavefronts.

In this notation the last term of (9.1.67) leads to

$$I_{\rm HP}(x_{\rm H}, y_{\rm H}) = h \left(x_{\rm H} + a, y_{\rm H}, -d \right) \left[U_{\rm O}(x_{\rm H}, y_{\rm H}) \otimes h(x_{\rm H}, y_{\rm H}, d) \right] \,. \tag{9.1.68}$$

The Fourier transform of the quadratic function $h(x_{\rm H}, y_{\rm H}, d)$ is still a quadratic function. After some manipulation, in particular by using the convolution theorem, the Fourier transform \mathcal{F} of $I_{\rm HP}(x_{\rm H}, y_{\rm H})$ takes the form [74Cat]

$$\mathcal{F}\{I_{\mathrm{H}P}(x_{\mathrm{H}}, y_{\mathrm{H}})\} = \lambda^2 d^2 \exp\{\mathrm{i}\,k\,a_{\mathrm{f}}^2\} \, U_{\mathrm{O}}(-x_{\mathrm{f}} - a_{\mathrm{f}}, -y_{\mathrm{f}})\,h\,(x_{\mathrm{f}}, y_{\mathrm{f}}, d) \,, \tag{9.1.69}$$

which represents an inverted image $U_{\rm O}$ of the object multiplied by a quadratic phase factor and shifted by $a_{\rm f}$. $x_{\rm f}$ and $y_{\rm f}$ denote the coordinates in the Fourier plane. Without loss of generality, we assume that the wavefront is reconstructed with unity magnification, i.e. $a_{\rm f} = a$. (Since the wavefront is not reconstructed physically we can change its size). By analogy, the Fourier transform of the third term of the intensity function will lead to the secondary reconstruction. Both the primary and secondary images are focused. The Fourier transform of a real positive function (the intensity) yields a Hermitian distribution in the Fourier plane, with its absolute value symmetrical with respect to the center.

Figure 9.1.54 shows a one-dimensional sketch of the hologram spectrum. The spectrum of the first term in (9.1.67) is, for a homogeneous wave, a delta function centered at the origin and does not disturb the reconstruction because its spatial extension is very small. A more fundamental limit is the spread of the central distribution due to the Fourier transform of $|U_0 \otimes h|^2$ which will



Fig. 9.1.53. Recording of an off-axis digital hologram.

Fig. 9.1.54. Spectrum of a hologram, the components are well separated.



Fig. 9.1.55. Digital reconstruction of a quasi-Fourier hologram.

represent a distribution having twice the extension of the reconstructed object. In order to avoid their superposition in the reconstructed images, the closest object point must be at least one object width from the reference. In practice this condition reduces the reconstructed object size almost by a factor of four.

Figure 9.1.55 shows an example of a reconstructed quasi-Fourier hologram of a small object $(3 \times 5 \text{ cm}^2)$. The object was located at the distance d = 140 cm from the sensor. The hologram was recorded on a 512×512 array and the Fourier transform calculated with an FFT algorithm implemented on a computer (PC). The bright area in the center is the superposition of the reference wave and a convolution term.

9.1.8.3.2.2 Fresnel hologram

A Fresnel hologram may be recorded by using the arrangement shown in Fig. 9.1.53 but we would like to point out that the reference does not need to be a point source located in the plane of the object. Simulating the diffraction of the reference wave used for the recording the digital reconstruction of the complex amplitude in a plane at a distance d from the hologram can be calculated. In digital holography we usually use a homogeneous plane or spherical wave for the recording which can be simulated easily. The wavefronts diffracted by the hologram are obtained by multiplying the intensity $I_{\rm H}$ recorded with the reference wave $U_{\rm R}$ and calculating the resulting diffraction. Mathematically we can write the wavefront in the plane (x, y) located at the distance d from the hologram in the form

$$U_{\rm r} = [U_{\rm R}I_{\rm H}] \otimes h . \tag{9.1.70}$$

As noted before, this calculation can be carried out using a Fourier-Transform (FT) algorithm. We obtain a reconstruction of the object wavefront without any additional quadratic phase term as for the quasi-Fourier hologram. It should be noted that the Fresnel approximation is valid only when the distance d is large with respect to the size of the recorded hologram. A digital reconstruction of the wavefront close to the hologram needs special considerations.

Recording a digital Fresnel hologram of a large object requires frequently a demagnification by means of an optical system. In this way the angle between the object and the reference beam is reduced. A positive lens [95Ped] or a negative lens [96Sch, 99Wag] can be helpful.

9.1.8.3.2.3 Image-plane hologram

For an image-plane digital hologram the object is projected onto the detector array by a lens system, as shown in Fig. 9.1.56. The arrangement is very similar to that one used for speckle pattern interferometry. Image-plane holography is very suitable to calculate the complex amplitude of a wavefront in the plane of the detector. In order to reduce the spatial frequency of the wavefront arriving on the sensor (increase the speckle size), an aperture (AP) has been inserted. The reference is a spherical wave originating from the point P located close to the aperture. A quasi-Fourier hologram of the aperture (with size and distance from the reference chosen in order to satisfy the sampling theorem) is recorded. By performing a Fourier transform of the intensity recorded with such an arrangement, we get two reconstructions (primary and secondary) of the reconstruction of the aperture together with the central convolution term. When the three terms are well separated (as in the case of Fig. 9.1.57), we can filter out the central term and one of the reconstructions of the aperture and keep the other. After this filtering operation an inverse Fourier transform is applied and the complex amplitude of the object wave field in the hologram plane, which is the image plane, is obtained.

It appears clearly from Fig. 9.1.57 that the most convenient form of the aperture is the rectangular one. This is the best compromise between the possibility of filtering and the possibility of using the maximum of the area. It is obvious that the aperture should be chosen as large as possible for two reasons: firstly for improved light collection, and secondly for collecting higher spatial frequencies leading to increased spatial resolution.

Unlike the quasi-Fourier and Fresnel holograms, the reconstructed image-plane hologram is localized in the image plane. This means that a certain area of the image-plane hologram contains only information about a certain part of the object.



Fig. 9.1.56. Arrangement for the recording of an image-plane hologram.



Fig. 9.1.57. Filtering procedure for obtaining the phase.

9.1.8.4 Digital holographic interferometry

9.1.8.4.1 Principle of digital holographic interferometry

The advantage of digital holography is the calculation of the phase of the reconstructed wavefront from the complex amplitude. The digital method allows not only the analysis of wavefronts without physical reconstruction but also a comparison of two or more wavefronts [76Fri, 85Tha, 94Sch, 95Ped, 96Goo, 96Ped, 96Sch, 97Ada, 97Ped1, 97Ped2, 97Ped3, 97Ped4, 98Ped, 98Seb, 99Ped1, 99Ped2, 99Ped3, 01Ped, 01Per, 01Sch, 01See, 02Kol].

Therefore, two digital holograms can be recorded for example without and with loading as in classical holographic interferometry and in speckle pattern interferometry. However, no physical reconstruction of the two stored holograms is required in digital holographic interferometry. We denote the two corresponding reconstructed wave fields without and with loading by $U_{\rm O}$ and $\overline{U}_{\rm O}$, respectively, as in Sect. 9.1.8.2. The wrapped phases $\phi_{\rm w}$ and $\overline{\phi}_{\rm w}$ of the reconstructed wavefronts can be obtained from the complex amplitudes by

$$\phi_{\rm w} = \arg\left(U_{\rm O}\right) \ , \tag{9.1.71}$$

$$\overline{\phi}_{\rm w} = \arg\left(\overline{U}_{\rm O}\right) , \qquad (9.1.72)$$

where arg denotes the argument of a complex amplitude. The dependence from the position (x, y, z) has been omitted for convenience. The phases ϕ_{w} and $\overline{\phi}_{w}$ vary between 0 and 2π . The difference between the wavefronts is obtained by the phase differences

$$\Delta \phi_{\rm w} = \overline{\phi}_{\rm w} - \phi_{\rm w} \,. \tag{9.1.73}$$

In order to have a phase difference in the interval from 0 to 2π it is necessary to add 2π if $\Delta \phi_{\rm w} < 0$. $\Delta \phi_{\rm w}$ is a wrapped phase map $(0 \le \Delta \phi_{\rm w} \le 2\pi)$ containing the information about the change of the wavefront between the two exposures.

When an object illuminated by coherent light is deformed, the phases of the reflected wavefront change as a function of the deformation [79Sch, 79Ves, 96Kre]. If we consider two states of the object (before and after deformation), the relation between the deformation vector \boldsymbol{L} and the change of the optical phase $\Delta \phi$ of the wavefront is given by

$$\Delta \phi = \frac{2\pi}{\lambda} \boldsymbol{L} \cdot \boldsymbol{s} , \qquad (9.1.74)$$

where s is the sensitivity vector given by the geometry of the set-up: $s = k_i - h$, where k_i and h are the unit vectors of illumination and observation, respectively. Digital holography permits the calculation of the phase difference modulo 2π of the phase difference $\Delta \phi_w$ between two wavefronts; the phase $\Delta \phi$ is obtained by unwrapping $\Delta \phi_w$. From (9.1.74) the projection of the deformation L on the sensitivity vector s is calculated from the phase $\Delta \phi$.

A few results of some applications of digital holographic interferometry will be presented. In particular the phase difference obtained by subtraction from different digital holograms recorded at different times will be shown leading to the measurement of deformation, vibration, or shape.

9.1.8.4.2 Digital holographic interferometry for dynamic deformations

9.1.8.4.2.1 Dimensional measurements

Figure 9.1.58 shows an interference pattern obtained by recording two digital image-plane holograms of a plane object (metal plate) submitted to a shock. A pulsed ruby laser with pulse width of 20 ns, which can emit two high-energy pulses separated by few microseconds has been used for the recording of the two holograms on a CCD. The laser pulses were fired shortly after the shock



Fig. 9.1.58. (a) Object submitted to a shock excitation. (b) Phase map obtained by double-pulse-exposed digital holography.

produced by a metal ball falling on the plate. After calculation and subtraction of the phases of the two reconstructed digital holograms, we get a phase map containing 2π discontinuities. Notice that the fringes are concentric around the point of the impact. A small defect was inserted in the plate and the shock wave propagating at the surface of the plate is disturbed by this defect. This example shows the principle for defect detection in mechanical parts with digital holographic interferometry [01Sch].

For this kind of investigations the holograms recorded by the CCD for each laser pulse must be separated. Two or even four separated holograms can be recorded on a CCD sensor with pulse separation in the microsecond range. For the recording of two digital holograms within few microseconds the procedure is the following. The first pulse is fired producing the electrical charges inside the photosensors corresponding to the recording of the first hologram. Since we record 2 separated holograms (and not a sum of 2 holograms) it is necessary to transfer the information of the first hologram before the second pulse is fired. In a commercially available CCD it is possible to transfer the image (or more correctly the electrical charges corresponding to the recorded image) from the photosensors to the shift register. The time necessary for the transfer is short (2...5 microseconds for a commercially available interline sensor) because the transfer is parallel from each photosensor to the adjacent register cell. After the charge transfer, the photosensors of the camera are ready for the new image that is recorded with the second pulse. The charges corresponding to the 2 holograms are transferred from the CCD into a computer and digitally evaluated later.

9.1.8.4.2.2 Digital holographic interferometry for deformation and vibration analysis of 3D objects

In some cases a three-dimensional analysis of the deformation or vibration is necessary. This can be obtained if the shape of the object under investigation and the deformation are known. For 3D vibration analysis, the complete determination of the vibration occurs by combination of three different measurements, namely:

- 1. 3D vibration of the surface obtained by digital holography,
- 2. shape measurement of the object,
- 3. 3D vibration on a point obtained using a laser Doppler vibrometer for calibration.

The system used for the measurement of the vibrations is schematically shown in Fig. 9.1.59. Light from a pulsed ruby laser is divided into two beams, one is coupled into an optical monomode



Fig. 9.1.59. Schematic arrangement for 3D deformation or vibration analysis of 3D objects.

fiber and serves as the reference beam and the other one is further divided into four beams used to illuminate the object from four different directions. Together with the reference beam, four different sensitivity vectors are formed.

Equation (9.1.74) gives us the deformation along one sensitivity vector. For the 3D case, the relation between the phase differences and the deformation L is

$$\Delta \phi_i = \frac{2\pi}{\lambda} \mathbf{L} \cdot \mathbf{s}_i . \tag{9.1.75}$$

The sensitivities s_i are given by the geometry of the set-up:

$$s_i = k_i - h$$
, $i = 1, 2, 3, \dots$, (9.1.76)

where k_i and h are the unit vectors of illumination and observation, respectively. The most convenient way to determine the deformation vector L is to decompose all vectors into their orthogonal components x, y, and z.

There are many known methods used for shape measurement, for instance: triangulation [97Hun], fringe projection, interferometric methods (holography, speckle interferometry). We will restrict our attention to interferometric methods based on digital holography [98Seb, 99Ped1, 99Ped2, 02Kol].

The two-wavelength method for contouring was used in the early years of holography [76Fri, 85Tha]. Two holograms of an object were recorded with two wavelengths (λ_1 and λ_2) on holographic film; for the reconstruction the hologram was illuminated by a laser using one wavelength. Contour fringes corresponding to the object shape were obtained on the reconstructed object. Digital holography can be used for shape measurement in an optical set-up where two digital holograms are recorded using two different wavelengths. There are different methods to change the wavelength of a laser e.g. a dye laser or a diode laser with external cavity discussed in Sect. 9.1.7.5.2 when only a small change of the wavelength is necessary. For very small wavelength change, tuning a laser diode by changing the temperature or the injection current may be used. When the form of the object has to be determined in a short time, e.g. few microseconds, a pulsed laser has to be used. The wavelength of a pulsed ruby laser may be changed by changing the spacing between the etalon plates of the laser cavity [99Ped1].

Two separated holograms are recorded at two different wavelengths λ_1 and λ_2 , and the phases $\phi_{\lambda 1}$ and $\phi_{\lambda 2}$ are calculated digitally. The phase difference contains the information about the shape



Fig. 9.1.60. Superposition of the 3D vibration amplitudes on the shape of the object. Amplitudes along the (a) x-direction, (b) y-direction, and (c) z-direction, respectively. The shape of the object is given in mm, the gray scale for the vibration amplitude is in μ m. It should be noticed that the gray scale is different for each picture.

of the object. In the case where the illumination is parallel to the observation, the fringes of equal phase obtained represent the intersection of the object surface with equidistant planes of spacing

$$\Lambda = \frac{\lambda_1 \lambda_2}{2 \left| \lambda_1 - \lambda_2 \right|} \,. \tag{9.1.77}$$

If for the illumination a diverging beam is used and the observation is not parallel to the illumination direction, a factor depending on the geometry of the arrangement needs to be introduced in (9.1.77).

An example of the 3D information obtained for a vibrating object is shown in Fig. 9.1.60. The object was a tool used in the ultrasound technology vibrating at a frequency of 20 kHz. The deformations in the three directions (x, y, z) are superimposed onto the shape measurements. The magnitude of the amplitude is coded by gray levels. The maximum of the vibrating amplitude is in the x-direction. This is as expected because the object was excited by a piezoelectric actuator along this direction. The shape of the object was measured by the two-wavelength method implemented by changing of the etalon plate separation. A precise 3D measurement of the vibrations in a point occurred by using a three-beam Doppler laser vibrometer. This measurement is necessary because in double-pulse holography we get only the relative deformation of the object in the time between the two pulses. Even in the case of a simple harmonic vibration, the information of the deformation is unknown. The total object deformation is obtained from the vector resultant from the data obtained from the phase maps for each sensitivity vector with the combination of the object shape obtained from the contouring measurements.

9.1.8.4.3 Pulsed digital holographic interferometry for endoscopic investigations

Holographic interferometry combined with endoscopy enhances the versatility of standard 2D endoscope imaging as it opens up possibilities to measure additional parameters, on hidden surfaces [01See, 02Kol, 03Kol, 03Ped2]. Combinations of the digital holography with an endoscope for carrying the image and a pulsed laser as light source allow measurements in an industrial environment (e.g. vibration measurements, non-destructive testing of technical objects) and in-vivo investigation of biological tissues [99Ped1, 99Sch, 00Kem, 01Sch, 01See, 02Kol].

Figure 9.1.61 shows a schematic illustration of rigid and flexible endoscopes combined with a measuring system based on pulsed holographic interferometry. The optical set-up consists of the pulsed laser, the interferometer unit with the CCD camera, and the endoscope unit. Figure 9.1.61a shows the arrangement for a rigid endoscope. However, the set-up can also be combined with a flexible fiber endoscope, as shown in Fig. 9.1.61b. Rigid and flexible endoscopes have a lot in common. The objective lens forms an image of the subject which in turn is transferred by the relay optics and magnified by a lens system on the sensor. To allow flexibility the image is carried by a bundle of optical fibers, instead of a system of lenses as for the rigid endoscopes. The resolution of a flexible endoscope depends on the number of fibers and their diameter. More fibers of smaller diameter give higher spatial resolution.

For both arrangements the recording procedure and the way to process the digital holograms is exactly the same. The pulsed laser emits short (20 ns) Q-switched pulses, which are divided at the beam-splitter in the reference and the object beam. The reference beam is conveyed to the interferometer unit with a single-mode optical fiber. An image-plane hologram is formed on the CCD detector as a result of the interference between the slightly off-axis reference beam and the object beam. The diverging output beam illuminates the object, the light is diffusely reflected back from the object surface towards the endoscope, which in turn brings the object information to the interferometer unit. The aperture serves to limit the spatial frequencies of the interference pattern in such a way that the detector resolves it. The dimensions of the aperture are chosen considering the resolution of the CCD detector and the distance between the aperture and the sensor. Two or more digital holograms, corresponding to the laser pulses, are captured at separate video frames of the CCD camera.

Figure 9.1.62 shows an example where the system was used to measure inside a pump; Fig. 9.1.62b shows an image of the inside of the pump, and Fig. 9.1.62c shows an interferogram. During the pumping operation, the piston is moving forward and backward (frequency 50 Hz). At the right side (piston) there are more fringes. This means that the piston movement was increased as compared to the other areas around. This example shows that by using the endoscopic technique it is possible to look inside a more or less closed object and investigate deformation and vibrating portions.

A dynamical deformation of in-vivo biological tissues was tested. The problem with measurements of biological tissues is that the reflectivity is not ideal and furthermore any disruption of



Fig. 9.1.61. Set-up with (a) rigid and (b) flexible fiber endoscopes for investigations together with pulsed digital holography.



Fig. 9.1.62. (a) Measurements inside a pump. (b) Image of the object. (c) Phase map obtained during the pump operation.



Fig. 9.1.63. In-vivo investigation inside the oral cavity. (a) Image of the investigated part (tongue).(b) Phase map corresponding to the deformation produced by a shock excitation of the tongue.



Fig. 9.1.64. Image of the prototype, diameter: 18 mm.

the biological tissue produces, together with the deformation, alterations of the microstructure of the surface. Consequently, the correlation between the holographic patterns recorded with the two laser pulses is reduced resulting in noise in the fringe patterns leading to poor quality. Figure 9.1.63 shows phase maps obtained from measurements performed inside the oral cavity (in vivo) using a rigid endoscope.

In order to measure at hidden surfaces, we can combine commercially available endoscopes with an interferometer based on digital holography. Recently, with the newer smaller CCD detector arrays, it has become possible to build the complete interferometric system small (CCD included). Figure 9.1.64 shows a picture of a built prototype. The chip has 659×494 pixels (pixel size $7.4 \times 7.4 \ \mu\text{m}^2$). The sensitive area is quite small ($4.8 \times 3.6 \ \text{mm}^2$), but the sensor is inserted on a mount which has much larger size ($12 \times 13 \ \text{mm}^2$), in effect limiting the size of our holographic head to a diameter of 18 mm. The prototype shown in Fig. 9.1.64 has been used to perform measurements inside a cavity, as shown in Fig. 9.1.65. A pulsed Nd:YAG laser was used for these measurements. Meanwhile the size of the interferometer has been reduced to a diameter of 6 mm.



Fig. 9.1.65. Vibration measurement of an object with a defect located inside a cavity. Vibration frequency: 2350 Hz. (a) Phase map. (b) Pseudo 3D representation of the vibration.

9.1.8.4.4 Temporal phase unwrapping of digital holograms

In Sect. 9.1.7.5 a technique to analyze dynamical deformation based on temporal speckle pattern interferometry is described. The method will be extended to digital holography. The complex amplitude of an object submitted to dynamical deformation is a function of the time. On the sensor plane this amplitude is

$$U_{\rm H}(x, y, t) = |U_{\rm H}(x, y, t)| \exp\left\{i\phi_{\rm H}(x, y, t)\right\}, \qquad (9.1.78)$$

where $|U_{\rm H}(x, y, t)|$ and $\phi_{\rm H}(x, y, t)$ are the amplitude and the phase of the wavefront as a function of the position and the time, respectively. We consider now the case where a sequence of K digital holograms (e.g. 100 or 1000) of an object submitted to deformation is recorded. For each hologram (frame), the intensity is integrated during the exposure period $\Delta \tau$. The intensity recorded in one camera frame can be written as

$$I_{\rm H}(m\Delta x, n\Delta y, k\Delta \tau) = \int_{(k-1)\Delta \tau}^{k\Delta \tau} I_{\rm H}(m\Delta x, n\Delta y, t) \, \mathrm{d}t \,, \qquad (9.1.79)$$

where k is an integer number indicating the recorded frame, k takes the value 1 for the first frame of the sequence and K for the last. The operation described by (9.1.79) corresponds to a time averaging of the hologram during the period $\Delta \tau$. If the deformation of the object is small during the frame exposure, the phase $\phi_{\rm H}$ recorded may be considered as a constant across the pixel and can be reconstructed from the intensity $I_{\rm H}(m\Delta x, n\Delta y, k\Delta \tau)$. The time necessary to record the sequence will be $K\Delta\tau$. Each hologram is then processed individually as shown in Fig. 9.1.66. This operation is nothing else than the calculation of the phase for each recorded hologram by using the method described in Fig. 9.1.57. From the hologram intensity $I_{\rm H}(m\Delta x, n\Delta y, k\Delta\tau)$ we will get the phase $\phi_{\rm H}(m\Delta x, n\Delta y, k\Delta\tau)$ at the point $(m\Delta x, n\Delta y)$ and at the time $k\Delta\tau$. The phase



Fig. 9.1.66. Procedure for calculating the phase from a sequence of digital holograms.



Fig. 9.1.67. Deformation measurement of (\mathbf{a}) a notched piece of a bone immersed in water. Measurement at two different times, namely (\mathbf{b}) 0.5 s and (\mathbf{c}) 1.5 s after the beginning of the loading process.

 $\phi_{\rm H}(m\Delta x, n\Delta y, k\Delta \tau)$ obtained from the evaluation of the digital holograms is in the interval $-\pi$ to π and is indefinite to an additive integer multiplied by 2π . The unwrapping procedure involves the detection of the phase steps. The unknown integer multiplied by 2π is detected by comparison of two successive phase changes. If the change is larger than π , the value of the integer multiple is increased by 1, if the change is lower than $-\pi$, the value is decreased by 1. Consider now only one pixel, the phase at this pixel as a function of the time may easily be unwrapped by a one-dimensional unwrapping algorithm. The temporal unwrapping can be done for all the points of the sensor and a two-dimensional evaluation of the phase as a function of the time can be obtained. Some pixels (due to the low modulation of the interference fringes) do not give a good measurement of the phase. In order to detect the pixels not giving good results, the intensity modulation as a function of the time for each pixel is calculated. The method measures the phase change for each pixel, by assigning the value 0 to the initial phase change and calculate all the deformations with respect to this initial phase.

Combining digital holography and the temporal phase unwrapping leads to the investigation of dynamical deformations [03Ped1]. This method can be applied successfully when the recorded sequence of digital holograms satisfies the sampling theorem in the spatial and temporal domain. For the spatial domain we need the period of the cycle produced by interference between the reference and the object beams to be larger than two times the sensor pixel, this allows to get the phase from the digital hologram. In the time domain we need to sample the hologram in order that from one hologram of the sequence to the next one the change of the phase is smaller than π .

Compared with the temporal speckle method described in Sect. 9.1.7.5 this method has the advantages that it calculates the phase and thus gives us the information about the direction of the deformation without the introduction of additional temporal phase modulation (heterodyne technique) [01Tiz].

The method has been used to measure the deformation of biological tissues. Figure 9.1.67 shows a measurement where the object, a piece of deer antler immersed in water, was loaded by using a special tension device. The camera used had 690×480 pixels with an acquisition rate of 30 full frames/second. The measurements were performed within a period of 4 seconds (during this period the sample was loaded), and 120 holograms (30 each second) were recorded. Figure 9.1.67 shows the results of the object deformation at two different times after the beginning of the loading process. Close to the notch we see a strange deformation. This is the location where the bone will break.

References for 9.1

- 1819Ara Arago, F., Fresnel, A.: Ann. Chim. Phys. 2 (1819) 288.
- 1852Sto Stokes, G.G.: Trans. Cambridge Phil. Soc. 9 (1852) 399–416.
- 21Mic Michelson, A.A., Pease, F.G.: J. Astrophys. **53** (1921) 249.
- 34Van Van Cittert, P.H.: Physica 1 (1934) 201–210.
- 38Zer Zernike, F.: Physica **5** (1938) 785–795.
- 39Van Van Cittert, P.H.: Physica 6 (1939) 1129–1138.
- 48Gab Gabor, D.: Nature (London) **161** (1948) 777.
- 51Hop Hopkins, H.H.: Proc. R. Soc. (London) A 208 (1951) 263–277.
- 53Hop Hopkins, H.H.: Proc. R. Soc. (London) A **217** (1953) 408–432.
- 54Wol Wolf, E.: Proc. R. Soc. (London) A **225** (1954) 96–111.
- 55Hop Hopkins, H.H.: Opt. Acta **2** (1955) 23–29.
- 55Wol Wolf, E.: Proc. R. Soc. (London) A **230** (1955) 246–265.
- 62Shu Shurcliff, W.A.: Polarized light, Cambridge: Harvard University Press, 1962.
- 64Lei Leith, E.N., Upatnieks, J.: J. Opt. Soc. Am. 54 (1964) 1295.
 64Mur Murty, M.V.R.: Appl. Opt. 3 (1964) 853.
- $\mathbf{\mathbf{M}}_{\mathbf{M}}(\mathbf{\mathbf{M}},\mathbf{\mathbf{M$
- 65Gol Goldfischer, L.I.: J. Opt. Soc. Am. **55** (1965) 247.
- 65Man Mandel, L., Wolf, E.: Rev. Mod. Phys. **37** (1965) 231–287.
- 66Fra Francon, M.: Optical interferometry, New York: Academic Press, 1966.
- 70Arc Archbold, E., Burch, J.M., Ennos, A.E.: Opt. Acta 17 (1970) 883–898.
- 70Hop Hopkins, H.H., Tiziani, H.J.: Proc. Int. Symposium on Holography, Besancon, 1970.
- 71But Butters, J.N., Leendertz, J.A.: J. Phys. E 4 (1971) 277–279.
- 71Col Collier, R.J., Burckhard, C.B., Lin, L.H.: Optical holography, Chap. 5, New York: Academic Press, 1971.
- 71Dur Durst, F., Withelaw, J.H.: J. Phys. E 4 (1971) 804.
- 72Tiz1 Tiziani, H.J.: Appl. Opt. 11 (1972) 2911–2917.
- 72Tiz2 Tiziani, H.J.: Opt. Commun. 5 (1972) 271–274.
- 72Wya Wyant, J.C., Bennett, V.P.: Appl. Opt. 11 (1972) 2833–2839.
- 74Bru Bruning, J.H., Herriott, D.R., Gallagher, J.E., Rosenfeld, D.P., White, A.D., Brangaccio, D.J.: Appl. Opt. 13 (1974) 2693–2703.
- 74Cat Cathey, W.T.: Optical information processing and holography, Chap. 9, New York: John Wiley & Sons, 1974.
- 74Hun Hung, Y.Y.: Opt. Commun. **11** (1974) 132–135.

75Goo 75Ste	Goodman, J.W.: Statistical properties of laser speckle patterns, Chap. 2, in: Dainty, J.C. (ed.): Laser speckle and related phenomena, Berlin: Springer-Verlag, 1975, p. 9–75. Stetson, K.A.: Opt. Eng. 14 (1975) 482–489.
76Fri	Friesem, A.A., Levy, U.: Appl. Opt. 15 (1976) 3009–3020.
78But 78Enn 78Erf 78Gre	 Butters, J.N., Jones, R.C., Wykes, C.: Electronic speckle pattern interferometry, in: Erf, R.K. (ed.): Speckle metrology, New York: Academic Press, 1978, p. 111–157. Ennos, A.E.: Prog. Opt. 16 (1978) 233–288. Erf, R.K.: Speckle metrology, New York: Academic Press, 1978. Gregory, D.A.: Topological speckle and structural inspection, in: Erf, R.K. (ed.): Speckle metrology, New York: Academic Press, 1978, p. 183–223.
79Fra 79Sch	Francon, M.: Laser speckle and application to optics, New York: Academic Press, 1979. Schumann, W., Dubas, M.: Holographic interferometry, Heidelberg: Springer-Verlag, 1979.
79Ves	Vest, C.M.: Holographic interferometry, New York: John Wiley & Sons, 1979.
80Tiz	Tiziani, H.J., Leonhardt, K., Klenk, J.: Opt. Commun. 34 (1980) 327–331.
82Doe 82Tak	Dörband, B.: Optik (Stuttgart) 60 (1982) 161. Takeda, M., Hideki, I., Kobayashi, S.: J. Opt. Soc. Am. A 72 (1982) 156–160.
83Sch	Schwider, J., Burow, R., Elssner, K.E., Grzanna, J., Spolaczyk, R., Merkel, K.: Appl. Opt. 22 (1983) 3421–3432.
84Dai 84Lok 84McK 84Per	 Dainty, J.C.: Introduction, in: Dainty, J.C. (ed.): Laser speckle and related phenomena, 2nd edition, Berlin: Springer-Verlag, 1984. Lokberg, O.J.: J. Acoust. Soc. Am. 75 (1984) 1783–91. Mc Kechnie, T.S.: Speckle reduction, in: Dainty, J.C. (ed.): Laser speckle and related phenomena, 2nd edition, Berlin: Springer-Verlag, 1984. Perry, G.: Speckle patterns in partially coherent light, in: Dainty, J.C. (ed.): Laser speckle and related phenomena, 2nd edition, Berlin: Springer-Verlag, 1984.
85Cre 85Doe 85Tha	Creath, K.: Appl. Opt. 24 (1985) 3053–3058. Dörband, B., Tiziani, H.J.: Appl. Opt. 24 (1985) 2604–2611. Thalmann, R., Dändliker, R.: Opt. Eng. 24 (1985) 930–935.
86Joe 86Sch	Joenathan, C., Narayanmurthy, C.S., Sirohi, R.S.: Opt. Commun. 56 (1986) 309–312. Schwider, J., Burow, R., Elssner, K.E., Grzanna, J., Spolaczyk, R.: Appl. Opt. 25 (1986) 1117–1121.
87Bor 87Tyr	Born, M., Wolf, E.: Principles of optics, 6th edition, Oxford: Pergamon Press, 1987. Tyrer, T.R.: Proc. SPIE (Int. Soc. Opt. Eng.) 746 (1987) 98–103.
88Dae 88Ste	Dändliker, R., Thalman, R., Prongue, D.: Opt. Lett. 13 (1988) 339. Stetson, K.A., Brohinsky, W.R.: J. Opt. Soc. Am. A 5 (1988) 1472–1476.
89Jon	Jones, R., Wykes, C.: Holographic and speckle interferometry, 2nd edition, Cambridge: Cambridge Univ. Press, 1989.
90Har	Hariharan, P.: Selected papers on interferometry, Bellingham, Washington: SPIE Press, 1990.

282	References for 9.1
90Joe 90Ker 90Kin	Joenathan, C., Tiziani, H.J., Pfister, P.: Appl. Opt. 29 (1990) 1905–1911. Kerr, D., Santoyo, F.M., Tyrer, J.R.: J. Opt. Soc. Am. 7 (1990) 820–826. Kino, G.S., Chim, S.S.C.: Appl. Opt. 29 (1990) 3775.
91Cre 91Sir 91Sod	Creath, K.: Proc. SPIE (Int. Soc. Opt. Eng.) 1553 (1991) 213. Sirohi, R.S.: Selected papers on speckle metrology, Washington: SPIE Press, 1991. Sodnik, S., Fischer, E., Ittner, T., Tiziani, H.J.: Appl. Opt. 30 (1991) 3139.
92Bro	Brosseau, C.: Fundamentals of polarized light: a statistical approach, New York: John Wiley & Sons. Inc. 1992
92Dre	Dresel T Häusler G Venzke H · Appl Opt 31 (1992) 919
92Gre	Greivenkamp, J.E., Bruning, J.H.: Phase shifting interferometry, in: Malacara, D. (ed.): Optical shop testing, 2nd edition, New York: John Wiley & Sons, Inc., 1992.
92Mal	Malacara, D.: Optical shop testing, New York: John Wiley & Sons, Inc., 1992.
92Rod	Rodriguez-Vera, R., Kerr, D., Mendoza-Santoyo, F.M.: J. Opt. Soc. Am. A 9 (1992) 2000–2008.
93Col	Collett, E.: Polarized light – fundamentals and applications, New York: Marcel Dekker, 1993.
93Dav	Davies, J.C., Buckberry, C.H.: Television holography and its applications, in: Williams, D.C. (ed.): Optical methods in engineering metrology, London: Chapman and Hall, 1993.
93Hun	Huntley, J.M., Saldner, J.: Appl. Opt. 32 (1993) 3047–3052.
93Kau	Kaufmann, G.H.: Automatic fringe analysis procedures in speckle metrology, in: Sirohi, R.S.: Speckle metrology, New York: Marcel Dekker, 1993.
93Lok	Lokberg, O.J.: Recent developments in video speckle interferometry, in: Sirohi, R.S.: Speckle metrology, New York: Marcel Dekker, 1993, p. 157–194.
93Ped	Pedrini, G., Pfister, B., Tiziani, H.J.: J. Mod. Opt. 40 (1993) 89–96.
93Pry	Pryputniewicz, R.J.: Proc. SPIE (Int. Soc. Opt. Eng.) 1821 (1993) 27–37.
93Rob	Robinson, D.W., Reid, G.T.: Interferogram analysis: Digital fringe pattern measurement
93Sch	Schwider, J., Falkensdorfer, O., Schreiber, H., Zöller, A., Streibl, N.: Opt. Eng. 32 (1993) 1883
93Sir	Sirohi, R.S.: Speckle metrology, New York: Marcel Dekker, 1993.
93Yam	Yamaguchi, I.: Theory and application of speckle displacement, in: Sirohi, R.S.: Speckle metrology, New York: Marcel Dekker, 1993.
94Joe 94Oka	Joenathan, C., Tiziani, H.J., Franze, B.: Appl. Opt. 33 (1994) 7305–7311. Okada, E., Enomoto, H., Minamitani, H.: Proc. IEEE: IECON 1994, Vol. 2 (1994)
040.1	890-893.
94Sch	Schnars, U.: J. Opt. Soc. Am. A 11 (1994) 2011–2015. Sreedhan P. Krishna Mohan K. Sirahi P.S. Opt. Eng. 22 (1004) 1080-05
94Zou	Zou, Y., Pedrini, G., Tiziani, H.J.: J. Mod. Opt. 41 (1994) 1637–52.
95Ped	Pedrini, G., Zou, Y., Tiziani, H.J.: J. Mod. Opt. 42 (1995) 367–374.
95Vdo	Vdovin, G., Sarro, P.M.: Appl. Opt. 34 (1995) 2968.
96Goo	Goodman, W.: Introduction to Fourier optics, New York: McGraw-Hill, 1996.
96Kre	Kreis, Th.: Holographic interferometry, principles and methods, Berlin: Akademie Ver-
96Lei	iag, 1996. Leibbrand, G.W.R., Harbeers, G., Kunst, P.J.: Appl. Opt. 35 (1996) 6151.

96Mei	Meinschmidt, P., Hinsch, K.D., Sirohi, R.S.: Selected papers on speckle pattern inter- ferometry – principles and practice. Washington: SPIE Press, 1996.
96Ped	Pedrini, G., Zou, Y., Tiziani, H.J.: J. Mod. Opt. 43 (1996) 395–407.
965cn	Schnars, U., Kreis, In., Jueptner, W.P.: Opt. Eng. 35 (1996) 977–982.
9011Z1	Tiziani, H.J., Maier, N., Rotne, A.: Opt. Commun. 123 (1996) 34.
9011Z2	11ziani, H.J., Rotne, A., Maler, N.: Appl. Opt. 45 (1996) 3225.
97Ada	Adams, M., Kreis, Th., Jueptner, W.: Proc. SPIE (Int. Soc. Opt. Eng.) 3098 (1997) 234.
97Bot	Bothe, T., Burke, J., Helmers, H.: Appl. Opt. 36 (1997) 5310–5316.
97Her	Hertwig, M.: Proc. SPIE (Int. Soc. Opt. Eng.) 3098 (1997) 90–105.
97Hun	Huntley, J.M., Saldner, H.: Opt. Eng. 36 (1997) 610–615.
97Ped1	Pedrini, G., Froening, Ph., Fessler, H., Tiziani, H.J.: Opt. Laser Technol. 29 (1997) 505–511.
97Ped2	Pedrini, G., Tiziani, H.J.: Opt. Laser Technol. 29 (1997) 249–256.
97Ped3	Pedrini, G., Tiziani, H.J., Zou, Y.: Opt. Lasers Eng. 26 (1997) 199–219.
97Ped4	Pedrini, G., Zou, Y., Tiziani, H.J.: Appl. Opt. 36 (1997) 786–792.
97Ras1	Rastogi, P.K.: Opt. Lasers Eng. 26 (1997) 83–278.
97Ras2	Rastogi, P.K.: Optical measurement techniques and applications, London: Artech House Boston 1007
07Sch	Schreiber H. Schwider I: Appl. Opt. 36 (1007) 5321-5324
97501 07Tiz1	Tiziani H I: Chap 2 in: Bastori DK (ad): Ontical manufung techniques and appli
371121	entions Boston: Artech House Inc. 1007
07Tiz2	Tiziani H I Franze B Haible P: I Mod Ont 44 (1007) 1485
9711ZZ	1121am, 11.9., Franze, D., Habble, I., J. Mod. Opt. 44 (1997) 1465.
98Fra	Franze, B., Tiziani, H.J.: J. Mod. Opt. 45 (1998) 861.
98Gre	Gren, P., Schedin, S., Li, X.: Appl. Opt. 37 (1998) 834–840.
98Joe1	Joenathan, C., Franze, B., Haible, P., Tiziani, H.J.: Appl. Opt. 37 (1998) 2608–2614.
98Joe2	Joenathan, C., Franze, B., Haible, P., Tiziani, H.J.: Appl. Opt. 37 (1998) 3385–3390.
98Mov	Movilla, J.M., Piquero, G., Martínez-Herrero, R., Mejías, P.M.: Opt. Commun. 149 (1998) 230.
98Ped	Pedrini, G., Froening, Ph., Fessler, H., Tiziani, H.J.: Appl. Opt. 37 (1998) 6262-6269.
98Seb	Sebacher, S., Osten, W., Jueptner, W.: Proc. SPIE (Int. Soc. Opt. Eng.) 3479 (1998)
	104–115.
98Ste	Steinchen, W., Yang, L.X., Kupfer, G., Mackel, P., Vossing, F.: J. Strain Anal. 33 (1998)
	171–182.
99Joe	Joenathan, C., Franze, B., Haible, P., Tiziani, H.J.: Appl. Opt. 38 (1999) 1169–1178.
99Moo	Moore, A.J., Hand, D.P., Barton, J.S., Jones, J.D.C.: Appl. Opt. 38 (1999) 1159–1162.
99Ped1	Pedrini, G., Fröning, Ph., Tiziani, H.J., Gusev, M.E.: Appl. Opt. 38 (1999) 3460-3467.
99Ped2	Pedrini, G., Fröning, Ph., Tiziani, H.J., Mendoza Santoyo, F.: Opt. Commun. 164
	$(1999) \ 257-268.$
99Ped3	Pedrini, G., Schedin, S., Tiziani, H.J.: Opt. Commun. 171 (1999) 29–36.
99Piq	Piquero, G., Movilla, J.M., Mejías, P.M., Martínez-Herrero, R.: Opt. Quantum Electron.
-	31 (1999) 223.
99Sch	Schedin, S., Pedrini, G., Tiziani, H.J.: Opt. Commun. 165 (1999) 183–188.
99Sir	Sirohi, R.S., Chau, F.S.: Optical methods of measurement: Whole field techniques, New
	York: Marcel Dekker, 1999.
99Wag	Wagner, C., Seebacher, S., Osten, W., Jüptner, W.: Appl. Opt. 38 (1999) 4812–4820.
0.017	
UUKem	Kemper, B., Dirksen, D., Avenhaus, W., Merker, A., von Bally, G.: Appl. Opt. 39 (2000) 3899–3905.

284	References for 9.1
00Mov	Movilla, J.M., Piquero, G., Mejías, P.M., Martínez-Herrero, R.: Opt. Quantum Electron. 32 (2000) 1333.
00Tot	Totzeck, M., Jacobsen, H., Tiziani, H.J.: Appl. Opt. 39 (2000) 6295–6305.
01Ge 01Gru	Ge, Z., Kobayashi, F., Matsuda, S., Takeda, M.: Appl. Opt. 40 (2001) 1649. Gruber M. Hagedorn D. Eckert W.: Appl. Opt. 40 (2001) 5052–5055
01Kum 01Leh	Kumar, R., Singh, S.K., Shakher, C.: Opt. Laser Technol. 33 (2001) 567–571. Lehmann, M.: Speckle statistics in the context of digital speckle interferometry, in: Rastogi, P.K. (ed.): Digital speckle pattern interferometry and related techniques, Chichester: John Wiley & Sons, 2001.
01Moo	Moore, A.J., Jones, J.D.C., Valera, J.D.R.: Dynamic measurement, in: Rastogi, P.K. (ed.), Digital speckle pattern interferometry and related techniques, Chichester: John Wiley & Sons. 2001.
01Ped	Pedrini, G., Tiziani, H.J.: Digital holographic interferometry, in: Rastogi, P.K. (ed.): Digital speckle pattern interferometry and related techniques, Chichester: John Wiley & Sons, 2001.
01Per	Perez-Lopez, C., Mendoza Santoyo, F., Pedrini, G., Schedin, S., Tiziani, H.J.: Appl. Opt. 40 (2001) 5106–5110.
01Ras1	Rastogi, P.K.: Digital speckle pattern interferometry and related techniques, Chichester: John Wiley & Sons, 2001.
01Ras2	Rastogi, P.K.: Surface displacements, derivatives and shapes, in: Rastogi, P.K. (ed.): Digital speckle pattern interferometry and related techniques, Chichester: John Wiley & Sons, 2001.
01Sch	Schedin, S., Pedrini, G., Tiziani, H.J., Aggarwal, A.K., Gusev, M.E.: Appl. Opt. 40 (2001) 100–103.
01See	Seebacher, S., Osten, W., Baumbach, T., Jüptner, W.: Opt. Lasers Eng. 36 (2001) 103–126.
01Tiz	Tiziani, H.J.: Optik (Stuttgart) 9 (2001) 370–380.
02Kol	Kolenovic, E., Lai, S., Osten, W., Jüptner, W.: A miniaturized digital holographic en- doscopic system for shape and deformation measurement, in: VDI Berichte 1694: In- ternational symposium on photonics and measurement, Düsseldorf: VDI Verlag GmbH, 2002, p. 79–84.
03Har 03Kol	Hariharan, P.: Optical interferometry, 2nd edition, London: Academic Press, 2003. Kolenovic, E., Osten, W., Klattenhoff, R., Lai, S., von Kopylow, Ch., Jüptner, W.: Appl. Opt. 42 (2003) 5167–5172.
03Mad	Madjarova, D., Kadono, V., Toyooka, H.: Opt. Express 6 (2003) 617.
03Ped1	Pedrini, G., Alexeenko, I., Osten, W., Tiziani, H.J.: Appl. Opt. 42 (2003) 5846–5854. Pedrini, C., Cusey, M., Schedin, S., Tiziani, H.J.: Opt. Lagara Eng. 40 (2003) 487, 400
031 euz 03Rei	Reichelt, S., Pruss, C., Tiziani, H.J.: Appl. Opt. 42 (2003) 4468–4479.
03Wya	Wyant, J.C.: Opt. Photonics News 15 (2003) 36.
04Joe	Joenathan, C., Tiziani, H.J.: Speckle metrology, in: Wiley: Optics Encyclopedia, 2004.
04Lie	Liesener, J., Tiziani, H.J.: Proc. SPIE (Int. Soc. Opt. Eng.) 5252 (2004) 264–271.
04Pru 04Tiz	 Pruss, C., Tiziani, H.J.: Opt. Commun. 233 (2004) 15–19. Tiziani, H.J., Totzeck, M.: High-precision optical metrology for surfaces, Chap. 5.1, in: Landolt-Börnstein, Group VIII, Vol. 1, Laser Physics and Applications, Subvol. C, Berlin: Springer-Verlag, 2004.

Index

ABCD law 154, 157 for the q-parameter 159 matrix, beam shaping 129 transformation, beam shaping 127 Absolute heterodyne interferometry 243 with rotational matched grating 243 Absorbing mask, beam shaping 127 Absorption, thin film 117 Acousto-optic Bragg cell 108 modulator 104 cell, Bragg 108 deflector 106 intensity modulator 106 light modulator 85 material 105 modulator 98,106 Bragg 104 intensity 106 light 85 Acousto-optic material BGO 105 fused silica 105 GaAs 105 GaP 105, 106, 108 Ge = 105GeAsSe glass 105 Hg_2Cl_2 105 $LiNbO_3 \quad 98, 102, 105, 108 \\$ PbMoO₄ 98, 105, 106 silica, fused 105 SiO_2 98 TeO_2 98, 105, 106, 108 TiO_2 105 Tl_3AsS_4 105 Tl_3AsSe_3 105 Active mode locking 85 Adaptive optics 246 interferometry with 244 AD^*P ((NH₄)D₂PO₄) electro-optics 92 ADP $((NH_4)H_2PO_4)$ electro-optics 88, 92, 95, 96 Advanced laser component 122

Amplification, irreversible 19 Amplified spontaneous emission (ASE) 67 Amplitude modulation 90,94 modulator 94 stabilization 59 Analysis beam transfer, beam shaping 132 deformation 273 fringe by phase shifting 232 of vibrating objects 257 phase by phase stepping 256 by spatial phase shifting 256in speckle interferometry 256vibration 273 Fizeau arrangement 263 Analytic signal 50 Angle half, of divergence, lens resonator 193 of divergence half, lens resonator 193 higher-order modes, circular symmetry 159 higher-order modes, rectangular symmetry 159of the Gaussian beam 158 of misalignment 168 Angular momentum operator 68 Anisotropic crystal 86 interaction 102 Annihilation of a photon 15 operator 10 Annular ring 178 Antibunching 24, 28 photon 4,59 Antistokes Raman scattering, coherent (CARS) 56Aperture -limited stable resonator 164 resonator with one 164 resonator with two 166 Approximation

dipole 14

Fresnel 268 rotating-wave 15,71 slowly varying envelope (SVEA) 71 Arago 221 and Fresnel 226 Archimedic spiral curve, beam shaping 140 Area cross-sectional, of the Gaussian beam 154 of coherence 54 small, pulse 76 ArF excimer, optical coating system 122 Arrangement Fizeau, for vibration analysis 263 Mirau 237 Array, lenslet, beam shaping 134 ASE (amplified spontaneous emission) 67 Aspheric surface 246 Asymmetric confocal unstable ring resonator 199 Atom –field interaction 14 laser 35 optics 14, 32, 35 nonlinear 36 -reservoir interaction 22 Rydberg 70 trapping 78 two-level 15 Autocorrelation function 25 Available power 187 Average ensemble 51 time-, technique 265 Ba ions, pair of superradiation 74 Band -pass filter 123 transmission, spectral 114 -width of modulation 103 Basic principles of interference 222 $BaTiO_3$ electro-optics 88,97 Beam coherent beam shaping 130 partially, beam shaping 132 deflector, electro-optic 94 Gaussian 154 angle of divergence 158 cross-sectional area 154 divergence 158 radius 153 integration, beam shaping 133 integrator beam shaping 128, 134, 136 multifaceted, beam shaping 135 parameter product 160, 162, 179, 182, 195

maximum 195 partially coherent, beam shaping 132 propagation 157 factor 160, 161 of higher-order modes 160 resonator, unstable 172 quality 161 resonator, unstable 178 radius at mirror, lens resonator 193 at the principal planes, lens resonator 193 Gaussian 153 higher-order modes, circular symmetry 159 higher-order modes, rectangular symmetry 159scanning 127 shaping 127, 138 ABCD-matrix 129 ABCD-transformation 127absorbing mask 127 Archimedic spiral curve 140 beam, partially coherent 132 beam integration 133 beam integrator 128, 134, 136 beam integrator, multifaceted 135 beam splitting 139 beam-transfer analysis 132 beam transformation 130 beamlet 128 coherence 138 coherence length 138 coherence length, transverse 128 coherent beam 130 Collins integral 130 diamond turning 141 element, manufacturing of 140 element, phase-diffusing 138 Fourier transformation 135 Gerchberg–Saxton iteration 130, 134gray-scale lithography 141 GRIN technique 140 imaging integrator 134 integrator, flies-eye 135 intensity, mutual 132 lenslet array 134 lithographic technique 140, 141 lithography, gray-scale 141 mirror 141 multifaceted mirror 128 phase element, diffractive 140 phase element, non-quadratic 127 phase mask 127 power density distribution 127 projection technique 127 Schell model 138 smoothing parameter ρ 139

structure function d = 127technique 129 technique, lithographic 140, 141 tophat distribution 133 tophat profile 128 transformation efficiency 131 Wigner distribution function 132 splitting, beam shaping 139 transfer analysis, beam shaping 132 transformation, beam shaping 130 two-, interference 223 waist, location of, lens resonator 193 Beamlet 128 beam shaping 128 shaping 127, 137, 140 Bell inequality 5, 31, 60 John S. 31 Benzene (C_6H_6) electro-optics 97 $\beta - ZnS$ electro-optics 95BGO acousto-optics 105 Biological tissue, in-vivo 276Biphoton 60 Birefringence 86 Bloch sphere 72Bogoliubov, Nikolai N. 34 Bohr, Niels 19, 26 Boltzmann photon statistics 12 Born, Max 3 Bose -Einstein condensate 35,77 condensate Na, superradiation 74 condensation 4, 33, 34 distribution 62Satyendra Nath 34 Boson 34 Boundary stability 150 Bragg acousto-optic cell 108modulator 104 cell 107 acousto-optic 108 condition 101, 102 diffraction 101 efficiency 102 modulator, acousto-optic 104 regime 101 Branch negative resonator 175 resonator, unstable 172

ring resonator, unstable 200 positive resonator 174 resonator, unstable 172 Broadening, linehomogeneous 187, 188 inhomogeneous 187, 190 Brown and Twiss 49,58 Robert Hanbury 4 Bunching anti- 24, 28 photon 59 Burnham–Chiao ringing 75 Burning, spatial hole 184, 196 Burshtein, Anatoly I. 26 $\mathrm{C}_{12}\mathrm{H}_{10}$ in diphenyl, superradiation 74 C₅H₇NO₂ (nitrotoluene) electro-optics 97 $C_6H_5NO_2$ (nitrobenzene) electro-optics 97 C_6H_6 (benzene) electro-optics 97 Calorimetric measurement 118 Cam, phase-, set-up 234 Carbon disulfide (CS_2) electro-optics 97 CARS (coherent Antistokes Raman scattering) 56 CCD (charge coupled device) 231 CdTe electro-optics 93,95 Cell Bragg 107 acousto-optic 108 Kerr 85 CGH (computer-generated hologram) 236, 244 CH_3CN superradiation 74 CH₂F superradiation 74 Chaotic light 55,62 source 55 Characterization of unstable resonators 172 Charge coupled device (CCD) 231Charging time, electrode 91 CHCl₃ (chloroform) electro-optics 97 Chloroform (CHCl₃) electro-optics 97 Choice, delayed, experiment 28 Circular symmetry eigenmodes 152
higher-order modes 159 higher-order modes, angle of divergence 159higher-order modes, beam radius 159 higher-order modes, waist radius 159 resonance frequencies 163 waveguide hollow 205 symmetric 209 Cittert, van 222 –Zernike theorem 222 Class, crystal 3m 92 **4**2m 92 **4**3m 93 Classical light 49 stochastic quantity 71 Classification, resonator, optical 149 CMOS (complementary metal oxide semiconductor) 231 CO_2 laser optical coating system 122 waveguide 201 Coating fluoride 124 optical 114 production of 114 system high-power broad-band 124 optical 122 optical, ArF-excimer 122 optical, CO_2 -laser 122 optical, F₂-excimer 122 optical, HeNe-laser 122 optical, KrF-excimer 122 optical, Nd:YAG-laser 122 Coblentz sphere 119 Coefficient elasto-optic 99 gain g_0 186 gain, small-signal g_0 184 thermal-lensing 191 Coherence 49, 52, 222 area of 54 beam shaping 138 complex degree of 223 degree of 52 complex 223 factor 224 first-order 52 full 52 higher-order 57 historical remarks 49 length 55, 222 beam shaping 138 longitudinal 53, 55

transverse 53.55 transverse, beam shaping 128 matrix 240 partial 52 pure spatial 222 temporal 222 spatial 53, 223 pure 222 temporal 52 pure 222 theory 51 higher-order 222time 56, 222 volume 54 Coherency matrix 228 Coherent Antistokes Raman scattering (CARS) 56 beam, beam shaping 130 interaction 56 matter waves 35 partially – beam, beam shaping 132 state 10,56 Stokes Raman scattering (CSRS) 56 Coincidence, delayed 59 Collins integral, beam shaping 130Communication, quantum Commutation relation 8 Complementary metal oxide semiconductor (CMOS) 231 Complex correlation factor 228 degree of coherence 223 Component, laser, advanced 122 Computer-generated hologram (CGH) 236, 244 Computing, quantum 4 Condensate, Bose–Einstein 35,77 Condensation, Bose–Einstein 4, 33, 34 Condition Bragg 101, 102 confocal 151 resonance, of the optical resonator 152 steady-state 185 Configuration dynamic retarder 90 waveguide resonator 206 Confined output coupler 199 Confocal condition 151 resonator 175 asymmetric unstable ring 199 off-axis unstable 212 Constant, Kerr 97 Contrast, speckle 251 Converging wave 173 Conversion, down-, parametric 60

Cooperative Rayleigh scattering 77 spontaneous emission 67 system 77 Corrector, null 245 interferometry with 244 Correlation factor, complex 228 fringe 255 function field 51 first-order 51 second-order 51,58 intensity spatial 58 temporal 59 Coulomb gauge 5 condition 5 Coupler, output, confined 199 Coupling, output, optimum 187 Cr: ruby superradiation 74 Creation of a photon 15 Critical parameter of laser components 117 Cross-sectional area of the Gaussian beam 154Cryptography, quantum 37, 38 Crystal anisotropic 86 class 3m 92 **42m** 92 **4**3m 93 liquid-, phase shifter 239 Cs vapor superradiation 74 CS_2 (carbon disulfide) electro-optics 97 CSRS (coherent Stokes Raman scattering) 56 CuCl electro-optics 93 Curvature, radius of 158 Curve Archimedic spiral, beam shaping 140 of constant magnification 175 of constant misalignment sensitivity 170 Damage laser-induced 120 threshold (LIDT) 117, 120 threshold 94 Damped radiation field 18 Damping 19 term, phenomenological 71 Dark spot in a speckle pattern 253Data recording intensity, five-frame 233intensity, four-frame 233

de Broglie wavelength 57 Decay, Weisskopf–Wigner 23 Decoherence 5, 19, 20, 51 Decomposition, mode 7 Deflector acousto-optic 106 beam-, electro-optic 94 Deformable membrane mirror 247 Deformation analysis 273 dynamic 272 digital holographic interferometry for 272 measurement 253, 260 Degeneracy parameter 54 Degenerate quantum gas, fermionic 4 Degree of coherence 52 complex 223 Dehmelt, Hans G. 27 Delaved -choice experiment 28 coincidence 59 Density distribution, power, beam shaping 127 Detector charge-coupled (CCD) 231 "which-path" 28 Device, modulator 87 DHI (double heterodyne interferometry) 243 Diagram, stability 151 Diamond turning, beam shaping 141 Dicke 67 Diffraction Bragg 101 efficiency 102loss 167 Raman–Nath 100 Rayleigh–Sommerfeld 267 integral 267 Diffractive phase element, beam shaping 140 Digital hologram 267 holographic interferometry 260, 272for dynamic deformations 272pulsed 276 holography 266 principle of 266 Diode optical 196 -pumped monolithic nonplanar ring oscillator 200 Nd:YAG 191 Nd:YAG, monolithic nonplanar ring oscillator 200 Nd:YLF 191 Nd:YVO₄ 191 Dipole approximation 14

interaction operator 68 moment electric 14 operator, total 68 operator 68 interaction-68 Dirac 67 Paul Adrian Maurice 3 Displacement measurement 253 Distance measurement, heterodyne interferometry for 241 Distribution Bose–Einstein 62 field, steady-state 152 function photon 62 Wigner, beam shaping 132 Glauber–Sudarshan 13 phase space 19 photon 10 Poisson 10, 62, 63 power density, beam shaping 127 tophat, beam shaping 133 Divergence angle half, lens resonator 193 higher-order modes, circular symmetry 159 higher-order modes, rectangular symmetry 159of the Gaussian beam 158 half angle lens resonator 193 of the Gaussian beam 158 Diverging wave 173 Donut mode 156 Double exposed hologram 265 exposure holography 265 heterodyne interferometry 243slit experiment 28,29 Young's 50 Down-conversion, parametric 60 Drei-Männer-Arbeit 3 Duality, wave-particle 28 Dynamic deformation 272 digital holographic interferometry for 272 retardation 89 retarder configuration 90 Effect

electro-optic Kerr 86,96 Kerr, in isotropic media 96 linear 86,87 Pockels 86,87

quadratic 86,96 quadratic, in isotropic media 96 photoelastic 86,98 vector – of interference 226 Effective Fresnel number 164 Efficiency Bragg diffraction 102 extraction 183, 189 maximum 188, 198 slope 187 transformation -, beam shaping 131 Ehrenreich and Woodruff relation 104 Eigenmode circular symmetry 152 of the optical resonator 151 of the resonator 153 rectangular symmetry 153 Eigensolution 151 Einstein Albert 3, 30, 34 Podolsky, Rosen (EPR) 30 paradox 60 Elasto-optic coefficient 99 Electric dipole moment 14 field -dipole interaction 14 strength 50 Electro-magnetic, transverse (TEM) 153 Electro-optic beam deflector 94 effect Kerr 86,96 Kerr, in isotropic media 96 linear 86,87 Pockels 86, 87 quadratic 86,96 quadratic, in isotropic media 96 light modulator 85 material 95 modulator 96 light 85 Electro-optics AD^*P ((NH₄)D₂PO₄) 92 ADP $((NH_4)H_2PO_4)$ 88, 92, 95, 96 BaTiO₃ 88,97 benzene (C_6H_6) 97 β -ZnS 95 $C_5H_7NO_2$ (nitrotoluene) 97 $C_6H_5NO_2$ (nitrobenzene) 97 C_6H_6 (benzene) 97 disulfide (CS_2) 97 CdTe 93,95 CHCl₃ (chloroform) 97 chloroform (CHCl₃) 97 CS_2 (carbon disulfide) 97

CuCl 93 GaAs 93,95 GaP 95GaSe 88 H_2O (water) 97 InAs 93 KD_2PO_4 (KD^*P) 92,94–96 $KD^*P(KD_2PO_4) = 92,94-96$ KDP (KH₂PO₄) 88, 91, 92, 94, 95 KH₂PO₄ (KDP) 88, 91, 92, 94, 95 KTa_{0.65}Nb_{0.35}O₃ 97 LiNbO₃ 88, 92-96 :MgO 96 LiTaO₃ 88, 92, 95, 96 MgO:LiNbO₃ 96 $(NH_4)D_2PO_4 (AD^*P) 92$ $(NH_4)H_2PO_4$ (ADP) 88, 92, 95, 96 nitrobenzene (C₆H₅NO₂) 97 nitrotoluene $(C_5H_7NO_2)$ 97 $Pb_{0.93}La_{0.07}(Zr_{0.65}Ti_{0.35})0_3 97$ water (H_2O) 97 ZnS 93 $\beta - 95$ ZnSe 95 ZnTe 95Electrode charging time 91 Electron-hole plasma 78 Electronic phase evaluation technique 231 speckle pattern interferometry 254 Element beam-shaping, manufacturing of 140 phase diffractive, beam shaping 140 -diffusing, beam shaping 138 non-quadratic, beam shaping 127 Emission, spontaneous 5,24 amplified (ASE) 67 cooperative 67 Endoscope, fiber 276 Endoscopic investigation 276Energy of the radiation field 7 output per pulse 189 Ensemble average 51 Entangled photon pair 60 state 4 Entanglement 30 EPR (Einstein, Podolsky, Rosen) 30 paradox 60 source 37 Equal -inclination fringe 225 -thickness fringe 225

Equation Gross-Pitaevskii 34 Helmholtz 6 Kirchhoff integral 151 master 18, 21, 22, 69 Maxwell–Bloch (MB) 71 quantum (QMB) 71 Schrödinger 15 Equivalent G-parameter 149, 164, 174 Eu vapor superradiation 74 Evaluation technique, phase, electronic 231Evaporation, thermal 121 Excimer laser ArF-, optical coating system 122 F_{2} -, optical coating system 122 KrF-, optical coating system 122 Excitation 54 Expansion, mode 5 Experiment delayed choice 28 double-slit 28, 29 Franson 61 Exponential method 118 Exposed, double, hologram 265 Exposure, double, holography 265 Extended multiatomic system 70 External modulation 85 Extraction efficiency 183, 189maximum 188, 198 F_2 excimer, optical coating system 122 Fabry–Perot interferometer 231 Factor coherence 224 correlation, complex 228 fill 187 loss 151, 152, 165, 166 per round trip 176 propagation beam 160, 161 Far off-resonant case 17 Faraday rotator 197 Fast Fourier transform (FFT) 268FEL (free electron laser) 78 Fermionic quantum gas, degenerate 4 Feshbach resonance 36 FFT (fast Fourier transform) 268Fiber endoscope 276 Field correlation function 51 distribution, steady-state 152mode 50 quantization 5,7 radiation –, damped 18 single-mode 50

state 8 strength, electric 50 Fill factor 187 Film, thin absorption 117 matrices M_m 112 polarizer 123 reflectance 113 system 111 technology 111 future trends in 123 Filter band-pass 123 rugate 124 First-order coherence 52 correlation function 51 Five-frame intensity data recording 233 Fizeau arrangement for vibration analysis 263 fringe 225 interferometer 230 Flashlamp-pumped Nd:glass 191 Nd:YAG 191Nd:YAP 191Nd:YLF 191Flies-eye integrator, beam shaping 135 Fluctuation, vacuum 5 Fluorescence, resonance 24 Fluoride coating 124 Fock state 9, 16, 17, 29 Focus, intracavity 199 Formulation, path integral 3 Four-frame intensity data recording 233 Four-wave mixing 36 Fourier hologram 264 lensless 269 quasi-, hologram 270 transform fast (FFT) 268 technique 232 temporal speckle pattern shearing interferometry (TSPSI) 258 transformation, beam shaping 135 Franson experiment 61 Free electron laser (FEL) 78 Frequency Rabi 15,16 resonance 163 circular symmetry 163 rectangular symmetry 163 Fresnel 221 and Arago 226approximation 268

hologram 264, 270 number 166 effective 164 Fringe analysis, by phase shifting 232 correlation 255 Fizeau 225 Haidinger 225 of equal inclination 225 of equal thickness 225 Full coherence 52 Function autocorrelation 25 correlation field 51 first-order 51 second-order 51,58 distribution photon 62 Wigner, beam shaping 132 structure, beam shaping 127 transfer 117 Fundamental mode 154, 157 operation 169 waveguide 205 Fused silica acousto-optics 105 Future trends in thin-film technology 123 g-parameter 149 G-parameter, equivalent 149, 164, 174 Ga vapor superradiation 74 GaAs acousto-optics 105 electro-optics 93,95 Gabor 263 Gain coefficient $g_0 = 186$ small-signal g_0 184GaP acousto-optics 105, 106, 108 electro-optics 95 Gas, quantum, degenerate fermionic 4 GaSe electro-optics 88 Gauge condition, Coulomb 5 Coulomb 5 condition 5 Gauss -Hermite mode 152, 153 -Laguerre mode 152 Gaussian beam 154 angle of divergence 158 cross-sectional area 154

divergence 158 radius 153 speckle 252 Ge acousto-optics 105GeAsSe glass acousto-optics 105General interference law 224 Gerchberg–Saxton iteration, beam shaping 130,134Glass GeAsSe, acousto-optics 105Glauber 49 Roy J. 4 state 56 -Sudarshan distribution 13 Gouy phase shift 158 Grating interferometer 236 matched, rotational 243 phase 99 moving 98 rotating 242 volume-index 99 Gray-scale lithography, beam shaping 141 Green–Twyman interferometer 229 GRIN technique, beam shaping 140 Gross–Pitaevskii equation 34 H_2O (water) electro-optics 97 Hahn and McCall, theorem of 72 Haidinger fringe 225 Haken, Hermann 4 Half angle of divergence, lens resonator 193 -wave voltage 89,90 Hamiltonian 8,15 interaction 15 Jaynes–Cummings–Paul (JCP) 18, 21 JCP (Jaynes–Cummings–Paul) 18, 21 Hartmann 246 sensor 246 -Shack 246 sensor 246 4 He, liquid, superfluidity of 34Heisenberg inequality 8 Werner 26 Helmholtz equation 6 HeNe laser, optical coating system 122 Hermite -Gauss mode 152, 153 polynomials 153 Heterodyne interferometry absolute 243 with rotational matched grating 243

double 243 for distance measurement 241 for velocity measurement 241 HF superradiation 74 Hg_2Cl_2 acousto-optics 105 High-precision metrology 221 High-reflecting mirror 120 Higher-order coherence 57 theory 222 Higher-order modes 159 beam propagation of 160 circular symmetry 159 angle of divergence 159 beam radius 159 waist radius 159 rectangular symmetry 159 angle of divergence 159 beam radius 159 waist radius 159Historical remarks coherence 49 Hole burning, spatial 184, 196 Hollow waveguide circular 205 rectangular 202 Hologram computer-generated (CGH) 236, 244 digital 267 double-exposed 265Fourier 264 lensless 269 quasi- 270 Fresnel 264, 270 image-plane 264, 271 lensless Fourier 269 quasi-Fourier 270 spectrum of 269 Holographic interferometry 221, 263 digital 260, 272 for dynamic deformations 272 principle of 265 pulsed digital 276 Holography digital 266 principle of 266 double-exposure 265 principle of 263 real-time 265 Homogeneous line broadening 187, 188 Homogeneously broadened laser 187 Hong–Ou–Mandel interferometer 60 Hopkins 222 Huygens 221 Hybrid mode 156

Image-plane hologram 264, 271 Imaging integrator, beam shaping 134 In-vivo biological tissue 276 InAs electro-optics 93 Incidence, oblique interferometer grating 236 prismatic 234 interferometry 234 Inclination, equal, fringes of 225 Incoherence 52 Independent sources 56 Induced superradiation 73 Inequality Bell 5, 31, 60 Heisenberg 8 Information, quantum 4,36 Inhomogeneous line broadening 187, 190 Inhomogeneously broadened laser 187 Insulator, Mott 36 Integral Collins, beam shaping 130 diffraction, Rayleigh–Sommerfeld 267equation, Kirchhoff 151 Integration, beam-, beam shaping 133Integrator beam beam shaping 128, 134, 136 multifaceted, beam shaping 135 flies-eye, beam shaping 135 imaging, beam shaping 134 Intensity correlation spatial 58 temporal 59 data recording five-frame 233 four-frame 233 interference 58 interferometer 4 interferometry, stellar 58 modulator, acousto-optic 106 mutual, beam shaping 132 saturation 185 Interaction anisotropic 102 atom–field 14 atom-reservoir 22 coherent 56 electric field-dipole 14 Hamiltonian 15 isotropic 101 operator, dipole 68 Interference 50

basic principles of 222 in a plane-parallel plate 224intensity 58 law, general 224 optical 50 particle 57 photon-pair 61 two-beam 223 vector effect of 226 Interferometer Fabry-Perot 231Fizeau 230 grating 236 Hong–Ou–Mandel 60 intensity 4 Mach–Zehnder 28,230Michelson 228 prismatic 234 Twvman–Green 229white-light 238 Interferometry 221 digital holographic 260, 272 for dynamic deformations 272 pulsed 276 for surface metrology 234heterodyne absolute 243 absolute, with rotational matched grating 243 double 243 for distance measurement 241for velocity measurement 241holographic 221, 263 digital 260, 272 digital, for dynamic deformations 272digital, pulsed 276 principle of 265 intensity, stellar 58 multiwavelength 236 oblique incidence 234phase-locked 234 polarization 239 pulsed digital holographic 276shearing 231 speckle pattern 221, 250, 253 pattern shearing, temporal Fourier-transform (TSPSI) 258 pattern, electronic 254 pattern, temporal (TSPI) 257phase analysis in 256 white-light 237 with a null corrector 244 with adaptive optics 244 Internal modulation 85 Intracavity focus 199

Investigation, endoscopic 276Ion -beam sputtering 116 trap 16 trapping 78 Irreversible amplification 19 Isotropic interaction 101 media Kerr effect in 96 quadratic electro-optic effect in 96 Iteration, Gerchberg–Saxton, beam shaping 130, 134Jaynes-Cummings-Paul (JCP) Hamiltonian 18, 21 model 14,15 JCP (Jaynes-Cummings-Paul) Hamiltonian 18, 21 model 14,15 Jordan, Pascual 3, 27 Jump, quantum 26, 27 KD_2PO_4 (KD^*P) electro-optics 92,94-96 KD^*P (KD_2PO_4) electro-optics 92,94-96 $KDP (KH_2PO_4)$ electro-optics 88, 91, 92, 94, 95 Kerr cell 85 constant 97 electro-optic effect 86,96 in isotropic media 96 John 96 KH₂PO₄ (KDP) electro-optics 88, 91, 92, 94, 95 Kirchhoff integral equation 151 KrF excimer, optical coating system 122 $\mathrm{KTa}_{0.65}\mathrm{Nb}_{0.35}\mathrm{O}_3$ electro-optics 97 Laguerre -Gauss mode 152 polynomials 152 Lamb shift 5,23Willis E. 4,23 $\lambda/4$ plate, rotating 242Laser ArF-excimer, optical coating system 122 atom 35 broadened homogeneously 187 inhomogeneously 187 calorimetric measurement 118

 CO_2 optical coating system 122 waveguide 201 component advanced 122 critical parameters of 117 quality parameters of 116, 117 excimer ArF-, optical coating system 122 F₂-, optical coating system 122 KrF-, optical coating system 122 F_2 -excimer, optical coating system 122 free-electron (FEL) 78 HeNe, optical coating system 122 homogeneously broadened 187 -induced damage 120 threshold (LIDT) 117, 120 inhomogeneously broadened 187 KrF-excimer, optical coating system 122light 55 Nd:YAG, optical coating system 122 power 186 slab waveguide 201, 210 waveguide Laue 222 M. von 54 Law ABCD 154, 157 for the q-parameter 159 interference, general 224 radiation, Planck's 54 Lax, Melvin 4 Leith and Upatnieks 263 Length, coherence 55, 222 beam shaping 138 longitudinal 53, 55 transverse 53, 55 beam shaping 128 Lens resonator 191, 192 beam radius at mirror 193 at the principal planes 193half angle of divergence 193 location of beam waist 193 waist radii 193 Lensing, thermal 190 coefficient 191 Lensless Fourier hologram 269 Lenslet array, beam shaping 134 Lewis, Gilbert N. 4 LIDT (laser-induced damage threshold) 117, 120 Light chaotic 55, 62

classical 49 laser 55 modulator acousto-optic 85 electro-optic 85 non-classical 50, 59 propagation 86 white interferometer 238interferometry 237LiNbO₃ acousto-optics 98, 102, 105, 108 electro-optics 88,92–96 :MgO electro-optics 96 Line broadening homogeneous 187, 188 inhomogeneous 187, 190 Linear electro-optic effect 86,87 resonator stable 147 unstable 147 Liouville operator 18,20 Liquid ⁴He, superfluidity of 34 -crystal phase shifter 239 LiTaO₃ electro-optics 88, 92, 95, 96 Lithographic technique, beam shaping 140, 141 Lithography, gray-scale, beam shaping 141Local realistic theory (LRT) 31 193Location of beam waist, lens resonator London, Fritz 34 Longitudinal coherence length 53, 55 Loss calculation, waveguide resonator 208 diffraction 167 factor 151, 152, 165, 166 per round trip 176 per round trip, unstable resonator 174 optical 117 resonator 209 unstable 176 round-trip, waveguide resonator 207 Louisell, William H. 4 Lowering operator 68 LRT (local realistic theory) 31 Mach–Zehnder interferometer 28,230 Magnetic trap 34 Magneto-optical trap 34 Magnetron sputtering technique 122 Magnification 173, 174 curves of constant 175 Mandel 222 Q-parameter 13

Manufacturing of beam-shaping elements 140 rapid 122 Maser, one-atom 4,20 Mask absorbing, beam shaping 127 phase, beam shaping 127 Master equation 18, 21, 22, 69 system 4 Material acousto-optic 105 electro-optic 95 Matrix ABCD, beam shaping 129 coherence 240 coherency 228 rav-transfer 149 thin-film 112 Matter waves, coherent 35 Maximum beam parameter product 195 extraction efficiency 188, 198output power 187 Maxwell -Bloch equations (MB) 71 equations, quantum (QMB) 71 system, undamped 72 James Clerk 3 MB (Maxwell–Bloch) equations 71 McCall and Hahn, theorem of 72 Measurement calorimetric 118 deformation 253, 260 displacement 253 distance, heterodyne interferometry for 241stress 253 velocity, heterodyne interferometry for 241vibration 253, 262 Media, isotropic Kerr effect in 96 quadratic electro-optic effect in 96 Membrane mirror 248 deformable 247 Meter, prototype of 221 Metrology high-precision 221 surface, interferometry for 234 MgO:LiNbO₃ electro-optics 96 Michelson 221 interferometer 228 Mirau arrangement 237 Mirror beam radius at, lens resonator 193

beam shaping 141 high-reflecting 120 membrane 248 deformable 247 moving 85 multifaceted, beam shaping 128 output-coupling 174 variable-reflectivity (VRM) 180, 181 Misalignment angle of 168 resonator 169 sensitivity 167, 170 curves of constant 170 Mixed state 9 Mixing four-wave 36 two-photon 60 Mode decomposition 7 donut 156 expansion 5 field 50 fundamental 154, 157 operation 169 waveguide 205 Gauss-Hermite 152, 153 Gauss-Laguerre 152 higher-order 159 beam propagation 160 circular symmetry 159circular symmetry, angle of divergence 159circular symmetry, beam radius 159circular symmetry, waist radius 159rectangular symmetry 159 rectangular symmetry, angle of divergence 159rectangular symmetry, beam radius 159 rectangular symmetry, waist radius 159hybrid 156 -locking, active 85 multi-, operation 170 transverse 194 normal 6 property, waveguide resonator 208structure resonator, unstable 176 transverse 152 transverse 155 TEM_{mn} 154 waveguide 203 fundamental 205 Model Jaynes–Cummings–Paul (JCP) 14, 15 JCP (Jaynes–Cummings–Paul) 14, 15 Schell, beam shaping 138

Modulation amplitude 90,94 bandwidth 103 external 85 internal 85 phase 88,94 polarization 89 sinusoidal 91 Modulator 85 acousto-optic 98,106Bragg 104 intensity 106 amplitude 94 Bragg, acousto-optic 104 device 87 electro-optic 96 intensity, acousto-optic 106 light acousto-optic 85 electro-optic 85 phase 94 traveling-wave 91 Mollow Benjamin R. 26 spectrum 25 triplet 4 Moment, dipole, operator, total 68 Momentum angular, operator 68 operator 8 Monolithic nonplanar ring oscillator, diode-laserpumped 200 Motion, quantum 16 Mott insulator 36 Moving mirror 85 phase grating 98 Multiatomic extended system 70 Multifaceted beam integrator, beam shaping 135 mirror, beam shaping 128 Multimode operation 170 transverse 194 Multiwavelength interferometry 236Mutual intensity, beam shaping 132 $\mathbf{N}\mathbf{a}$ BE-condensate, superradiation 74 vapor, superradiation 74 Nd:glass flashlamp-pumped 191 Nd:YAG diode-pumped 191 monolithic nonplanar ring oscillator 200flashlamp-pumped 191 laser, optical coating system 122

superradiation 74 Nd:YAP flashlamp-pumped 191 Nd:YLF diode-pumped 191 flashlamp-pumped 191 Nd:YVO₄ diode-pumped 191 Negative-branch resonator 175 ring, unstable 200 unstable 172 Newton ring 221 NH_3 superradiation 74 $(NH_4)D_2PO_4$ (AD^*P) electro-optics 92 $(NH_4)H_2PO_4$ (ADP) electro-optics 88, 92, 95, 96 Nitrobenzene $(C_6H_5NO_2)$ electro-optics 97 Nitrotoluene $(C_5H_7NO_2)$ electro-optics 97 Non-classical light 50, 59 Non-classicality 13 Non-destructive testing 221 Nonlinear atom optics 36 Nonplanar ring monolithic, diode-laser-pumped 200 resonator 200 Normal mode 6 Null corrector 245 interferometry with 244 Number Fresnel 166 effective 164 occupation 54 photon 63

O_2

centers in KCl, superradiation 74 in KCl, superradiation 74 Object, vibrating, analysis of 257 Oblique incidence interferometer grating 236 prismatic 234 interferometry 234 Occupation number 54 Off-axis confocal unstable resonator 212One-atom maser 4,20 Operation fundamental-mode 169 multimode 170 transverse 194 unidirectional 199 Operator angular momentum 68

annihilation 10 dipole 68 interaction 68 moment, total 68 interaction, dipole 68 Liouville 18,20 lowering 68 momentum 8 angular 68 Pauli 15 position 8 raising 68 total dipole moment 68 Optical coating 114 production of 114 system 122 system, ArF-excimer 122 system, CO₂-laser 122 system, F₂-excimer 122 system, HeNe-laser 122 system, KrF-excimer 122 system, Nd:YAG-laser 122 diode 196 interference 50 laser component, quality parameters of 116, 117 loss 117 resonator 147 classification 149 eigenmodes 151 resonance condition 152 testing 221, 228 Optics adaptive 246 interferometry with 244 atom 14, 32, 35 nonlinear 36 quantum 3 Optimum output coupling 187 reflectance 188 Oscillator, ring, monolithic nonplanar, diode-laserpumped 200 Output coupler, confined 199 coupling mirror 174 optimum 187 energy per pulse 189 power maximum 187 resonator, stable 184Paradox

Einstein, Podolsky, Rosen (EPR) 60 EPR (Einstein, Podolsky, Rosen) 60

Parametric down-conversion 60 Partial coherence 52 Partially coherent beam, beam shaping 132 Particle interference 57 Path integral formulation 3 Pattern, speckle dark spot 253 interferometry 221, 250, 253electronic 254 shearing, temporal, Fourier-transform (TSPSI) 258temporal (TSPI) 257 Paul trap 16 Pauli operator 15 Wolfgang 27 $Pb_{0.93}La_{0.07}(Zr_{0.65}Ti_{0.35})0_3$ electro-optics 97 PbMoO₄ acousto-optics 98, 105, 106 Perot –Fabry interferometer 231 Phase analysis by phase stepping 256 by spatial phase shifting 256in speckle interferometry 256cam, set-up 234 -diffusing element, beam shaping 138element diffractive, beam shaping 140 non-quadratic, beam shaping 127 evaluation technique, electronic 231 grating 99 moving 98 -locked interferometry 234 mask, beam shaping 127 modulation 88,94 modulator 94 shift, Gouy 158 shifter, liquid-crystal 239shifting fringe analysis by 232 spatial, phase analysis by 256space distribution 19 state 9 stepping, phase analysis by 256unwrapping, temporal 278 Phenomenological damping term 71 Photoelastic effect 86,98 Photon 4.57 annihilation of 15 antibunching 4,59 bunching 59 counting statistics 62 creation of 15

distribution 10 function 62 number 63 state 9,17 state preparation 16pair entangled 60 interference 61 statistics Boltzmann 12 steady-state 21 Planar ring resonator 197 Planck Max 3 radiation law 54 Plane image-, hologram 264, 271 principal, beam radius at, lens resonator 193 Plane-parallel plate, interference in 224 Plasma, electron-hole 78 Plate $\lambda/4$, rotating 242 plane-parallel, interference in 224 Pockels electro-optic effect 86,87 Friedrich Carl Alwin 87 Podolsky, Boris 30 Poisson distribution 10, 62, 63 Polarization interferometry 239 modulation 89 TE-227TM- 227 Polarizer sample compensator analyzer (PSCA) 239 thin-film 123 Polynomial Hermite 153 Laguerre 152 Position of waist 157operator 8 Positive-branch resonator 174 unstable 172 Potential, vector 5 Power available 187 density distribution, beam shaping 127 laser 186 output maximum 187 resonator, stable 184 Preparation, state photon number 16Schrödinger cat 17

Principal plane, beam radius at, lens resonator 193Principle of digital holography 266of holographic interferometry 265of holography 263 polarizer-sample-compensator-analyzer (PSCA) 239PSCA (polarizer sample compensator analyzer) 239Prismatic interferometer 234 Probability, transition 69 Product, beam parameter 160, 162, 179, 182, 195 maximum 195 Production of optical coatings 114 Profile reflectivity, super-Gaussian 181 tophat, beam shaping 128 Projection technique, beam shaping 127Propagation beam 157 factor 160, 161 of higher-order modes 160 resonator, unstable 172 of light 86 Property mode, waveguide resonator 208of speckles 251 resonator 148 Prototype of the meter 221 PSCA (polarizer sample compensator analyzer) 239Pulse 2π secanthyperbolic 72 small-area 76 superradiant 69 tipping 76 Pulsed digital holographic interferometry 276 Pure coherence spatial 222 temporal 222 state 9 q-parameter 160 ABCD law for 159 Q-switching 85 QED (quantum electrodynamics) 3 QMB (quantum Maxwell–Bloch) equations 71 Quadratic electro-optic effect 86,96 in isotropic media 96 Quality beam 161 resonator, unstable 178 parameter of optical laser components 116, 117 surface 117 Quantity, classical stochastic 71

Quantization, field 5,7 Quantum communication 4 computing 4 cryptography 37, 38 electrodynamics (QED) 3 gas, degenerate fermionic 4 information 4,36 jump 26,27 method 19 Maxwell–Bloch (QMB) equations 71 motion 16 optics 3 state engineering 16 teleportation 36 Quarter-wave optical thickness (QWOT) 112 Quasi-Fourier hologram 270 QWOT (quarter-wave optical thickness) 112Rabi frequency 15, 16 Radiation field damped 18 energy of 7 law, Planck's 54 Radius beam at mirror, lens resonator 193 at the principal planes, lens resonator 193 Gaussian 153 higher-order modes, circular symmetry 159 higher-order modes, rectangular symmetry 159Gaussian beam 153 of curvature 158 waist 157 higher-order modes, circular symmetry 159 higher-order modes, rectangular symmetry 159lens resonator 193 Raising operator 68 Raman–Nath diffraction 100 regime 100 Ramsey set-up 17 Range, Rayleigh 157, 159 Rapid manufacturing 122 Ray transfer matrix 149 Rayleigh range 157, 159 scattering 77 cooperative 77 –Sommerfeld diffraction 267 integral 267 \mathbf{Rb} atomic beam, superradiation 74 superradiation 74

Real-time holography 265 Recording, intensity data five-frame 233 four-frame 233 Rectangular symmetry eigenmodes 153 higher-order modes 159 higher-order modes, angle of divergence 159higher-order modes, beam radius 159 higher-order modes, waist radius 159 resonance frequencies 163 resonator, unstable 179 waveguide, hollow 202 Reflectance optimum 188 spectrum 113 thin-film 113 Reflectivity profile, super-Gaussian 181 variable - mirror (VRM) 180, 181 Regime Bragg 101 Raman–Nath 100 Relation commutation 8 Woodruff and Ehrenreich 104 Reservoir theory 4,18 Resonance condition of the optical resonator 152 Feshbach 36 fluorescence 24 frequency 163 circular symmetry 163 rectangular symmetry 163 Resonator basic properties 148 configuration, waveguide 206confocal 175 eigenmode 153 lens 191, 192 beam radius at mirror 193 beam radius at the principal planes 193half angle of divergence 193 location of beam waist 193 waist radii 193 linear stable 147 unstable 147 loss 209misalignment 169 negative-branch 175 nonplanar, ring 200 optical 147 classification 149 eigenmodes 151

resonance condition 152 planar, ring 197 positive-branch 174 ring 148, 196 nonplanar 200 planar 197 unstable 198 unstable, asymmetric confocal 199 unstable, negative-branch 200 stable 150 aperture-limited 164 linear 147 output power 184 unconfined 151 unconfined, stable 151unstable 150, 172 applications of 183 beam propagation 172 beam quality 178 characterization of 172 linear 147 loss 176 loss factor per round trip 174 mode structure 176 negative-branch 172off-axis confocal 212 positive-branch 172 rectangular symmetry 179ring 198 ring, asymmetric confocal 199 ring, negative-branch 200 waveguide 148, 201, 205 configuration 206 loss calculations 208 mode properties 208round-trip loss 207 with one aperture 164 with two apertures 166 Response time 88 Retardation, dynamic 89 Retarder, dynamic, configuration 90 Retherford, Robert C. 23 Ring annular 178 monolithic, nonplanar, diode-laser-pumped 200 Newton 221 nonplanar monolithic, diode-laser-pumped 200 resonator 200 resonator 148, 196 nonplanar 200 planar 197 unstable 198 unstable, asymmetric confocal 199 unstable, negative-branch 200Ringing, Burnham–Chiao 75

Rosen, Nathan 30 Rotating grating 242 $\lambda/4$ plate 242 wave approximation (RWA) 15,71 Rotational matched grating 243 Rotator, Faraday 197 Round-trip loss factor 176 waveguide resonator 207 Rugate filter 124 Running wave 6 RWA (rotating wave approximation) 71 Rydberg atom 70 Saturation intensity 185 Scanning, beam 127 Scattering Antistokes Raman, coherent (CARS) 56 coherent Antistokes Raman (CARS) 56 Stokes Raman (CSRS) 56 Rayleigh 77 Stokes Raman, coherent (CSRS) 56 total 119 Schell model, beam shaping 138 Schrödinger cat state preparation 17 equation 15 Erwin 26 picture 7 Scraper 199 Scully, Marlan O. 4 Secanthyperbolic 2π pulse 72 Second-order correlation function 51, 58 strain tensor 86 Self-induced transparency (SIT) 72 Sensitivity, misalignment 167, 170 curves of constant 170 Sensor Hartmann 246 Shack-Hartmann 246 Set-up phase cam 234 Ramsey 17 Shack-Hartmann 246 sensor 246 Shaping beam 127, 138 ABCD-matrix 129 ABCD-transformation 127 absorbing mask 127 Archimedic spiral curve 140 beam integration 133 beam integrator 128, 134, 136

beam integrator, multifaceted 135 beam splitting 139 beam transformation 130 beam, partially coherent 132 beam-transfer analysis 132 beamlet 128 coherence 138 coherence length 138 coherence length, transverse 128 coherent beam 130 Collins integral 130 diamond turning 141 element, manufacturing of 140 element, phase-diffusing 138 Fourier transformation 135 Gerchberg–Saxton iteration 130, 134 gray-scale lithography 141 GRIN technique 140 imaging integrator 134 integrator, flies-eye 135 integrator, imaging 134 intensity, mutual 132 lenslet array 134 lithographic technique 140, 141 lithography, gray-scale 141 mirror 141 multifaceted mirror 128 phase element, diffractive 140 phase element, non-quadratic 127 phase mask 127 power density distribution 127 projection technique 127 Schell model 138 smoothing parameter $\rho = 139$ structure function d = 127technique 129 technique, lithographic 140, 141 tophat distribution 133 tophat profile 128 transformation efficiency 131 Wigner distribution function 132 beamlet 127, 137, 140 Shearing interferometry 231 Shift Lamb 5,23 phase, Gouy 158 Shifter, phase, liquid-crystal 239Shifting, phase fringe analysis by 232 spatial, phase analysis by 256Shutter 85 Signal, analytic 50 Silica, fused acousto-optics 105 Single-mode field 50

Sinusoidal modulation - 91 SiO_2 acousto-optics 98 $/Ta_2O_5$ 113 SIT (self-induced transparency) 72Skribanowitz 67 Slab waveguide laser 201,210 Slit, double, Young's 50Slope efficiency 187 Slowly varying envelope approximation (SVEA) 71Small area pulse 76 Small-signal gain coefficient g_0 184 Smoothing parameter ρ , beam shaping 139Sommerfeld -Rayleigh diffraction 267 integral 267 Source chaotic 55 EPR (Einstein, Podolsky, Rosen) 37 independent 56 Spatial coherence 53, 223 pure 222 hole burning 184, 196 intensity correlation 58 phase shifting, phase analysis by 256Speckle application 253 contrast 251 Gaussian 252 interferometry, phase analysis in 256pattern dark spot 253 interferometry 221, 250, 253 interferometry, electronic 254 interferometry, temporal (TSPI) 257shearing interferometry, temporal Fouriertransform (TSPSI) 258 properties of 251 Spectral transmission band 114 Spectrum Mollow 25 of a hologram 269 reflectance 113 Sphere Bloch 72 Coblentz 119 Ulbricht 119 Spiral curve, Archimedic, beam shaping 140 Splitting beam, beam shaping 139 Zeeman 242 Spontaneous emission 5,24 amplified (ASE) 67 cooperative 67

Spot, dark, in a speckle pattern 253 Sputtering ion-beam 116 technique, magnetron 122 Squeezed state 4, 11, 12 vacuum 23 Squeezing 5 Stability 117 boundary 150 diagram 151 Stabilization, amplitude 59 Stable resonator 150 aperture-limited 164 linear 147 output power 184 unconfined 151 Standing wave 6 State coherent 10,56 entangled 4 field 8 Fock 9, 16, 17, 29 Glauber 56 mixed 9 phase 9 photon number 9,17 preparation 16 pure 9 squeezed 4, 11, 12 steady, condition 185 thermal 12 trapping 22 vector 9 Statistics photon counting 62 steady-state 21 sub-Poissonian 21,63 super-Poissonian 21,63 Steady-state condition 185 field distribution 152 photon statistics 21Stellar intensity interferometry 58 Stepping, phase, phase analysis by 256 Stitching technique 249 Stochastic quantity, classical 71 Stokes 226 Raman scattering, coherent (CSRS) 56 Strain tensor, second-order 86 Stress measurement 253 Structure function d, beam shaping 127 mode resonator, unstable 176

transverse 152 ${\it Sub-Poissonian}$ behavior 12 statistics 21,63 Subradiance 69 Super-Gaussian reflectivity profile 181 Super-Poissonian statistics 21,63 Superfluid 36 Superfluidity of liquid ⁴He 34 Superfluorescence 67 Superluminescence 67 Superradiance 67 Superradiant pulse 69 Superradiation atomic beam. Rb 74 Ba ions, pair of 74 $C_{12}H_{10}$ in diphenyl 74 CH₃CN 74 CH_3F 74 Cr: ruby 74 Cs vapor 74 Eu vapor 74 Ga vapor 74 HF 74 induced 73 Na BE-condensate 74 vapor 74 Nd:YAG 74NH₃ 74 O_2 centers in KCl 74 in KCl 74 Rb 74 atomic beam 74 time 70 74Tl vapor triggered 73vapor Cs 74 Eu 74 Ga 74 Na 74 Tl 74 Surface aspheric 246 metrology, interferometry for 234 quality 117 SVEA (slowly varying envelope approximation) 71Switching, Q- 85 Symmetric, circular waveguide 209Symmetry circular eigenmodes 152 higher-order modes 159

higher-order modes, angle of divergence 159higher-order modes, beam radius 159 higher-order modes, waist radius 159 resonance frequencies 163 rectangular eigenmodes 153higher-order modes 159 higher-order modes, angle of divergence 159higher-order modes, beam radius 159 higher-order modes, waist radius 159 resonance frequencies 163 resonator, unstable 179 System cooperative 77 extended multiatomic 70 master 4 Maxwell–Bloch, undamped 72 optical coating 122 ArF-excimer 122 CO₂-laser 122 F_2 -excimer 122 HeNe-laser 122 KrF-excimer 122 Nd:YAG-laser 122 thin-film 111 Ta_2O_5/SiO_2 113 TE-polarization 227 Technique beam shaping 129 Fourier-transform 232 GRIN, beam shaping 140lithographic, beam shaping 140, 141 magnetron sputtering 122 phase evaluation, electronic 231 projection, beam shaping 127 sputtering, magnetron 122 stitching 249 time-average 265 Technology, thin-film 111 future trends in 123 Teleportation 5 quantum 36 TEM (transverse electro-magnetic) 153 TEM_{mn} transverse mode 154 Temporal coherence 52pure 222 Fourier-transform speckle pattern shearing interferometry (TSPSI) 258 intensity correlation 59 interferometry Fourier-transform speckle pattern shearing (TSPSI) 258 speckle pattern (TSPI) 257phase unwrapping 278 speckle pattern interferometry (TSPI) 257

Tensor, strain, second-order 86 TeO_2 acousto-optics 98, 105, 106, 108 Term, damping, phenomenological 71 Testing non-destructive 221 optical 221, 228 Theorem of McCall and Hahn 72 van Cittert–Zernike 53, 222 Wiener-Khintchine 24,52 Theory coherence 51 higher-order 222 local realistic (LRT) 31reservoir 18 Thermal evaporation 121 lensing 190 coefficient 191 state 12 Thickness, equal, fringes of 225 Thin film absorption 117 matrices M_m 112 polarizer 123 reflectance 113 system 111 technology 111 future trends in 123 Threshold damage 94 laser-induced (LIDT) 117, 120 Time -average technique 265charging, electrode 91coherence 56, 222 real-, holography 265 response 88 superradiation 70 TiO_2 acousto-optics 105 Tipping pulse 76 Tissue, biological, in-vivo 276Tl vapor superradiation 74 $\mathrm{Tl}_3\mathrm{AsS}_4$ 105acousto-optics Tl_3AsSe_3 acousto-optics 105TM-polarization 227Tophat distribution, beam shaping 133profile, beam shaping 128 Total dipole moment operator 68

scattering 119 Transducer 103 Transfer beam-, analysis, beam shaping 132 function 117 Transform, Fourier fast (FFT) 268 technique 232 Transformation ABCD, beam shaping 127 beam, beam shaping 130 efficiency, beam shaping 131 Fourier, beam shaping 135 Transition probability 69 Transmission band, spectral 114 Transparency, self-induced (SIT) 72Transverse coherence length 53, 55 beam shaping 128 electro-magnetic (TEM) 153 mode 155structure 152 TEM_{mn} 154 multimode operation 194 Trap 16ion magnetic 34 magneto-optical 34 Paul 16 Trapping atoms 78 ions 78 states 22 Traveling-wave modulator 91 Triggered superradiation 73 Trip, round, loss factor 176 Triplet, Mollow 4 TSPI (temporal speckle pattern interferometry) 257TSPSI (temporal Fourier-transform speckle pattern shearing interferometry) 258 Turning, diamond, beam shaping 141 Twiss and Brown 49,58 Richard 4 Two-beam interference 223Two-level atom 15 Two-photon mixing 60 wave packet 60 Twyman 221 -Green interferometer 229Ulbricht sphere 119 Unconfined stable resonator 151 Undamped Maxwell–Bloch system 72

Unidirectional operation 199 Unstable resonator 150, 172applications of 183 beam propagation 172 beam quality 178 characterization of 172 linear 147 loss 176loss factor per round trip 174 mode structure 176 negative-branch 172 off-axis confocal 212 positive-branch 172 rectangular symmetry 179ring 198 ring, asymmetric confocal 199 ring, negative-branch 200 ring resonator 198 asymmetric confocal 199 negative-branch 200 Unwrapping, phase, temporal 278 Upatnieks and Leith 263 Vacuum fluctuation 5 squeezed 23 van Cittert 222 –Zernike theorem 53, 222 Variable reflectivity mirror (VRM) 180, 181 Vector effect of interference 226 potential 5 state 9 Velocity measurement, heterodyne interferometry for 241 Vibrating object, analysis of 257 Vibration analysis 273 Fizeau arrangement 263 measurement 253, 262 Visibility 52 Voltage, half-wave 89,90 Volume coherence - 54 -index grating 99 VRM (variable-reflectivity mirror) 180, 181 Waist beam, location of, lens resonator 193position 157 radius 157 higher-order modes, circular symmetry 159 higher-order modes, rectangular symmetry 159lens resonator 193 Walk-off 92

Water (H_2O) electro-optics 97 Wave converging 173 diverging 173 matter, coherent 35 packet, two-photon 60 –particle duality 28 rotating, approximation (RWA) 71 running 6 standing 6 traveling-, modulator 91 Waveguide circular hollow 205 symmetric 209 CO_2 laser 201 fundamental mode 205hollow circular 205 rectangular 202 laser CO_2 201slab 201, 210 203mode fundamental 205 rectangular hollow 202 resonator 148, 201, 205 configuration 206 loss calculations 208 mode properties 208 round-trip loss 207 slab, laser 201, 210symmetric, circular 209 Wavelength de Broglie 57 multi-, interferometry 236 Weisskopf and Wigner 67 –Wigner decay 23 "Which-path" detector 28 White light interferometer 238 interferometry 237 Wiener-Khintchine theorem 24, 52 Wigner and Weisskopf 67 distribution function, beam shaping 132 Wolf 222 Woodruff and Ehrenreich relation 104 Young

double slit 50 experiment 221 Thomas 28, 221 Zeeman splitting 242 Zehnder 230 Zernike 222 -van Cittert theorem 222 ZnS $\beta-$ electro-optics 95 electro-optics 93 ZnSe electro-optics 95 ZnTe electro-optics 95