Continuum Mechanics

George Backus

Insitute for Geophysics and Planetary Physics

University of California, San Diego

S

Samizdat

Press

Continuum Mechanics

George Backus

S

Samizdat Press Golden · White River Junction

Published by the Samizdat Press

Center for Wave Phenomena

Department of Geophysics

Colorado School of Mines

Golden, Colorado 80401

and

New England Research

76 Olcott Drive

White River Junction, Vermont 05001

Contents

Logic

$\Rightarrow$ "implies" (e.g. $A \Rightarrow B$ where A and B are sentences.)

$\Leftrightarrow$ "implies and is implied by" (e.g. $A \Leftrightarrow B$ where A and B are sentences.)

iff "if and only if" — same as $\Leftrightarrow$

:= "is defined as"

$\forall$ "for every"

$\exists$ "there exists at least one"

$\exists\mathbf{1}$ "there exists exactly one"

$\ni$ "such that"

Sets (same as "Classes")

$(a_1, \ldots, a_6)$ is an (ordered) sequence of objects. $(a_1, \ldots, a_n)$ is also called an ordered n-tuple. Order is important, and $(a_1, a_2, a_3) \neq (a_2, a_1, a_3)$. However, duplication, e.g. $a_1 = a_2$, is permitted.

$\{a_1, \ldots, a_6\}$ is a set of objects. Order is irrelevant, and $\{a_1, a_2, a_3\} = \{a_2, a_1, a_3\} = \{a_3, a_1, a_2\}$, etc. Duplication is not permitted. All a are different.

$(a_1, \ldots, \not{a}_4, \ldots, a_6) := (a_1, a_2, a_3, a_5, a_6)$.

$\{a_1, \ldots, \not{a}_4, \ldots, a_6\} := \{a_1, a_2, a_3, a_5, a_6\}$.

$\emptyset$ is the empty set, the set with no objects in it.

$\mathcal{R}$:= set of all real numbers.

$\mathcal{R}^{\circ+}$:= set of all non-negative real numbers.

$\mathcal{R}^{+}$:= set of all positive real numbers.

$\mathcal{C}$:= set of all complex numbers.

$\{1, \ldots, n\}$:= set of integers from 1 to n inclusive.

$\{x : p(x)\}$ is the set of all objects x for which the statement $p(x)$ is true. For example, $\{x : x = 2y \text{ for some integer } y\}$ is the set of even integers. It is the same as $\{2x : x \text{ is an integer}\}$.

$\in$ "is a member of." Thus, "$a \in A$" is read "the object a is a member of the set A. The phrase "$a \in A$" can also stand for "the object a which is a member of A"

$\notin$ "is not a member of"

$\subseteq$ "is a subset of." $A \subseteq B$ means that A and B are sets and every member of A is a member of B.

$\subset$ "is a proper subset of." $A \subset B$ means that $A \subseteq B$ and $A \neq B$ (i.e., B has at least one member not in A).

$A \cap B$:= the set of all objects in both A and B. (Read "A meet B.")

$A \cup B$:= the set of all objects in A or B or both. (Read "A join B.")

$A \backslash B$:= the set of all objects in A and not B. (Read "A minus B.")

Cartesian products

$A \times B$:= the set of all ordered pairs of (a, b) with $a \in A$ and $b \in B$.

$A_1 \times \ldots \times A_n$:= the set of all ordered n-tuples $(a_1, \ldots, a_n)$ with $a_1 \in A_1$, $a_z \in A_2, \ldots, a_n \in A_n$.

$\times^n A$:= the set of all ordered n tuples $(a_1, \ldots, a_n)$ with $a_1, a_2, \ldots, a_n$ all members of A. Same as $A \times \ldots \times A$ (n times).

Functions

Function is an ordered pair (D_f, f). D_f is a set, called the "domain" of the function, and f is a rule which assigns to each $d \in D_f$ an object $f(d)$. This object is called the "value" of f at d. The function (D_f, f) is usually abbreviated simply as f. In the expression $f(d)$, d is the "argument" of f.

Range of function is a set R_f consisting of all objects which are values of f. Two equivalent definitions of R_f are

$$\begin{aligned} R_f &:= \{x : \exists d \in D_f \ni x = f(d)\} \\ &\text{or} \\ R_f &:= \{f(d) : d \in D_f\}. \end{aligned}$$

$f = g$ for functions means $D_f = D_g$ and for each $d \in D$, $f(d) = g(d)$.

$f(W)$ for a function f and a set W is defined as $\{f(w) : w \in W\}$. If $W \cap D_f = \emptyset$ then $f(W) = \emptyset$.

$d \mapsto f(d)$ is a shorthand way of describing a function. For example, if $D_f = \mathcal{R}$ and $f(x) = x^2$ for all $x \in \mathcal{R}$, we can describe f by saying simply $x \mapsto x^2$ under f. In particular, $2 \mapsto 4$ under f.

$f|W$ is the "restriction" of f to W. It is defined whenever W is a subset of D_f. It is the function (W, g) such that for each $w \in W$, $g(w) = f(w)$. Its domain is W.

"Extension." If g is a restriction of f, then f is an extension of g to D_f.

$f(\cdot, v_0)$, $f(u_0, \cdot)$. Suppose U and V are sets and f is a function whose domain is $U \times V$. Suppose $u_0 \in U$ and $v_0 \in V$. Define two functions g and h as follows:

$$\begin{aligned} D_g &= U \text{ and}, \forall u \in U,\ g(u) = f(u, v_0) \\ D_h &= V \text{ and}, \forall v \in V,\ h(v) = f(u_0, v). \end{aligned}$$

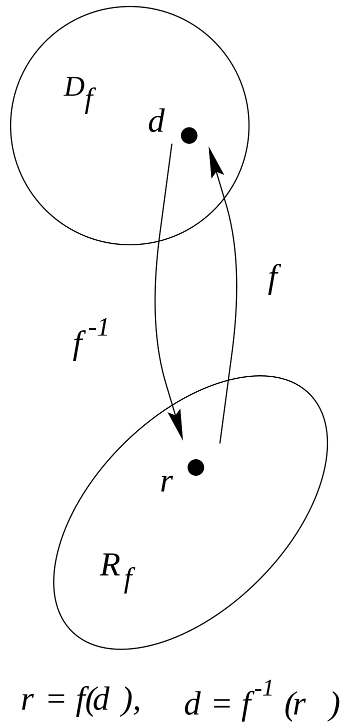


Figure D-1:

Then g is often written $f(\cdot, v_0)$, and h is often written $f(u_0, \cdot)$. The dot shows where to put the argument of the function.

$\sum_{a \in A} f(a)$ is defined when $A \subseteq D_f$ and when $f(A)$ consists of objects which can be added (e.g., real numbers or vectors). The symbol stands for the sum of all the values $f(A)$ for $a \in A$. The number of terms in the sum is the number of objects in A, so A must be finite unless some kind of convergence is assumed.

I_U**,** the identity function on the set U, has domain U and effect $u \mid\rightarrow u$. That is, $I_U(u) = u$, $\forall u \in U$.

"Invertible." A function f is invertible if $f(d) = f(d') \Rightarrow d = d'$, $\forall d, d' \in D_f$. That is, if $r \in R_f$, there is exactly one $d \in D_f$ such that $r = f(d)$. In other words, for each $r \in R_f$, the equation $f(d) = r$ has exactly one solution $d \in D_f$.

"Inverse," f^{-1}**.** If f is invertible, its inverse f^{-1} is defined to be the function whose domain is R_f and such that for each $r \in R_f$, $f^{-1}(r)$ is that unique $d \in D_f$ such that

$f(d) = r$. That is, $d = f^{-1}(r)$ is the unique solution of $r = f(d)$. Note $D_{f^{-1}} = R_f$, $R_{f^{-1}} = D_f$. See Fig. D-1.

For example, if the domain is $\mathcal{R}^+$, both the functions $x \mapsto x^2$ and $x \mapsto x^3$ are invertible, and their inverses are $x \mapsto x^{1/2}$ and $x \mapsto x^{1/3}$. However, if the domain is $\mathcal{R}$, the function $x \mapsto x^2$ is not invertible.

Note that if f is invertible, so is f^{-1}, and $(f^{-1})^{-1} = f$.

Mappings

A Mapping is an ordered triple (U, V, f) in which U and V are sets, f is a function, $D_f = U$, and $R_f \subseteq V$. We say that "the function f maps U into V." The mapping is often abbreviated simply as f if the context makes clear what U and V are.

$f : U \to V$ is the usual way of writing the mapping (U, V, f) if one wants to note explicitly what U and V are. The symbol "$f : U \to V$" is read "the mapping f of U into V." It can also stand for the sentence "f is a function which maps U into V." In this latter usage it is equivalent to "$f(U) \subseteq V$."

$F(U \to V)$:= the set of all functions mapping U into V.

Injective. If f is invertible, $f : U \to V$ is an "injective" mapping or an "injection."

Surjective. If $R_f = V$, $f : U \to V$ is a "surjective" mapping or a "surjection."

Bijection. If $f : U \to V$ is both injective and surjective, it is "bijective," or a "bijection." Note: If $f : U \to V$ is a bijection, so is $f^{-1} : V \to U$.

Composition $g \circ f$: Suppose $f : U \to V$ and $g : V \to W$. Then $g \circ f : U \to W$ is the mapping defined by requiring for each $u \in U$ that $(g \circ f)(u) = g[f(u)]$. The function $g \circ f$ is called the "composition" of g with f. Note that the order in which the functions are applied or evaluated runs backward, from right to left.

Note: If $f : U \to V$ and $g : V \to W$ and $h : W \to X$ then

$$h \circ (g \circ f) = (h \circ g) \circ f. \tag{D-1}$$

Note: If $f : U \to V$ and $g : V \to W$ are bijections, so is $g \circ f : U \to W$, and $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

Note: If $f : U \to V$ is a bijection, $f^{-1} \circ f = I_U$ and $f \circ f^{-1} = I_V$.

Note: Suppose $f : U \to V$ and $g : V \to U$. Then

i) if $g \circ f = I_U$, then f is an injection;

ii) if $f \circ g = I_V$, then f is an surjection;

iii) if $g \circ f = I_U$, and $f \circ g = I_V$ then f is a bijection and $g = f^{-1}$.

Note: If $f : U \to V$ then $f \circ I_U = I_V \circ f = f$.

Permutations

A permutation is a bijection $\pi : 1, \ldots, n \to \{1, \ldots, n\}$. It is an "$n$-permutation" or a "permutation of degree n."

S_n := the set of all n-permutations. It has $n!$ members. The product $\pi\sigma$ of two permutations π and σ is defined to be their composition, $\pi \circ \sigma$. If $e = I_{\{1,\ldots,n\}}$ then $e\pi = \pi e = \pi$ and $\pi^{-1}\pi = \pi\pi^{-1} = e$ for all $\pi \in S_n$. These facts, and the above notes on pg. iv, make S_n a group under the multiplication defined by $\pi\sigma = \pi \circ \sigma$. It is called the symmetric group of degree n. Its identity is e.

(i, j) transposition or interchange. If $i \neq j$ and $1 \leq i, j \leq n$ (i.e. $\{i, j\} \subseteq \{1, \ldots, n\}$) then (i, j) stands for the permutation $\pi \in S_n$ such that $\pi(i) = j$, $\pi(j) = i$, and $\pi(k) = k$ if $k \notin \{i, j\}$. This permutation is called the "transposition" or "interchange" of i and j. Note that $(ij) \circ (ij) = e$, the identity permutation.

Theorem 1 *(See, e.g., Birkhoff and MacLane, A Survey of Modern Algebra.) If $\pi \in S_n$, π is the product of $n - 1$ or fewer transpositions. There are many ways to* *write π as a product of transpositions, but all have the same parity (i.e., all involve an even number of transpositions or all involve an odd number of transpositions.)*

even, odd. If $\pi \in S_n$, π is even or odd according as it can be written as the product of an even or an odd number of transpositions.

sgn π := $+1$ if π is even, := -1 if π is odd.

Note: $\operatorname{sgn}(\pi_1 \circ \pi_2) = (\operatorname{sgn}\pi_1)(\operatorname{sgn}\pi_2)$, and $\operatorname{sgn} e = 1$. $\operatorname{sgn}\pi$ is read "signum of π."

Note: $\operatorname{sgn}\pi = \operatorname{sgn}\pi^{-1}$ because $\pi \circ \pi^{-1} = e$, and $1 = \operatorname{sgn} e = \operatorname{sgn}\pi \circ \pi^{-1} = (\operatorname{sgn}\pi)(\operatorname{sgn}\pi^{-1})$. But $\operatorname{sgn}\pi = \pm 1$.

Arrays

An array of order q is a function f whose domain is

$$D_f = \{1, \ldots, n_1\} \times \ldots \times \{1, \ldots, n_q\}.$$

The array is said to have dimension $n_1 \times n_2 \times \ldots \times n_q$.

$f_{i_1 \ldots i_q}$ If $i_1 \in \{1, \ldots, n_1\}$, $i_2 \in \{1, \ldots, n_2\}$, $\ldots$, $i_q \in \{1, \ldots, x_q\}$, then $f(i_1, \ldots, i_q)$ is often written $f_{i_1, i_2, \ldots i_q}$ or $f^{i_1 \ldots i_q}$ or $f_{i_1}{}^{i_2 i_3}{}_{i_4 \ldots i_q}$ etc. There are 2^q ways to write it. The object $f(i_1, \ldots, i_q)$ is called the entry in the array at location $(i_1, \ldots, i_q)$ or address $(i_1, \ldots, i_q)$. The integers $i_1, \ldots, i_q$ are the "indices", and each can be either a subscript or a superscript. What counts is their right-left order. We will never use symbols like f^{ij}_{kl}.

$\delta^{(n)}$ is the n-dimensional Kronecker delta. It is an $n \times n$ array with $\delta^{(n)}_{ij} = 0$ if $i \neq j$, $\delta^{(n)}_{ij} = 1$ if $i = j$. Usually the ${}^{(n)}$ is omitted, and it is written simply δ_{ij}, or δ^{ij}, or $\delta_i{}^j$, or $\delta^i{}_j$.

$\varepsilon^{(n)}$ is the n-dimensional alternating symbol. It is an $\overbrace{n \times n \times \ldots \times n}^{n \text{ factors}}$ array, with n^n entries, all 0 or $+1$ or -1. It is defined as follows:

$\varepsilon^{(n)}_{i_1 \ldots i_n} = 0$ if any of $i_1, \ldots, i_n$ are equal; i.e., if $\{i_1, \ldots, i_n\} \neq \{1, \ldots, n\}$.

If $\{i_1, \ldots, i_n\} = \{1, \ldots, n\}$, then there is a unique $\pi \in S_n \ni i_1 = \pi(1), \ldots, i_n = \pi(n)$. In that case, $\varepsilon^{(n)}_{i_1 \ldots i_n} = \varepsilon^{(n)}_{\pi(1) \ldots \pi(n)} = \mathrm{sgn}\pi$.

Usually the $\varepsilon^{(n)}$ is written without the n as $\varepsilon_{i_1 \ldots i_n}$, or $\varepsilon^{i_1}{}_{i_2 \ldots i_n}$, or $\varepsilon^{i_1 \ldots i_n}$, etc.

Example: $\varepsilon_{11} = 0$, $\varepsilon_{12} = 1$, $\varepsilon_{21} = -1$, $\varepsilon_{22} = 0$. $\quad (n = 2)$

Example $(n = 3)$**:**

$$\varepsilon_{123} = \varepsilon_{231} = \varepsilon_{312} = 1$$

$$\varepsilon_{213} = \varepsilon_{132} = \varepsilon_{321} = -1$$

all other $\varepsilon_{ijk} = 0$. Thus, $\varepsilon_{111} = 0$, $\varepsilon_{112} = 0$, etc.

Einstein index conventions. i) If an index appears once, without parentheses, in each term of an equation, the equation is true for all possible values of the index. Thus, "$\delta_{ij} = \delta_{ji}$" says that "$\delta_{ij} = \delta_{ji}$" for all $i, j \in \{1, \ldots, n\}$. This is a true statement. "$\varepsilon_{ijk} = \varepsilon_{ijk}$" says that "$\varepsilon_{ijk} = \varepsilon_{jik}$ for $i, j, k \varepsilon \{1, 2, 3, \}$. This is a false statement. The correct statement is $\varepsilon_{ijk} = -\varepsilon_{jik}$.

ii) If an index appears twice in any expression but a sum or difference, a sum over all possible values of that index is understood. Thus, if A_{ij} is an $n \times n$ array, A_{ii} stands for $\sum_{i=1}^n A_{ii}$. If u_i and v_j are first order arrays of dimension n, $u_i v_i$ stands for $\sum_{i=1}^n u_i v_i$, and u_i/v_i stands for $\sum_{i=1}^n u_i/v_i$, but $u_i + v_i$ and $u_i - v_i$ stand for the i'th entries in two arrays, not for $\sum_i (u_i + v_i)$ or $\sum_i (u_i - v_i)$.

iii) If an index appears three or more times in an expression other than a sum or difference, a mistake has been made. Thus, if A_{ijkl} is an $n \times n \times n \times n$ array, A_{iiii} is never written alone. The sum is written out as $\sum_{i=1}^n A_{iiii}$.

iv) Parentheses "protect" an index from the conventions. Thus, $A_{(ii)}$ refers to a particular element in the array, the one at location (ii). And $v_{(i)} = w_{(i)}$ means equality only for the particular value of i under discussion, not for all possible values of i. The element A_{iiii} in an $n \times n \times n \times n$ array is written $A_{(iiii)}$ to make clear that no sum is intended. Strictly speaking, the definition of the Kronecker delta, using these conventions, reads $\delta_{(ij)} = 0$ if $i \neq j$, and $\delta_{(ii)} = 1$ for all i.

An array A of order r is symmetric (antisymmetric) in its p'th and q'th indices if they have the same dimension, $n_p = n_q$, and

$$\begin{aligned} A_{i_1} & \ldots \; i_{p-1} i_p i_{p+1} \ldots i_{q-1} i_q i_{q+1} \ldots i_r \\ &= \pm A_{i_1} \ldots i_{p-1} i_q i_{p+1} \ldots i_{q-1} i_p i_{q+1} \ldots i_r. \end{aligned}$$

(The $+$ is for symmetric A, the $-$ is for antisymmetric A). The array is totally symmetric (antisymmetric) if it is symmetric (antisymmetric) in every pair of indices.

Some properties of δ_{ij} and $\varepsilon_{i_1 \ldots i_n}$ are listed below.

1. $\delta_{ij} = \delta_{ji}$
2. $\delta_{ii}^{(n)} = n$
3. $\delta_{ij} A_{jk_1 \ldots k_p} = A_{ik_1 \ldots k_p}$ for any array A of suitable dimension. $\delta_{ij} A_{k_1 j k_2 \ldots k_p} = A_{k_1 i k_2 \ldots k_p}$, etc. In particular, $\delta_{ij}\delta_{jk} = \delta_{ik}$.
4. $\varepsilon_{i_1 \ldots i_n}$ is totally antisymmetric.
5. If $A_{i_1 \ldots i_n}$ is an $\overbrace{n \times n \times \ldots \times n}^{n \text{ factors}}$ array which is totally antisymmetric, there is a constant α such that
$$A_{i_1 \ldots i_n} = \alpha \varepsilon_{i_1 \ldots i_n}.$$
6. $\varepsilon_{i_1 \ldots i_n} \varepsilon_{j_1 \ldots j_n} = \sum_{\pi \in S_n} (\mathrm{sgn}\pi)\, \delta_{i_1 j_{\pi(1)}} \ldots \delta_{i_1 j_{\pi(n)}}$ (The Polar Identity). The two special cases of interest to us are

   $n = 2$**:** $\varepsilon_{ij}\varepsilon_{kl} = \delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk}$ which implies $\varepsilon_{ij}\varepsilon_{kj} = \delta_{jk}$ and $\varepsilon_{ij}\varepsilon_{ij} = 2$;

   $n = 3$**:**
$$\begin{aligned} \varepsilon_{ijk}\varepsilon_{lmn} &= \delta_{il}\delta_{jm}\delta_{kn} + \delta_{im}\delta_{jn}\delta_{kl} + \delta_{in}\delta_{jl}\delta_{km} \\ &\quad -\delta_{im}\delta_{jl}\delta_{kn} - \delta_{il}\delta_{jn}\delta_{km} - \delta_{in}\delta_{jm}\delta_{kl} \end{aligned}$$

which implies $\varepsilon_{ijk}\varepsilon_{lmk} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}$ and $\varepsilon_{ijk}\varepsilon_{ljk} = 2\delta_{il}$ and $\varepsilon_{ijk}\varepsilon_{ijk} = 6$.

Some applications of $\varepsilon_{i_1 \ldots l_n}$ are these. If A_{ij} is a real or complex $n \times n$ array whose determinant is $\det A$, then

7.

$$A_{i_1 j_1} A_{i_2 j_2} \ldots A_{i_n j_n} \varepsilon_{j_1 \ldots j_n} = (\det A)\, \varepsilon_{i_1 \ldots i_n}.$$

In real three-space, suppose $\hat{x}_1, \hat{x}_2, \hat{x}_3$ is a right-handed triple of mutually perpendicular unit vectors (that is, $\hat{x}_1 \cdot (\hat{x}_2 \times \hat{x}_3) = 1$). Suppose $\vec{u} = u_i \hat{x}_i$ and $\vec{v} = v_j \hat{x}_j$. Then

$$\begin{aligned} \vec{u} \cdot \vec{v} &= u_i v_i \\ \vec{u} \times \vec{v} &= \hat{x}_i \varepsilon_{ijk} u_j v_k, \quad \text{or} \\ (\vec{u} \times \vec{v})_i &:= \hat{x}_i \cdot (\vec{u} \times \vec{v}) = \varepsilon_{ijk} u_j v_k. \end{aligned}$$

If $\vec{r} = r_i \hat{x}_i$ is the position vector in real three-space, $\mathcal{R}^3$, and $\vec{w} : \mathcal{R}^3 \to \mathcal{R}^3$ is a vector field, with

$$\begin{aligned} \vec{w}(\vec{r}\,) &= w_i(\vec{r}\,)\hat{x}_i, \quad \text{then} \\ \vec{\nabla} \cdot \vec{w} &= \text{divergence of } \vec{w} = \frac{\partial w_i}{\partial r_i} \\ \vec{\nabla} \times \vec{w} &= \text{curl of } \vec{w} = \hat{x}_i \varepsilon_{ijk} \frac{\partial w_k}{\partial r_j}. \end{aligned}$$

A very important general property of arrays is the following:

Remark 1 Suppose $A_{ijk_1 \ldots k_p}$ is antisymmetric in i and j, while $S_{ijl_1 \ldots l_q}$ is symmetric in i and j. Then

$$A_{ijk_1 \ldots k_P}\ \ S_{ijl_1 \ldots l_q} = 0.$$

Proof:

The subscripts k and l are irrelevant, so we omit them. We have $A_{ij}S_{ij} = -A_{ji}S_{ji}$ because $A_{ij} = -A_{ji}$ and $S_{ij} = S_{ji}$. Now replace the summation index

j by i and the index i by j. Then

$$A_{ji}S_{ji} = A_{ij}S_{ij}.$$

Putting these two results together gives $A_{ij}S_{ij} = -A_{ij}S_{ij}$, so $A_{ij}S_{ij} = 0$.

Vector Space Facts

Fields: A field is a set F of objects which can be added, subtracted, multiplied and divided according to the ordinary rules of arithmetic. Examples are $\mathcal{R}$, $\mathcal{C}$, the set of rational numbers, and the set of numbers $a + b\sqrt{2}$, where a and b are rational. We will use only the fields $\mathcal{R}$ and $\mathcal{C}$, and usually $F = \mathcal{R}$.

Definition of vector spaces: A vector space is an ordered quadruple $(V,\ F,\ ad,\ sc)$ with these properties:

i) V is a non-empty set. Its members are called vectors.

ii) F is a field. Its members are called scalars.

iii) $ad:\ V \times V \to V$. This function is called vector addition.

iv) $sc:\ F \times V \to V$. This function is called multiplication by scalars.

The remaining properties make more sense if we introduce a new notation. For any $\vec{u}, \vec{v} \in V$ and $c \in F$ we define

$$\begin{aligned} \vec{u} + \vec{v} : &= ad(\vec{u}, \vec{v}) \\ c\vec{v} &= \vec{v}c := sc(c, \vec{v}) \end{aligned}$$

The additional properties which make (V, F, ad, sc) a vector space are properties of ad and sc as follows (letters with $\to$ are vectors, those without $\to$ are scalars):

v) $\vec{u} + \vec{v} = \vec{v} + \vec{u}$

vi) $\vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$

vii) $(a+b)\vec{u} = a\vec{u} + b\vec{u}$

viii) $a(\vec{u}+\vec{v}) = a\vec{u} + a\vec{v}$

ix) $a(b\vec{u}) = (ab)\vec{u}$

x) $0\vec{u} = 0\vec{v}$ for any $\vec{u}, \vec{v} \in V$

xi) $1\vec{u} = \vec{u}$

The vector $0\vec{u}$ is called the zero vector, written $\vec{0}$. The vectors $(-1)\vec{v}$ and $\vec{u}+(-1)\vec{v}$ are written $-\vec{v}$ and $\vec{u}-\vec{v}$ respectively.

Example 1: F is a field. V is the set of all ordered n-tuples of members of F, so a typical vector is

$$\vec{u} = (u_1, \ldots u_n) \text{ with } u_1, \ldots, u_n \in F.$$

The addition and multiplication functions are

$$\begin{aligned} \vec{u}+\vec{v} &:= ad(\vec{u},\vec{v}) := (u_1+v_1, \ldots, u_n+v_n) \\ c\vec{v} &:= sc(c,\vec{v}) := (cv_1, \ldots, cv_n) \end{aligned}$$

The zero vector is $\vec{0} = (0, \ldots, 0)$. The space V is usually written F^n.

Example 2: $F = \mathcal{R}$. V is the set of all ordered n tuples of members of $\mathcal{R}^+$, so a typical vector is

$$\vec{u} = (u_1, \ldots, u_n) \text{ with } u_1, \ldots, u_n \in \mathcal{R}^+.$$

The addition and multiplication functions are

$$\begin{aligned} \vec{u}+\vec{v} &:= ad(\vec{u},\vec{v}) := (u_1v_1, \ldots, u_nv_n) \\ c\vec{v} &:= sc(c,\vec{v}) := (v_1^c, \ldots, v_n^c). \end{aligned}$$

The zero vector is $\vec{0} = (1, 1, \ldots, 1)$.

Example 3: Let (V, F, ad, sc) be any vector space and let S be any non-empty set. Define a new vector space (V_S, F, ad_S, sc_S) as follows:

$$V_S = \mathcal{F}(\mathcal{S} \to \mathcal{V}), \text{ the set of all functions} f: \ \mathcal{S} \to \mathcal{V}.$$

To define ad_S and sc_S we note that if f and g are functions in V_S and $c \in F$ then $ad_S(f,g)$ and $sc_S(c,f)$ are supposed to be functions in V_S, i.e., $ad_S(f,g) : S \to V$ and $sc_S(c,f) : S \to V$. To define these functions, we must say what values they assign to each $s \in S$. The definitions we adopt are these:

$$[ad_S(f,g)](s) = f(s) + g(s) \tag{D-2}$$

$$[sc_S(c,f)](s) = cf(s). \tag{D-3}$$

Following the convention for vector spaces, we write $ad_S(f,g)$ as $f+g$ and $sc_S(c,f)$ as cf. Thus, $f+g$ and cf are functions whose domains are S and whose ranges are subsets of V. The definitions (1) and (2) read thus: for any $s \in S$

$$(f+g)(s) = f(s) + g(s) \tag{D-4}$$

$$(cf)(s) = cf(s). \tag{D-5}$$

These are not "obvious facts." They are definitions of the vector-valued functions $f+g$ and cf.

In V_S, the zero vector is the function which assigns to each $s \in S$ the zero vector in V.

Example 4: F = real numbers = $\mathcal{R}$. V = set of all <u>continuous</u> real-valued functions on the closed unit interval $0 \leq x \leq 1$. ad and sc are defined as in Example 3, page D-12. (This is a vector space, because if f and g are continuous functions so is $f+g$ and so is cf for any real c.)

Example 5: $F = \mathcal{R}$. $V =$ set of all continuous <u>positive</u> real valued functions on $0 \leq x \leq 1$. ad and sc as defined in Examples 3 and 4. This is *not* a vector space. There are scalars $c \in F$ and vectors $f \in V$ such that $cf \not\in V$. For example, if $f \in V$ then $(-1)f \not\in V$. In other words, the function sc in this case is *not* a mapping from $F \times V$ to V.

Notation: Usually the vector space (V, F, ad, sc) will be called simply the vector space V. We will almost never use the notations $ad(\vec{u}, \vec{v})$ or $sc(c, \vec{v})$, but will use $\vec{u} + \vec{v}$

and $c\vec{v}$ and $\vec{v}c$. V will be called a vector space over F, or a real or complex vector space if $F = \mathcal{R}$ or $\mathcal{C}$.

Because of vector space rules v) and vi), there is no ambiguity about $\sum_{i=1}^{n} \vec{u}_i$ when $\vec{u}_1, \ldots, \vec{u}_n \in V$. For example, if $n = 4$, all of $\vec{u}_1 + [u_2 + (\vec{u}_3 + u_4)]$, $(\vec{u}_2 + \vec{u}_4) + (\vec{u}_3 + \vec{u}_1)$, $[\vec{u}_1 + \vec{u}_2) + \vec{u}_3] + \vec{u}_4, [\vec{u}_4 + (\vec{u}_2 + \vec{u}_1)] + \vec{u}_3$, etc., are the same.

When $\vec{u}_1, \ldots, \vec{u}_n \in V$ and $a_1, \ldots, a_n \in F$, we do not even need the $\sum$. We can use the index conventions to write $\sum_{i=1}^{n} a_i \vec{u}_i$ as $a_i \vec{u}_i$.

Facts about $\vec{0}$. If $a \in F$, then $a\vec{0} = \vec{0}$ because $a\vec{0} = a(0\vec{0}) = (a0)\vec{0} = 0\vec{0} = \vec{0}$. Also, if $a \in F$, $\vec{v} \in V$, and $a \neq 0$, $\vec{v} \neq \vec{0}$, then $a\vec{v} \neq \vec{0}$. For suppose $a\vec{v} = \vec{0}$ and $a \neq 0$. Then $a^{-1} \in F$, and $\vec{v} = 1\vec{v} = (a^{-1}a)\vec{v} = a^{-1}(a\vec{v}) = a^{-1}\vec{0} = \vec{0}$. Finally, $\vec{0} + \vec{u} = \vec{u}$ because $\vec{0} + \vec{u} = 0\vec{u} + 1\vec{u} = (0+1)\vec{u} = 1\vec{u} = \vec{u}$.

Linear mappings: Suppose V and W are both vector spaces over the same field F, and $L \in \mathcal{F}(V \to W)$. We call L a "linear mapping" if for any $\vec{u}, \vec{v} \in V$ and $c \in F$ it is true that

$$\begin{aligned} L(\vec{u} + \vec{v}) &= L(\vec{u}) + L(\vec{v}) \\ L(c\vec{u}) &= cL(\vec{u}). \end{aligned}$$

$\boxed{\mathcal{L}(V \to W) \text{ is the set of all linear mappings from } V \text{ to } W.}$[1]

It is a vector space over F if it's *ad* and *sc* are defined as in example 3, page D-12.

If $L \in \mathcal{L}(V \to W)$ is a bijection, then $L^{-1} \in \mathcal{L}(W \to V)$. And then L is called an "isomorphism" and V and W are said to be isomorphic. Intuitively speaking, two isomorphic vector spaces are really the same space. Each member of V corresponds to exactly one member of W, and vice versa, and *ad* and *sc* can be applied in either V or W with the same result. For example, if $F = \mathcal{R}$, then the two spaces in examples 1 and 2 on page D-12 are isomorphic. If V is the space in Example 2,

[1]Note that if $L \in \mathcal{L}(V \to W)$ and $M \in \mathcal{L}(U \to V)$, then $L \circ M \in \mathcal{L}(U \to W)$.

then the isomorphism $L: \mathcal{R}^n \to V$ is given by

$$L(u_1, \ldots, u_n) = (e^{u_1}, \ldots, e^{u_n})$$

and

$$L^{-1}(v_1, \ldots, v_n) = (ln\ v_1, \ldots ln\ v_n).$$

Subspaces: A subspace of vector space (V, F, ad, sc) is a vector space $(\tilde{V}, F, \tilde{ad}, \tilde{sc})$ with these properties:

i) $\tilde{V} \subseteq V$

ii) $\widetilde{ad} = ad\ |\tilde{V} \times \tilde{V}$

iii) $\widetilde{sc} = sc\ |F \times \tilde{V}$.

If $\tilde{V}$ is any subset of V, then $(\tilde{V}, F, ad \left|\tilde{V} \times \tilde{V}, sc\ \right| F \times \tilde{V})$ is a subspace of (V, F, ad, sc) if and only if

a) $ad(\tilde{V} \times \tilde{V}) \subseteq \tilde{V}$ and also

b) $sc(F \times \vec{V}) \subseteq \tilde{V}$.

Intuitively, subspaces are lines, planes and hyperplanes in V passing through the origin. (If $\tilde{V}$ is a subspace of V, and $\vec{v} \in \tilde{V}$, then $0\vec{v} \in \tilde{V}$ by iii), but $0\vec{v} = \vec{0}$. Thus the zero vector in V is the zero vector in every subspace of V.)

Finite dimensional vector spaces. V is a vector space over field F. U is an arbitrary non-empty subset of V.

Linear combination: If $\vec{u}_1, \ldots, \vec{u_n} \in V$ and $a_1, \ldots, a_n \in F$, the vector $a_i\vec{u}_i$ is called a linear combination of $\{\vec{u}_1, \ldots \vec{u}_n\}$.

Span: The set of all linear combinations of finite subsets of U is written $sp\, U$. It is called the span of U. It is a subspace of V, so it is called the subspace spanned or generated by U. If U is a subspace, $U = sp\, U$.

Finite dimensional. A vector space V is "finite dimensional" if $V = sp\ U$ for some *finite* subset $U \subset V$.

Linear dependence. A set $U \subseteq V$ is "linearly dependent" if there are finitely many vectors $\vec{u}_1, \ldots, \vec{u}_n \in U$ and scalars $a_1, \ldots, a_n$ such that

i) at least one of $a_1, \ldots, a_n$ is not 0, and also

ii) $a_i \vec{u}_i = \vec{0}$.

It is an important theorem that if $(\vec{u}_1, \ldots, \vec{u}_n)$ is a sequence such that $\{\vec{u}_1, \ldots, \vec{u}_n\}$ is linearly dependent, then there is an $m \in \{1, \ldots, n\}$ such that $\vec{u}_m \in sp\{\vec{u}_1, \ldots, \vec{u}_{m-1}\}$.

Linear independence. A set $U \subseteq V$ is linearly independent if it is not linearly dependent. Linear independence of U means that whenever $\{\vec{u}_1, \ldots \vec{u}_n\} \subseteq U$ and $a_i \vec{u}_i = \vec{0}$, then $a_1 = \ldots = a_n = 0$.

Basis. Any linearly independent subset of vector space V which spans V is called a basis for V. The following are important theorems about bases.

Theorem 2 If V is finite dimensional, it has a basis, and all its bases are finite and contain the same number of vectors. This number is called the dimension of V, written $\dim V$.

Theorem 3 If V is finite dimensional and W is a subspace of V, W is finite dimensional and $\dim W \leq \dim V$. Every basis for W is a subset of a basis for V.

Theorem 4 If $\{\vec{u}_1, \ldots, \vec{u}_n\}$ spans V then $\dim V \leq n$. If $\dim V = n$, then $\{\vec{u}_1, \ldots, \vec{u}_n\}$ is linearly independent, hence a basis for V.

Theorem 5 If V is finite dimensional and U is a linearly independent subset of V, U is finite, say $U = \{\vec{u}_1, \ldots, \vec{u}_n\}$, and $n \leq \dim V$. If $n = \dim V$, U spans V and so is a basis for V.

Theorem 6 Suppose $B = \{\vec{b}_1, \ldots \vec{b}_n\}$ is a basis for V. Then there are functions $c_B^1, \ldots, c_B^n \in \mathcal{L}(V \to F)$ such that for any $\vec{v} \in V$

$$\vec{v} = c_B^j(\vec{v})\vec{b}_j.$$

The scalars $v^1 = c_B^1(\vec{v}), \ldots, v^n = c_B^n(\vec{v})$, which are uniquely determined by $\vec{v}$ and B, are called the coordinates of $\vec{v}$ relative to the basis B. The linear functions $c_B^1, \ldots, c_B^n$ are called the coordinate functionals for the basis B. (A function whose values are scalars is called a "functional.") Clearly $\vec{b}_i = \delta_i{}^j \vec{b}_j$. Since the coordinates of any vector relative to B are unique, it follows that

$$c_B^j(\vec{b}_i) = \delta_i{}^j \tag{D-6}$$

Theorem 7 Suppose V and W are vector spaces over F, and $B = \{\vec{b}_1, \ldots, \vec{b}_n\}$ is a basis for V, and $\{\vec{w}_1, \ldots, \vec{w}_n\} \varepsilon W$. Then $\exists 1 L \in \mathcal{L}(V \to W) \ni L(\vec{b}_i) = \vec{w}_i$. (In other words, L is completely determined by $L|B$ for one basis B, and $L|B$ can be chosen arbitrarily.) (The L whose existence and uniqueness are asserted by the theorem is easy to find. For any $\vec{v} \in V$, $L(\vec{v}) = c_B^i(\vec{v})\vec{w}_i$.) This L is an injection if $\{\vec{w}_1, \ldots, \vec{w}_n\}$ is linearly independent and a surjection if $W = sp\{\vec{w}_1, \ldots, \vec{w}_n\}$.

Linear operators: If $L \in \mathcal{L}(V \to V)$, L is called a "linear operator on V." If $B = \{\vec{b}_1, \ldots, \vec{b}_n\}$ is any basis for V, then the array $L_i{}^j = c_B^j[L(\vec{b}_i)]$ is called the matrix of L relative to B. Clearly, $L(\vec{b}_i) = L_i{}^j \vec{b}_j$. (The matrix of L relative to B is often defined as the transpose of our $L_i{}^j$. It is that matrix $\ni L(\vec{b}_i) = \vec{b}_j L^j{}_i$. Our definition is the more convenient one for continuum mechanics.)

The determinant of the matrix of L relative to B depends only on L, not on B, so it is called the determinant of L, written $\det L$. A linear operator L is invertible iff $\det L \neq 0$. If it is invertible it is an isomorphism (a surjection as well as an injection).

If $L \in \mathcal{L}(V, V)$, $\vec{v} \in V$, $\lambda \in F$, and $\vec{v} \neq \vec{0}$, and if

$$L(\vec{v}) = \lambda \vec{v} \tag{D-7}$$

then λ is an "eigenvalue" of L, and $\vec{v}$ is an "eigenvector" of L belonging to eigenvalue λ. The eigenvalues of L are the solution $\lambda \in F$ of the polynomial equation

$$\det(L - \lambda I_V) = 0. \tag{D-8}$$

Linear operators on a finite dimensional complex vector space always have at least one eigenvalue and eigenvector. On a real vector space this need not be so.

If L and $M \in \mathcal{L}(V,V)$, then $L \circ M \in \mathcal{L}(V,V)$ and $\det(L \circ M) = (\det L)(\det M)$.

Euclidean Vector Spaces

Definition: A Euclidean vector space is an ordered quintuple $(V, \mathcal{R}, ad, sc, dp)$ with these properties:

i) $(V, \mathcal{R}, ad, sc)$ is a finite-dimensional (real) vector space

ii) dp is a functional on $V \times V$ which is symmetric, positive-definite and bilinear.

In more detail, the requirements on dp are these:

a) $dp: \ V \times V \to \mathcal{R}$. ($dp$ is a functional)

b) $dp(\vec{u}, \vec{v}) = dp(\vec{v}, \vec{u})$ (dp is symmetric)

c) $dp(\vec{u}, \vec{u}) > 0$ if $\vec{u} \neq \vec{0}$. (dp is positive definite)

d) for any $\vec{u}, \ \vec{v} \in V$, $dp(\cdot, \vec{v}\,)$ and $dp(\vec{u}, \cdot) \in \mathcal{L}(V \to R)$. ($dp$ is bilinear).

dp is called the dot-product functional, and $dp(\vec{u}, \vec{v})$ is usually written $\vec{u} \cdot \vec{v}$. In this notation, the requirements on the dot product are

a)$'$ $\vec{u} \cdot \vec{v} \in \mathcal{R}$ (dp is a functional)

b)$'$ $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$ (dp is symmetric)

c)$'$ $\vec{u} \cdot \vec{u} > 0$ if $\vec{u} \neq \vec{0}$ (dp is positive definite)

d)$'$ for any $\vec{u}, \vec{u}_1, \ldots, \vec{u}_n, \ \vec{v} \in V$ and $a_1, \ldots, a_n \in \mathcal{F}$

$$\left.\begin{aligned} (a_i \vec{u}_i) \cdot \vec{v} &= a_i (\vec{u}_i \cdot \vec{v}) \\ \vec{v} \cdot (a_i \vec{u}_i) &= a_i (\vec{v} \cdot \vec{u}_i) \end{aligned}\right\} dp \text{ is bilinear}$$

Example 1: $(V, \mathcal{R}, ad, sc)$ is as in Example 1, page D-12, and $\vec{u} \cdot \vec{v} = u_i v_i$. For $n = 3$ this is ordinary three-space.

Example 2: $(V, \mathcal{R}, ad, sc)$ is as in Example 4, page D-13. If f and g are two functions in V, their dot product is $f \cdot g = \int_0^1 dx f(x) g(x)$. Note that without continuity, $f \neq 0$ need not imply $f \cdot f > 0$. For example, let $f(x) = 0$ in $0 \leq x \leq 1$ except at $x = \frac{1}{2}$. There let $f(\frac{1}{2}) = 1$. Then $f \neq 0$ but $f \cdot f = 0$.

Length and angle. Schwarz and triangle inequalities. The length of a vector $\vec{v}$ in a Euclidean space is defined as $\|\vec{v}\,\| := \sqrt{\vec{v} \cdot \vec{v}}$. All nonzero-vectors have positive length. From the definitions on page D-18,

$$|\vec{u} \cdot \vec{v}| \leq \|\vec{u}\| \, \|\vec{v}\| \qquad \text{(Schwarz inequality)}$$

$$|\vec{u} + \vec{v}| \leq \|\vec{u}\| + \|\vec{u}\| \qquad \text{(triangle inequality)}$$

If $\vec{u} \neq \vec{0}$ and $\vec{v} \neq \vec{0}$ then $\vec{u} \cdot \vec{v} / \|\vec{u}\,\| \|\vec{v}\,\|$ is a real number between -1 and 1, so it is the cosine of exactly one angle between 0 and π. This angle is called the angle between $\vec{u}$ and $\vec{v}$, and is written $\angle(\vec{u}, \vec{v})$. Thus, by the definition of $\angle(\vec{u}, \vec{v})$,

$$\vec{u} \cdot \vec{v} = \|\vec{u}\| \|\vec{v}\| \cos \angle(\vec{u}, \vec{v}).$$

In particular, $\vec{u} \cdot \vec{v} = 0 \Leftrightarrow \angle(\vec{u}, \vec{v}) = \frac{\pi}{2}$. If $\|\vec{u}\| = 1$, we will write $\hat{u}$ for $\vec{u}$.

Orthogonality: If V is a Euclidean vector space, and $\vec{u}, \vec{v} \in V$, $P \subseteq V$, $Q \subseteq V$, then

$$\begin{array}{lll} \vec{u} \perp \vec{v} & \text{means} & \vec{u} \cdot \vec{v} = 0. \\ \vec{u} \perp Q & \text{means} & \vec{u} \cdot \vec{q} = 0 \text{ for all } \vec{q} \in Q. \\ P \perp Q & \text{means} & \vec{p} \cdot \vec{q} = 0 \text{ for all } \vec{p} \in P \text{ and } \vec{q} \in Q. \end{array}$$

$Q^\perp := \{\vec{x} : \vec{x} \in V \text{ and } \vec{x} \perp Q\}$. This set is called "the orthogonal complement of Q." It is a subspace of V. Obviously $Q \perp Q^\perp$. Slightly less obvious, but true in finite dimensional spaces, is $(Q^\perp)^\perp = sp\, Q$.

Theorem 8 Q spans $V \Leftrightarrow Q^\perp = \left\{\vec{0}\right\}$. In particular, if $\left\{\vec{b}_1, \ldots, \vec{b}_n\right\}$ is a basis for V and $\vec{v} \cdot \vec{b}_i = 0$, then $\vec{v} = \vec{0}$. Also, if $\vec{v} \cdot \vec{x} = 0$ for all $\vec{x} \in V$, then $\vec{v} = \vec{0}$.

Orthonormal sets and bases. <u>An orthonormal set Q</u> in V is any set of mutually perpendicular unit vectors. I.e., $\vec{u} \in Q \Rightarrow \|\vec{u}\| = 1$; and $\vec{u}, \vec{v} \in Q$, $\vec{u} \neq \vec{v} \Rightarrow \vec{u} \cdot \vec{v} = 0$.

An orthonormal set is linearly independent, so in a finite dimensional Euclidean space, any orthonormal set is finite, and can have no more than $\dim V$ members. An orthonormal basis in V is any basis which is an orthonormal set. If $B = \{\hat{x}_1, \ldots, \hat{x}_n\}$ is an orthonormal basis for V, the coordinate functionals c_B^i are given simply by

$$c_B^i(\vec{v}\,) = \vec{v} \cdot \hat{x}_i, \quad \forall \vec{v} \in V. \tag{D-9}$$

That is, for any $\vec{v}$, $\vec{v} = (\vec{v} \cdot \hat{x}_i)\hat{x}_i$.

Theorem 9 Every Euclidean vector space has an orthonormal basis. If U is a subspace of Euclidean vector space V, U is itself a Euclidean vector space, and every orthonormal basis for U is a subset of an orthonormal basis for V.

Linear functionals and dual bases: Let V be a Euclidean vector space. For any $\vec{v} \in V$ we can define a $\lambda_{\vec{v}} \in \mathcal{L}(V \to \mathcal{R})$ by requiring simply

$$\lambda_{\vec{v}}(\vec{y}) = \vec{v} \cdot \vec{y}, \quad \forall \vec{y} \in V. \tag{D-10}$$

Theorem 10 Let V be a Euclidean vector space. The mapping $\vec{v} \mapsto \lambda_{\vec{v}}$ defined by equation (9) is an isomorphism between V and $\mathcal{L}(V \to \mathcal{R})$. (Therefore we can think of the linear functionals on V, the members of $\mathcal{L}(V \to \mathcal{R})$, simply as vectors in V.)

Proof:

We must show that the mapping $\vec{v} \mapsto \lambda_{\vec{v}}$ is an injection, a surjection, and linear. Linearity is easy to prove. We want to show $\lambda_{a_i\vec{v}_i} = a_i\lambda_{\vec{v}_i}$. But for any $\vec{y} \in V$,

$$\begin{aligned} \lambda_{a_i\vec{v}_i}(\vec{y}) &= (a_i\vec{v}_i) \cdot \vec{y} = a_i(\vec{v}_i \cdot \vec{y}) = a_i\left[\lambda_{\vec{v}_i}(\vec{y}\,)\right] \\ &= (a_i\lambda_{\vec{v}_i})(\vec{y}\,). \end{aligned}$$

(See definition of $a_i\lambda_{\vec{v}_i}$ in example 3, page D-12) Therefore the two functions $\lambda_{a_i\vec{v}_i}$ and $a_i\lambda_{\vec{v}_i}$ are equal.

To prove that $\vec{v} \mapsto \lambda_{\vec{v}}$ is injective, suppose $\lambda_{\vec{u}} = \lambda_{\vec{v}}$. We want to prove $\vec{u} = \vec{v}$. For any $\vec{y} \in V$, we know $\lambda_{\vec{u}}(\vec{y}) = \lambda_{\vec{v}}(\vec{y})$, so $\vec{u} \cdot \vec{y} = \vec{v} \cdot \vec{y}$, or $(\vec{u} - \vec{v}) \cdot \vec{y} = 0$. In particular, setting $\vec{y} = \vec{u} - \vec{v}$ we see $(\vec{u} - \vec{v}) \cdot (\vec{u} - \vec{v}) = 0$, so $(\vec{u} - \vec{v}) = \vec{0}$, $\vec{u} = \vec{v}$.

To prove that $\vec{v} \mapsto \lambda_{\vec{v}}$ is surjective, let $\lambda \in \mathcal{L}(V \to \mathcal{R})$. We want to find a $\vec{v} \in V \ni \lambda = \lambda_{\vec{v}}$. Let $\{\hat{x}_1, \ldots, \hat{x}_n\}$ be an orthonormal basis for V and define $\vec{v} = \lambda(\hat{x}_i)\hat{x}_i$. Then for any $\vec{y} \in V$, $\lambda_{\vec{v}}(\vec{y}) = \vec{v} \cdot \vec{y} = \lambda(\hat{x}_i)(\hat{x}_i \cdot \vec{y}) = \lambda(\hat{x}_i) y_i = \lambda(y_i \hat{x}_i) = \lambda(\vec{y})$. Thus, $\lambda_{\vec{v}} = \lambda$. QED.

Even in a Euclidean space, it is sometimes convenient to work with bases which are not orthonormal. The advantage of an orthonormal basis $\hat{x}_1, \ldots, \hat{x}_n$ is that for any $\vec{v} \in V$, $\vec{v} = (\vec{v} \cdot \hat{x}_i)\hat{x}_i$. An equation almost as convenient as this can be obtained for an arbitrary basis. We need some definitions. All vectors are in a Euclidean vector space V.

Definition: Two sequences of vectors in V, $(\vec{u}_1, \ldots, \vec{u}_m)$ and $(\vec{v}_1, \ldots, \vec{v}_m)$, are dual to each other if $\vec{u}_i \cdot \vec{v}_j = \delta_{ij}$.

Remark 2 If $(\vec{u}_1, \ldots, \vec{u}_m)$ and $(\vec{v}_1, \ldots, \vec{v}_m)$ are dual to each other, each is linearly independent.

Proof:

If $a_i \vec{u}_i = \vec{0}$, then $(a_i \vec{u}_i) \cdot \vec{v}_j = 0$, so $a_i(\vec{u}_i \cdot \vec{v}_j) = 0$, so $a_i \delta_{ij} = 0$, so $a_j = 0$.

Definition: An ordered basis for V is a sequence of vectors $(\vec{b}_1, \ldots, \vec{b}_n)$ such that $\{\vec{b}_1, \ldots, \vec{b}_n\}$ is a basis for V.

Remark 3 If two sequences of vectors are dual to each other and one is an ordered basis, so is the other. (Proof. If $(\vec{v}_1, \ldots, \vec{v}_m)$ is a basis, $m = \dim V$. Then, since $(\vec{u}_1, \ldots, \vec{u}_m)$ is linearly independent, it is also a basis.

Remark 4 If $(\vec{b}_1, \ldots, \vec{b}_n)$ is an ordered basis for V, it has *at most* one dual sequence $(\vec{b}^1, \ldots, \vec{b}^n)$.

Proof:

If $(\vec{b}^1, \ldots, \vec{b}^n)$ and $(\vec{v}_1, \ldots, \vec{v}_n)$ are both dual sequences to $(\vec{b}_1, \ldots, \vec{b}_n)$, then $\vec{b}^i \cdot \vec{b}_j = \delta_{ij}$, and $\vec{v}_i \cdot b_j = \delta_{ij}$, so $(\vec{b}^i - \vec{v}_i) \cdot \vec{b}_j = 0$. Then $(\vec{b}^i - \vec{v}_i) \cdot (a_j \vec{b}_j) = 0$ for any $a_i, \ldots, a_n \in \mathcal{R}$. But any $\vec{u} \in V$ can be written $\vec{u} = a_j \vec{b}_j$, so $(\vec{b}^i - \vec{v}_i) \cdot \vec{u} = 0$ for all $\vec{u} \in V$. Hence, $\vec{b}^i - \vec{v}_i = 0$, $\vec{b}^i = \vec{v}_i$.

Remark 5 If $(\vec{b}_1, \ldots, \vec{b}_n)$ is an ordered basis for V, it has *at least* one dual sequence $(\vec{b}^1, \ldots, \vec{b}^n)$.

Proof:

Let $B = \{\vec{b}_1, \ldots \vec{b}_n\}$, and let c^i_B be the coordinate functions for this basis. They are linear functionals on V, so according to theorem (10) there is exactly one $\vec{b}^i \in V$ such that $\lambda_{\vec{b}^i} = c^i_B$. Then $\lambda_{\vec{b}^i}(\vec{y}) = c^i_B(\vec{y}\,)$, $\forall \vec{y} \in V$, so $\vec{b}^i \cdot \vec{y} = c^i_B(\vec{y}\,)$, $\forall \vec{y} \in V$. In particular, $\vec{b}^i \cdot \vec{b}_j = c^i_B(\vec{b}_j) = \delta^i{}_j$ [see (5)]. Thus $(\vec{b}^1, \ldots, \vec{b}^n)$ is dual to $(\vec{b}^1, \ldots, \vec{b}_n)$.

From the foregoing, it is clear that each ordered basis $B = (\vec{b}_i, \ldots, \vec{b}_n)$ for V has exactly one dual sequence $B^D = (\vec{b}^1, \ldots, \vec{b}^n)$, and that this dual sequence is also an ordered basis for V. It is called the dual basis for $(\vec{b}_1, \ldots, \vec{b}_n)$. It is characterized and uniquely determined by

$$\vec{b}^i \cdot \vec{b}_j = \delta^i{}_j. \tag{D-11}$$

There is an obvious symmetry: each of B and B^D is the dual basis for the other.

If $\vec{v} \in V$, and $B = (\vec{b}_1, \ldots, \vec{b}_n)$ is any ordered basis, we can write

$$\vec{v} = v^j \vec{b}_j, \tag{D-12}$$

where $v^j = c^j_B(\vec{v})$ are the coordinates of $\vec{v}$ relative to B. Then $\vec{b}^i \cdot \vec{v} = v^j \vec{b}^i \cdot \vec{b}_j = v^j \delta^i_j = v^i$. Thus

$$v^i = \vec{b}^i \cdot \vec{v}. \tag{D-13}$$

Since $(\vec{b}^1, \ldots, \vec{b}^n) = B^D$ is also an ordered basis for V, we can write

$$\vec{v} = v_j \vec{b}^j. \tag{D-14}$$

Because duality is symmetrical, it follows immediately that

$$v_j = \vec{v} \cdot \vec{b}_j. \tag{D-15}$$

An orthonormal basis is its own dual basis; it is self-dual. If $\vec{b}_i = \hat{x}_i$ then $\vec{b}^i = \hat{x}_i$, and (11), (12) are the same as (13) and (14). We have

$$v_j = v^j = \hat{x}_j \cdot \vec{v}.$$

In order to keep indices always up or down, it is sometimes convenient to write an orthonormal basis $(\hat{x}_1, \ldots, \hat{x}_n)$ as $(\hat{x}^1, \ldots, \hat{x}^n)$, with $\hat{x}_i = \hat{x}^i$. Then

$$v_j = v^j = \hat{x}_j \cdot \vec{v} = \hat{x}^j \cdot \vec{v}.$$

If $(\vec{b}_1, \ldots, \vec{b}_n) = B$ is an ordered basis for V, and $B^D = (\vec{b}^1, \ldots, \vec{b}^n)$ is its dual basis, and $\vec{v} \in V$, we can write

$$\vec{v} = v^i \vec{b}_i = v_i \vec{b}^i. \tag{D-16}$$

The v^i are called the contravariant components of $\vec{v}$ relative to B, and the v_i are the covariant components of $\vec{v}$ relative to B (and also the contravariant components of $\vec{v}$ relative to B^D). The contravariant components of $\vec{v}$ relative to B are the coordinates of $\vec{v}$ relative to B. The covariant components of $\vec{v}$ relative to B are the coordinates of $\vec{v}$ relative to B^D.

The covariant metric matrix of B^D is

$$g_{ij} = \vec{b}_i \cdot \vec{b}_j. \tag{D-17}$$

The contravariant metric matrix of B^D is

$$g^{ij} = \vec{b}^i \cdot \vec{b}^j. \tag{D-18}$$

Clearly g_{ij} is the contravariant and g^{ij} the covariant metric matrix of B^D.

We have $v_i = \vec{b}_i \cdot \vec{v} = \vec{b}_i \cdot (\vec{b}_j v^j) = (\vec{b}_i \cdot \vec{b}_j) v^j$ so

$$v_i = g_{ij} v^j. \tag{D-19}$$

Similarly, or by the symmetry of duality,

$$v^i = g^{ij} v_j. \tag{D-20}$$

The metric matrices can be used to raise or lower the indices on the components of $\vec{v}$.

Transposes: Suppose V and W are Euclidean vector spaces and $L \in \mathcal{L}(V \to W)$. Then there is exactly one $L^T \in \mathcal{F}(W \to V)$ such that

$$L(\vec{v}) \cdot \vec{w} = \vec{v} \cdot L^T(\vec{w}), \quad \forall (\vec{v}, \vec{w}) \in V \times W. \tag{D-21}$$

This L^T is linear, i.e. $L^T \in \mathcal{L}(W \to V)$. Furthermore, $(L^T)^T = L$, and if $a_1, \ldots, a_n \in \mathcal{R}$ and $L_1, \ldots, L_n \in \mathcal{L}(V \to W)$, then

$$(a_i L_i)^T = a_i (L_i^T) \tag{D-22}$$

The linear mapping L^T is called the "transpose" of L (also sometimes the adjoint of L, but adjoint has a second, quite different, meaning for matrices, so we avoid the term).

If $L \in \mathcal{L}(V \to W)$ and $M \in \mathcal{L}(U \to V)$, then $L \circ M \in \mathcal{L}(U \to W)$ and it is easy to verify that

$$(L \circ M)^T = M^T \circ L^T. \tag{D-23}$$

If $L \in \mathcal{L}(V \to W)$ is bijective, then $L^{-1} \in \mathcal{L}(W \to V)$. It is easy to verify that $L^T : W \to V$ is also bijective and

$$(L^{-1})^T = (L^T)^{-1}. \tag{D-24}$$

Orthogonal operators or isometries: If V is a vector space, the members of $\mathcal{L}(V \to V)$ are called "linear operators on V." If V is Euclidean and $L \in \mathcal{L}(V \to V)$, then L is called "orthogonal" or an "isometry" if $\|L(\vec{u}) - L(\vec{v})\| = \|\vec{u} - \vec{v}\|$ for all $\vec{u}$ and $\vec{v}$ in V. If $L \in \mathcal{L}(V \to V)$, then each of the following conditions is necessary and sufficient for L to be an isometry:

a) $\|L\vec{v}\| = \|\vec{v}\|, \quad \forall\, \vec{v} \in V$

b) $L(\vec{u}) \cdot L(\vec{v}) = \vec{u} \cdot \vec{v}, \quad \forall \vec{u}, \vec{v} \in V$

c) $L^T \circ L = I_V$

d) $L \circ L^T = I_V$

e) $L^T = L^{-1}$

f) For every orthonormal basis $B \subseteq V$, $L(B)$ is orthonormal. That is, if $B = \{\hat{x}_1, \ldots, \hat{x}_n\}$ is any orthonormal basis, then $\{L(\hat{x}_1), \ldots, L(\hat{x})\}$ is also an orthonormal basis.

g) There is at least one orthonormal basis $B \ni L(B)$ is orthonormal.

h) For every orthonormal basis $B \subseteq V$, the matrix of L relative to B has orthonormal rows and columns (i.e., $L^i{}_j L^i{}_k = \delta_{jk}$, and $L^i{}_j L^k{}_j = \delta^{ik}$).

i) There is at least one orthonormal basis $B \subseteq V$ relative to which the matrix of L has either orthonormal rows $(L^i{}_j L^k{}_j = \delta^{ik})$ or orthonormal columns $(L^i{}_j L^i{}_k = \delta_{ij})$.

If L is orthogonal, det $L = \pm 1$. If $\det L = +1$, L is called "proper." Otherwise L is "improper."

$\Theta(V) :=$ set of all orthonormal operators on V.

$\Theta^+(V) :=$ set of all proper orthonormal operators on V.

If L and $M \in \Theta(V)$ or $\Theta^+(V)$, the same is true of $L \circ M$ and L^{-1}. Also $I_V \in \Theta(V)$ and $\Theta^+(V)$. Therefore, $\Theta(V)$ and $\Theta^+(V)$ are both groups if multiplication is defined to mean composition. $\Theta(V)$ is called the orthogonal group on V, while $\Theta^+(V)$ is the proper orthogonal group.

If $L \in \Theta^+(V)$, L is a "rotation." This implies that there is an orthonormal basis $\hat{x}_1, \ldots, \hat{x}_{2m}, \hat{x}_{2m+1}$ and angles $\Theta_1, \ldots, \Theta_m$ between 0 and π such that

$$\left.\begin{aligned} L(\hat{x}_{2i-1}) &= \hat{x}_{2i-1}\cos\theta_i + \hat{x}_{2i}\sin\theta_i \\ L(\hat{x}_{2i}) &= -\hat{x}_{2i-1}\sin\theta_i + \hat{x}_{2i}\cos\theta_i \\ L(\hat{x}_{2n+1}) &= \hat{x}_{2n+1}. \end{aligned}\right\} \forall i \in \{1, \ldots, m\}$$

The vector $\hat{x}_{2n+1}$ is omitted in the above definition if $\dim V$ is even, and included if $\dim V$ is odd. If included, it is the "axis" of the rotation.

If $L \in \Theta(V)$ is improper, L is the product of a rotation and a "reflection." A reflection is an L for which there is an orthonormal basis $\hat{x}_1, \ldots, \hat{x}_n$ such that $L(\hat{x}_1) = -\hat{x}_1$, $L(\hat{x}_n) = \hat{x}_k$, if $k \geq 2$.

Symmetric operators: A symmetric operator on Euclidean space V is an $L \in \mathcal{L}(V, V) \ni L^T = L$. (Antisymmetric means $L^T = -L$). Each of the following conditions is necessary and sufficient for a linear operator L to be symmetric.

a) $L^T = L$

b) $L(\vec{u}) \cdot \vec{v} = L(\vec{v}) \cdot \vec{u}, \quad \forall \vec{u},\ \vec{v} \in V$

c) The matrix of L relative to every orthonormal basis for V is symmetric

d) The matrix of L relative to at least one orthonormal basis for V is symmetric.

e) There is an orthonormal basis for V consisting entirely of eigenvectors of L. That is, there is an orthonormal ordered basis $(\hat{x}_1, \ldots, \hat{x}_n)$ for V, and a sequence of real numbers $(\lambda_1, \ldots, \lambda_n)$ (not necessarily different) such that

$$L(\hat{x}_{(i)}) = \lambda_{(i)}\hat{x}_{(i)}, \quad \forall i \in \{1, \ldots, n\}. \tag{D-25}$$

Facts: The eigenvalues $\{\lambda_1, \ldots, \lambda_n\}$ are uniquely determined by L, but the eigenvectors are not. For any real λ let $[\![L; \lambda]\!]$ denote the set of all $\vec{v} \in V$ such that $L(\vec{v}) = \lambda\vec{v}$.

Then $[\![L;\lambda]\!]$ is a subspace of V, and $\dim[\![L;\lambda]\!] > 0 \Leftrightarrow \lambda$ is an eigenvalue of L. In that case, $[\![L;\lambda]\!]$ is called the eigenspace of L with eigenvalue λ. The eigenspaces are uniquely determined by L. If $\lambda_1 \neq \lambda_2$, $[\![L;\lambda_1]\!] \perp [\![L;\lambda_2]\!]$.

Positive-definite (semidefinite) symmetric operators. If $L \in \mathcal{L}(V \to V)$ is symmetric, L is said to be positive definite when

$$L(\vec{v}) \cdot \vec{v} > 0 \text{ for all nonzero } \vec{v} \in V \tag{D-26}$$

and positive semidefinite when

$$L(\vec{v}) \cdot \vec{v} \geq 0 \text{ for all } \vec{v} \in V. \tag{D-27}$$

A symmetric linear operator is positive definite (semidefinite) iff all its eigenvalues are positive (non-negative).

Every symmetric positive (semi) definite operator L has a unique positive (semi) definite symmetric square root, i.e., a symmetric positive (semi) definite operator $L^{1/2}$ such that

$$L^{1/2} \circ L^{1/2} = L \tag{D-28}$$

Polar Decomposition Theorem: (See Halmos, p. 169 ff). Suppose V is Euclidean and $L \in \mathcal{L}(V \to V)$. Then there exist orthogonal operators $\mathcal{O}_1$ and $\mathcal{O}_2$ and positive semidefinite symmetric operators S_1 and S_2 such that

$$L = \mathcal{O}_1 \circ S_1 = S_2 \cdot \mathcal{O}_2. \tag{D-29}$$

Both S_1 and S_2 are uniquely determined by L (in fact $S_1 = (L^T L)^{1/2}$ and $S_2 = (LL^T)^{1/2}$). If L^{-1} exists, $\mathcal{O}_1$ and $\mathcal{O}_2$ are uniquely determined by L and are equal, and S_1 and S_2 are positive definite. If $\det L > 0$ then $\mathcal{O}_1$ is proper.

The existence part of the polar decomposition theorem can be restated as follows: for any $L \in \mathcal{L}(V \to V)$ there are orthonormal bases $(\hat{x}_1, \ldots, \hat{x}_n)$ and $(\hat{y}_1, \ldots, \hat{y}_n)$ in V and non-negative numbers $\lambda_1, \ldots, \lambda_n$ such that

$$L(\hat{x}_{(i)}) = \lambda_{(i)}\hat{y}_{(i)}, \quad i = 1, \ldots, n. \tag{D-30}$$

If L is invertible, all $\lambda_{(i)} > 0$.

Stated this way, the theorem remains true for any $L(V \to W)$ if V and W are Euclidean spaces and $\dim V \le \dim W$. If $\dim V > \dim W$, then (29) holds for $i = 1, \ldots, m$, with $m = \dim W$, while for $i = m+1, \ldots, n$ (with $n = \dim V$) (29) is replaced by

$$L(\hat{x}_i) = \vec{0}. \tag{D-31}$$

Five Prerequisite Proofs

The Polar Identity

Let $S_n :=$ group all permutations on $\{1, \ldots, n\}$.

$$\begin{aligned} \varepsilon_{i_1,\ldots,i_n} &:= n \text{ dimensional alternating symbol} \\ &:= 0 \text{ if } \{i_1, \ldots, i_n\} \neq \{1, \ldots, n\} \\ &:= \operatorname{sgn}\pi \text{ if } i_1 = \pi(1), \ldots, i_n = \pi(n), \pi \in S_n. \end{aligned}$$

Then the polar identity is

$$\sum_{\pi \in S_n} (\operatorname{sgn}\pi)\delta_{i_1 j_{\pi(1)}} \cdots \delta_{i_n j_{\pi(n)}} = \varepsilon_{i_1\ldots i_n}\varepsilon_{j_1\ldots j_n}. \tag{PP-1}$$

Lemma PP-1 Suppose $A_{i_1\ldots i_n}$ is an $n \times \ldots \times n$ dimensional n'th order array which is totally antisymmetric (i.e. changes sign where any two indices are interchanged). Then

$$A_{i_1\ldots i_n} = (A_{12\ldots n})\, \varepsilon_{i_1\ldots i_n}. \tag{PP-2}$$

Proof:

If any two of $\{i_1, \ldots, i_n\}$ are equal, $A_{(i_1,\ldots i_n)} = 0$. If they are all different, there is a $\pi \in S_n$ such that $i_1 = \pi(1), \ldots, i_n = \pi(n)$. This π is a product of transpositions (interchanges), and by carrying out these interchanges in succession on $A_{12\ldots n}$ one obtains $A_{\pi(1)\ldots\pi(n)}$. Each interchange produces a sign change, so the final sign is $+$ or $-$ according as π is even or odd. Thus, $A_{\pi(1)\ldots\pi(n)} = (\operatorname{sgn}\pi)A_{1\ldots n} = \varepsilon_{\pi(1)\ldots\pi(n)}A_{1\ldots n}$. Therefore, equation PP-2, holds both when $\{i_1, \ldots, i_n\}$ contains a repeated index and when it does not.

Lemma PP-2 Suppose W is a vector space and $f : S_n \to W$ and $\sigma \in S_n$. Then

$$\sum_{\pi \in S_n} f(\pi) = \sum_{\pi \in S_n} f(\pi^{-1}) = \sum_{\pi \in S_n} f(\sigma\pi). \tag{PP-3}$$

Proof:

Both the mapping $\pi \mapsto \pi^{-1}$ and the mapping $\pi \mapsto \sigma\pi$ are bijections of S_n to itself. Therefore, each of the sums in (PP-3) contains exactly the same terms. These terms are vectors in W, so they can be added in any order.

Lemma PP-3 Define

$$A_{i_1 \dots i_n j_1 \dots j_n} = \sum_{\pi \in S_n} (\operatorname{sgn} \pi) \delta_{i_1 j_{\pi(1)}} \dots \delta_{i_n j_{\pi(n)}} \tag{PP-4}$$

Then

$$A_{i_1 \dots i_n j_1 \dots j_n} = A_{j_1 \dots j_n i_1 \dots i_n} \tag{PP-5}$$

and for any $\sigma \in S_n$,

$$A_{i_1 \dots i_n j_{\sigma(1)} \dots j_{\sigma(n)}} = (\operatorname{sgn} \sigma) A_{i_1 \dots i_n j_1 \dots j_n} \tag{PP-6}$$

Proof:

To prove (PP-5) note that by lemma 2, since $\operatorname{sgn} \pi^{-1} = \operatorname{sgn} \pi$,

$$A_{i_1 \dots i_n j_1 \dots j_n} = \sum_{\pi \in S_n} (\operatorname{sgn} \pi) \delta_{i_1 j_{\pi^{-1}(1)}} \dots \delta_{i_n j_{\pi^{-1}}}(n).$$

The set of n pairs

$$\left\{ \begin{pmatrix} 1 \\ \pi^{-1}(1) \end{pmatrix}, \dots, \begin{pmatrix} n \\ \pi^{-1}(n) \end{pmatrix} \right\}$$

is the same as the set

$$\left\{ \begin{pmatrix} \pi(1) \\ 1 \end{pmatrix}, \dots, \begin{pmatrix} \pi(n) \\ n \end{pmatrix} \right\}$$

because a pair

$$\begin{pmatrix} i \\ j \end{pmatrix}$$

is in the first set of pairs iff $j = \pi^{-1}(i)$ and in the second set of pairs iff $i = \pi(j)$. Therefore

$$\delta_{i_1 j_{\pi^{-1}(1)}} \ldots \delta_{i_n j_{\pi^{-1}(n)}} = \delta_{i_{\pi(1)} j_1} \ldots \delta_{i_{\pi(n)} j_n}$$

and

$$\begin{aligned} A_{i_1 \ldots i_n j_1 \ldots j_n} &= \sum_{\pi \in S_n} (\operatorname{sgn} \pi) \delta_{i_{\pi(1)} j_1} \ldots \delta_{i_{\pi(n)} j_n} \\ &= \sum_{\pi \in S_n} (\operatorname{sgn} \pi) \delta_{j_1 i_{\pi(1)}} \ldots \delta_{j_n i_{\pi(n)}} \\ &= A_{j_1 \ldots j_n i_1 \ldots i_n}. \end{aligned}$$

This proves equation (PP-5).

To prove equation (PP-6), let $\sigma \in S_n$ and let $j_1 = k_{\sigma(1)}, \ldots, j_n = k_{\sigma(n)}$. Then from equation (PP-4),

$$A_{i_1 \ldots i_n k_{\sigma(1)} \ldots k_{\sigma(n)}} = \sum_{\pi \in S_n} (\operatorname{sgn} \pi) \delta_{i_1 k_{\sigma\pi(1)}} \ldots \delta_{i_n k_{\sigma\pi(n)}}.$$

Now $(\operatorname{sgn} \sigma)^2 = 1$ and so $(\operatorname{sgn} \pi) = (\operatorname{sgn} \sigma)^2 (\operatorname{sgn} \pi) = (\operatorname{sgn} \sigma)\,[(\operatorname{sgn} \sigma)(\operatorname{sgn} \pi)] = (\operatorname{sgn} \sigma)(\operatorname{sgn} \sigma\pi)$, and

$$\begin{aligned} A_{i_1 \ldots i_n k_{\sigma(1)} \ldots k_{\sigma(n)}} &= (\operatorname{sgn} \sigma) \sum_{\pi \in S_n} (\operatorname{sgn} \sigma\pi) \delta_{i_1 k_{\sigma\pi(1)}} \ldots \delta_{i_n k_{\sigma\pi(n)}} \\ &= (\operatorname{sgn} \sigma) \sum_{\pi \in S_n} (\operatorname{sgn} \pi) \delta_{i_1 k_{\pi(1)}} \ldots \delta_{i_n k_{\pi(n)}} \qquad \text{by Lemma PP-2} \\ &= (\operatorname{sgn} \sigma) A_{i_1 \ldots i_n k_1 \ldots k_n}. \end{aligned}$$

This proves equation (PP-6).

Now we can prove the polar identity. Let $A_{i_1 \ldots i_n j_1 \ldots j_n}$ be as in (PP-4). Then fix $i_1, \ldots, i_n$. By lemma 1 and equation (PP-6),

$$A_{i_1 \ldots i_n j_1 \ldots j_n} = A_{i_1 \ldots i_n 12 \ldots n} \varepsilon_{j_1 \ldots j_n}.$$

By equation (PP-5), interchanging the i's and j's gives

$$A_{i_1 \ldots i_n j_1 \ldots j_n} = A_{12 \ldots n\, j_1 \ldots j_n} \varepsilon_{i_1 \ldots i_n}. \tag{PP-7}$$

Set $i_1 = 1, \ldots, i_n = n$, and equate these two. The result is

$$A_{12\ldots n\, j_1\ldots j_n}\varepsilon_{12\ldots n} = A_{12\ldots n12\ldots n}\varepsilon_{j_1\ldots j_n}$$

so

$$A_{12\ldots n\, j_1\ldots j_n} = A_{12\ldots n\; 12\ldots n}\varepsilon_{j_1\ldots j_n}. \tag{PP-8}$$

Substituting (PP-8) in (PP-7) gives

$$A_{i_1\ldots i_n j_1\ldots j_n} = A_{12\ldots n\; 12\ldots n}\varepsilon_{i_1\ldots i_n}\varepsilon_{j_1\ldots j_n}. \tag{PP-9}$$

Then from the definition (PP-4), with $i_r = r$, $j_s = s$

$$A_{12\ldots n12\ldots n} = \sum_{\pi\in S_n} (\operatorname{sgn}\pi)\delta_{1\pi(1)}\ldots\delta_{n\pi(n)}.$$

There are $n!$ terms in this sum, and a term vanishes unless $\pi(1) = 1$, $\pi(2) = 2, \ldots \pi(n) = n$; that is, unless π is the identity permutation e. But $\operatorname{sgn} e = 1$, so only one of the $n!$ terms in the sum is nonzero, and this term is 1. Therefore, $A_{12\ldots n12\ldots n} = 1$, and equation (PP-9) becomes the polar identity (PP-1).

Spectral Decomposition of a Symmetric Linear Operator S on a Euclidean Space V.

Definition PP-1 $S : V \to V$ is symmetric iff $S = S^T$, that is iff $\vec{u} \cdot S(\vec{v}) = S(\vec{u}) \cdot \vec{v}$ for all $\vec{u}$, $\vec{v} \in V$.

Theorem PP-11 Let S be a symmetric linear operator on Euclidean space V. Then V has an orthonormal basis consisting of eigenvectors of S.

Lemma PP-4 Suppose S is symmetric, U is a subspace of V, and $S(U) \subseteq U$. Then $S(U^\perp) \subseteq U^\perp$ and $S|U^\perp$ is symmetric as a linear operator on $U^\perp$.

Proof:

> Suppose $\vec{w} \in U^\perp$. We want to show $S(\vec{w}) \in U^\perp$. That is, for any $\vec{u} \in U$, we want to show $\vec{u} \cdot S(\vec{w}) = 0$. But $\vec{u} \cdot S(\vec{w}) = S(\vec{u}) \cdot \vec{w}$. By hypothesis, $S(\vec{u}) \in U$ and $\vec{w} \in U^\perp$, so $S(\vec{u}) \cdot \vec{w} = 0$. Thus, $S(\vec{w}) \in U^\perp$. Therefore $S(U^\perp) \subseteq U^\perp$. And if the definition (PP-1) above is satisfied for all $\vec{u}$, $\vec{v} \in V$, it certainly holds for all $\vec{u}$, $\vec{v} \in U^\perp$. Hence, $S|U^\perp$ is symmetric.

Lemma PP-5 Suppose $S : V \to V$ is a symmetric linear operator on Euclidean space V. Then S has at least one eigenvalue σ such that $S(\vec{v}) = \sigma\vec{v}$, for some $\vec{v} \neq \vec{0}$.

Proof:

Let $\hat{x}_1, \ldots, \hat{x}_n$ be an orthonormal basis for V, and let S_{ij} be the matrix of S relative to this basis, so $S(\hat{x}_i) = S_{ij}\hat{x}_j$. Then $S_{ij} = S(\hat{x}_i) \cdot \hat{x}_j = \hat{x}_i \cdot S(\hat{x}_j) = S(\hat{x}_j) \cdot \hat{x}_i = S_{ji}$, so S_{ij} is a symmetric matrix. A real number σ is an eigenvalue of S iff $\det(S - \sigma I) = 0$, where I is the identity operator on V. (If the determinant is 0, then $S - \sigma I$ is singular, so there is a nonzero $\vec{v} \ni (S - \sigma I)(\vec{v}) = \vec{0}$, or $S(\vec{v}) - \sigma\vec{v} = \vec{0}$, or $S(\vec{v}) = \sigma\vec{v}$.) The matrix of $S - \sigma I$ relative to the basis $(\hat{x}_1, \ldots, \hat{x}_n)$ is $S_{ij} - \sigma\delta_{ij}$, so we want to prove that the n'th degree polynomial in σ, $\det(S_{ij} - \sigma\delta_{ij})$, has at least one real zero. It certainly has a complex zero, σ, so there is a complex n-tuple $(r_1, \ldots, r_n)$ such that

$$(S_{ij} - \sigma\delta_{ij})r_j = 0, \text{ or } S_{ij}r_j = \sigma r_i.$$

Let $*$ denote complex conjugation. Then

$$r_i^* S_{ij} r_j = \sigma(r_i r_i^*). \tag{PP-10}$$

Taking the complex conjugate of this equation gives (S_{ij} is real)

$$r_i S_{ij} r_j^* = \sigma^*(r_i^* r_i).$$

Changing the summation indices on the left gives

$$r_j S_{ji} r_i^* = \sigma^*(r_i^* r_i). \tag{PP-11}$$

But $S_{ij} = S_{ji}$, so the left hand sides of (PP-10) and (PP-11) are the same. Also, $r_i^* r_i = r_i r_i^* \neq 0$. Hence, $\sigma^* = \sigma$. Thus σ is real, and we have proved lemma (PP-5).

Now we can prove theorem PP-1. Let σ be any (real) eigenvalue of S. Let $\hat{v}$ be a unit eigenvector with this eigenvalue. Then $S(sp\{\hat{v}\}) \subseteq sp\{\vec{v}\}$ so by lemma PP-4, $S(\{\hat{v}\}^\perp) \subseteq \{\hat{v}\}^\perp$ and $S|\{\hat{v}\}^\perp$ is symmetric. We proceed by induction on $\dim V$. Since $\dim\{\hat{v}\}^\perp < \dim V$, we may assume theorem PP-4 true on $\{\hat{v}\}^\perp$. Adjoining $\hat{v}$ to the orthonormal basis for $\{\hat{v}\}^\perp$ which consists of eigenvectors of $S|\{\hat{v}\}^\perp$ gives an orthonormal basis for V consisting of eigenvectors of S.

Definition PP-2 Let S be a linear operator on Euclidean space V. A "spectral decomposition of S" is a collection of real numbers $\sigma_1 < \ldots < \sigma_m$ and a collection of subspaces of V, $U_1, \ldots, U_m$, such that

i) $S|U_{(i)} = \sigma_{(i)}\ I|U_{(i)}$ for $i = 1, \ldots, m$

ii) $V = U_1 \oplus \ldots \oplus U_m$. (This means

a) $U_i \perp U_j$ if $i \neq j$. That is, if $\vec{u}_i \in U_i$ and $\vec{u}_j \in U_j$ then $\vec{u}_i \cdot \vec{u}_j = 0$;

b) every $\vec{v} \in V$ can be written in exactly one way as $\vec{v} = \vec{u}_1 + \cdots + \vec{u}_m$ with $\vec{u}_i \in U_i$.)

Corollary PP-1 If S is a symmetric linear operator on Euclidean space V, S has *at least* one spectral decomposition.

Proof:

Order the *different* eigenvalues of S which theorem PP-1 produces as $\sigma_1 < \cdots < \sigma_m$. Let U_i be the span of all the orthonormal basis vectors turned up by that theorem which have eigenvalue σ_i.

Corollary PP-2 S has *at most* one spectral decomposition.

Proof:

Let $\sigma_1', < \cdots < \sigma_u'$ and $U_1', \ldots U_u'$ be a second spectral decomposition. Let σ' be one of $\sigma_1', \ldots, \sigma_u'$, and let $\vec{v}\,'$ be a corresponding nonzero eigenvector. Then $S(\vec{v}\,') = \sigma'\vec{v}\,'$. But $\vec{v}\,' = \vec{u}_1 + \cdots + \vec{u}_m$ where $\vec{u}_i \in U_i$, and $S(\vec{v})' = \sigma'\vec{v}\,' = \sigma'\vec{u}_1 + \ldots + \sigma'\vec{u}_m$. Hence $S(\vec{v}\,') = S(\vec{u}_1) + \cdots + S(\vec{u}_m) = \sigma_1\vec{u}_1 + \cdots + \sigma_m\vec{u}_m$. Hence,

$$(\sigma' - \sigma_1)\vec{u}_1 + \cdots + (\sigma' - \sigma_m)\vec{u}_m = \vec{0}. \tag{PP-12}$$

Since $\vec{v}\,' \neq \vec{0}$, there is at least one $\vec{u}_j \neq \vec{0}$. Then, dotting this $\vec{u}_j$ into (PP-12) gives $(\sigma' - \sigma_{(j)})\vec{u}_{(j)} \cdot \vec{u}_{(j)} = 0$, so $\sigma' - \sigma_j = 0$, or $\sigma' = \sigma_j$. Thus every σ_i' is included in $\{\sigma_1, \ldots, \sigma_m\}$. By symmetry, every σ_i is included in $\{\sigma'_1, \ldots, \sigma'_u\}$. Therefore the two sets of eigenvalues are equal, and $u = m$. Since the eigenvalues are ordered by size, $\sigma'_i = \sigma_i$ for $i = 1, \ldots, m$. Now suppose $\vec{u}\,' \in U'_i$. Then $\vec{u}_i' = \vec{u}_1 + \cdots + \vec{u}_m$ and $S(\vec{u}_i') = \sigma_{(i)}\vec{u}_{(i)}' = \sigma_i(\vec{u}_1 + \cdots + \vec{u}_m) = S(\vec{u}_1) + \cdots + S(\vec{u}_m) = \sigma_{(1)}\vec{u}_1 + \cdots + \sigma_m\vec{u}_m$ so$(\sigma_i - \sigma_1)\vec{u}_1 + \cdots + (\sigma_i - \sigma_m)\vec{u}_m = \vec{0}$. Dot in $\vec{u}_j$ and we get $(\sigma_i - \sigma_j)\vec{u}_{(j)} \cdot \vec{u}_{(j)} = 0$. Since $\sigma_i \neq \sigma_j$, $\vec{u}_{(j)} \cdot \vec{u}_{(j)} = 0$, and $\vec{u}_j = \vec{0}$. Thus $\vec{u}_i' = \vec{u}_i$. Thus, $\vec{u}_i' \in U_i$. Thus $U'_i \subseteq U_i$. By symmetry, $U_i \subseteq U'_i$. Thus $U_i = U'_i$. QED.

Corollary PP-3 If a linear operator on a Euclidean space has a spectral decomposition, it is symmetric.

Proof:

Suppose S has spectral decomposition $(\sigma_1, U_1), \cdots, (\sigma_m, U_m)$. Let $\vec{u}, \vec{v} \in V$. We can write $\vec{u} = \sum_{i=1}^m \vec{u}_i$, $\vec{v} = \sum_{i=1}^m \vec{v}_i$ with $\vec{u}_i$, $\vec{v}_i \in U_i$. Then

$$\begin{aligned}
S(\vec{u}) &= \sum_{i=1}^m S(\vec{u}_i) = \sum_{i=1}^m \sigma_i \vec{u}_i \\
S(\vec{v}) &= \sum_{i=1}^m S(\vec{v}_i) = \sum_{i=1}^m \sigma_i \vec{v}_i \\
\vec{u} \cdot S(\vec{v}) &= \left(\sum_{i=1}^m \vec{u}_i\right) \cdot \left(\sum_{j=1}^m \sigma_j \vec{v}_j\right) = \sum_{i,j=1}^m \sigma_j(\vec{u}_i \cdot \vec{v}_j) \\
&= \sum_{i=1}^m \sigma_i(\vec{u}_i \cdot \vec{v}_i) \\
S(\vec{u}) \cdot \vec{v} &= \left(\sum_{i=1}^m \sigma_i \vec{u}_i\right) \cdot \left(\sum_{j=1}^m \vec{v}_j\right) = \sum_{i,j=1}^m \sigma_i(\vec{u}_i \cdot \vec{v}_j) \\
&= \sum_{i=1}^m \sigma_i(\vec{u}_i \cdot \vec{v}_i).
\end{aligned}$$

Hence $\vec{u} \cdot S(\vec{v}) = S(\vec{u}) \cdot \vec{v}$.

Corollary PP-4 Two linear operators with the same spectral decomposition are equal.

Proof:

Suppose S and S' have the same spectral decomposition $(\sigma_1, U_1), \ldots, (\sigma_m, U_m)$. Let $\vec{v} \in V$. Then $\vec{v} = \sum_{i=1}^m \vec{u}_i$ with $\vec{u}_i \in U_i$. Then

$$\begin{aligned} S(\vec{v}) &= \sum_{i=1}^m S(\vec{u}_i) = \sum_{i=1}^m \sigma_i \vec{u}_i = \sum_{i=1}^m S'(\vec{u}_i) \\ &= S'\left(\sum_{i=1}^m \vec{u}_i\right) = S'(\vec{v}). \end{aligned}$$

Since this is true for all $\vec{v} \in V$, $S = S'$.

Positive Definite Operators and Their Square Roots

Definition PP-3 A symmetric linear operator S is positive definite if $\vec{v} \cdot S(\vec{v}) > 0$ for all $\vec{v} \neq \vec{0}$. (S is positive semi-definite if $\vec{v} \cdot S(\vec{v}) \geq 0$ for all $\vec{v} \in V$.)

Corollary PP-5 A symmetric linear operator S is positive (semi) definite iff all its eigenvalues are (non-negative) positive.

Proof:

- **$\Rightarrow$.** Suppose S has an eigenvalue $\sigma \leq 0$. Let $\vec{v}$ be a nonzero vector with $S(\vec{v}) = \sigma \vec{v}$. Then $\vec{v} \cdot S(\vec{v}) = \sigma(\vec{v} \cdot \vec{v}) \leq 0$, so S is not positive definite.
- **$\Leftarrow$.** Let $(\sigma_1, U_1), \ldots, (\sigma_m, U_m)$ be the spectral decomposition of S. Let $(\vec{v} \in V$, $\vec{v} \neq \vec{0}$. Then $\vec{v} = \sum_{i=1}^m \vec{u}_i$ with $\vec{u}_i \in U_i$ and at least one $\vec{u}_i \neq \vec{0}$. Then $S(\vec{v}) = \sum_{i=1}^m S(\vec{u}_i) = \sum_{i=1}^m \sigma_i \vec{u}_i$, so

$$\begin{aligned} \vec{v} \cdot S(\vec{v}) &= \left(\sum_{i=1}^m \vec{u}_i \right) \cdot \left(\sum_{j=1}^m \sigma_j u_j \right) = \sum_{i,j=1}^m \sigma_j \vec{u}_i \cdot \vec{u}_j \\ &= \sum_{i=1}^m \sigma_i \left(\vec{u}_i \cdot \vec{u}_i \right) > 0. \end{aligned}$$

Note: The proofs require an obvious modification if S is positive semi-definite rather than positive definite.

Theorem PP-12 Suppose S is a symmetric, positive (semi) definite linear operator in Euclidean space V. Then there is exactly one symmetric positive (semi) definite linear operator $Q : V \to V$ such that $Q^2 = S$. (Q is called the positive square root of S, and is written $S^{1/2}$.)

Proof:

Existence. Let $(\sigma_1, U_1), \ldots, (\sigma_m, U_m)$ be the spectral decomposition of S. Since $\sigma_i > 0$, let $\sigma_i^{1/2}$ be the positive square root of σ_i. Let Q be the linear operator on V whose spectral decomposition is $(\sigma_1^{1/2}, U_1), \ldots, (\sigma_m^{1/2}, U_m)$. Then Q is symmetric by corollary PP-5, and positive definite by corollary PP-6. Let $\vec{v} \in V$. Then $\vec{v} = \sum_{i=1}^m \vec{u}_i$ with $\vec{u} \in U_i$. Then

$$\begin{aligned} Q(\vec{v}) &= \sum_{i=1}^m Q(\vec{u}_i) = \sum_{i=1}^m \sigma_i^{1/2} \vec{u}_i \\ Q^2(\vec{v}) &= \sum_{i=1}^m \sigma_i^{1/2} Q(\vec{u}_i) = \sum_{i=1}^m \sigma_i^{1/2} \sigma_i^{1/2} \vec{u}_i \\ &= \sum_{i=1}^m \sigma_i \vec{u}_i = \sum_{i=1}^m S(\vec{u}_i) = S(\vec{v}). \end{aligned}$$

Hence, $Q^2 = S$.

Uniqueness. Suppose Q is symmetric, positive definite, and $Q^2 = S$. Let $(q_1, \tilde{U}_1), \ldots, (q_u, \tilde{U}_u)$ be the spectral decomposition of Q. Then if $\vec{u}_i \in \tilde{U}_i$ we have $S(\vec{u}_i) = Q^2 \vec{u}_i = QQ(\vec{u}_i) = Q(q_{(i)} \vec{u}_{(i)}) = q_{(i)}^2 \vec{u}_{(i)}$. Thus q_i^2 is an eigenvalue of S. Therefore $(q_1^2, \tilde{U}_1), \ldots, (q_u^2, \tilde{U}_u)$ is a spectral decomposition for S. Therefore $u = m$, $\tilde{U} = U_i$, $q_i^2 = \sigma_i$. Since $q_i > 0$, $q_i = \sigma_i^{1/2}$. Therefore Q has the spectral decomposition $(\sigma_1^{1/2}, U_1), \ldots, (\sigma_m^{1/2}, U_m)$, and is the operator found in the existence part of the proof.

The Polar Decomposition Theorem

Theorem PP-13 Suppose V is a Euclidean vector space, $L : V \to V$ is a linear operator, and $\det L > 0$. Then there are unique linear mappings R_1, S_1, R_2, S_2 with these properties:

i) $L = R_1 S_1$

ii) $L = S_2 R_2$

iii) R_1 and R_2 are proper orthogonal (i.e. $R^{-1} = R^T$ and $\det R = +1$)

iv) S_1 and S_2 are positive definite and symmetric.

Further, $R_1 = R_2$, so $S_1 = R_2^T S_2 R_2$.

Before giving the proof we need some lemmas.

Lemma PP-6 If $\det \neq 0$, $L^T L$ and LL^T are symmetric and positive definite.

Proof:

It suffices to consider $L^T L$, because $LL^T = (L^T)^T (L^T)$ and $\det L^T = \det L$. First, $(L^T L)^T = L^T (L^T)^T = L^T L$, so $L^T L$ is symmetric. Second, if $\vec{v} \neq \vec{0}$, $L(\vec{v}) \neq \vec{0}$ because L^{-1} exists. But then $L(\vec{v}) \cdot L(\vec{v}) > 0$. And $L(\vec{v}) \cdot L(\vec{v}) = \vec{v} \cdot L^T[L(\vec{v})] = \vec{v} \cdot [(L^T L)(\vec{v})]$. Thus, $\vec{v} \cdot [(L^T L)(\vec{v})] > 0$, so $L^T L$ is positive definite.

Lemma PP-7 If S is positive definite and symmetric, $\det S > 0$.

Proof:

Let $\hat{x}_1, \ldots \hat{x}_n$ be an orthonormal basis for V consisting of eigenvectors for S, namely $s_1, \ldots, s_n$. (Some of these may be the same.) Relative to basis $\hat{x}_1, \ldots, \hat{x}_n$, the matrix of S is diagonal, and its diagonal entries are $s_1, \ldots, s_n$. Therefore $\det S = s_1 s_2 \ldots s_n$. But since S is positive definite, all its eigenvalues are positive. Therefore $\det S > 0$.

Lemma PP-8 If A and B are linear operators on V and $AB = I$ then B^{-1} exists and $B^{-1} = A$. Hence, $BA = I$ and A^{-1} exists and $A^{-1} = B$.

Proof:

If $AB = I$ then $(\det A)(\det B) = \det AB = \det I = 1$, so $\det B \neq 0$. Hence, B^{-1} exists. By its definition, $B^{-1}B = BB^{-1} = I$. And if we multiply $AB = I$ on the right by B^{-1} we have $(AB)B^{-1} = IB^{-1}$, or $A(BB^{-1}) = B^{-1}$, or $AI = B^{-1}$, or $A = B^{-1}$. QED.

Lemma PP-9 If L is a linear operator on Euclidean space and L^{-1} exists so does $(L^T)^{-1}$, and $(L^T)^{-1} = (L^{-1})^T$.

Proof:

Take transposes in the equation $LL^{-1} = I$. The result is $(L^{-1})^T L^T = I$. By lemma PP-8, $(L^T)^{-1}$ exists and equals $(L^{-1})^T$.

Proof of Theorem PP-3:

First we deal with R_1 and S_1. To prove these *unique*, we assume they exist and have the asserted properties. Then $L = R_1 S_1$ implies $L^T = (R_1 S_1)^T = S_1^T R_1^T = S_1 R_1^{-1}$ so $L^T L = S_1 R_1^{-1} R_1 S_1 = S_1 I S_1 = S_1^2$. Now $L^T L$ is a symmetric positive definite operator, and so is S_1. Therefore, S_1 must be the unique symmetric positive definite square root of $L^T L$, i.e.

$$S_1 = (L^T L)^{1/2}. \tag{PP-13}$$

Since $\det S_1 > 0$, S_1^{-1} exists. Then $L = R_1 S_1$ implies

$$R_1 = L S_1^{-1}. \tag{PP-14}$$

Thus S_1 and R_1 are given explicitly and uniquely in terms of L. To prove the *existence* of an R_1 and S_1 with the required properties, we *define* them by (PP-8) and (PP-9). Then S_1 is symmetric positive definite as required, and we need show only that R_1 is proper orthogonal, i.e. that $\det R_1 = +1$ and $R_1^T = R_1^{-1}$. But $\det R_1 = \det L \det(S_1^{-1}) = \det L / \det S_1 > 0$ since $\det L > 0$ and $\det S_1 > 0$. Therefore, if we can show $R_1^T = R_1^{-1}$, we automatically have $\det R_1 = 1$. By lemma PP-8 it suffices to prove $R_1^T R_1 = I$. But $R_1^T = (LS_1^{-1})^T = (S_1^{-1})^T L^T = (S_1^T)^{-1} L^T = S_1^{-1} L^T$ so $R_1^T R_1 = S_1^{-1} L^T L S_1^{-1} = S_1^{-1} S_1^2 S_1^{-1} = I$.

Next we deal with R_2 and S_2. To prove them *unique*, we suppose they exist. Then $L = S_2 R_2$ so $L^T = (S_2 R_2)^T = R_2^T S_2^T = R_2^{-1} S_2$. But if $\det L > 0$ then $\det L^T = \det L > 0$, so there is only one way to write $L^T = \tilde{R}_1 \tilde{S}_1$ with $\tilde{R}_1$ proper orthogonal and $\tilde{S}_1$ symmetric. If R_2 is proper orthogonal, so is R_2^{-1}, and we must have $R_2^{-1} = \tilde{R}_1$ so $R_2 = \tilde{R}_1^{-1}$. Then R_2 and S_2 are uniquely determined by L. In fact, $S_2 = \tilde{S}_1 = (L^T)^T L^T = L L^T$, and $R_2 = (\tilde{R}_1)^{-1} = (L^T S_2^{-1})^{-1} = S_2 (L^T)^{-1}$. To prove *existence* of R_2 and S_2 with the required properties, we observe that $L^T = \tilde{R}_1 \tilde{S}_1$ with $\tilde{R}_1$ proper orthogonal and $\tilde{S}_1$ symmetric. Then $(L^T)^T = L = (\tilde{R}_1 \tilde{S}_1)^T = \tilde{S}_1{}^T \tilde{R}_1^T = \tilde{S}_1 R_1^{-1}$. Therefore we can take $S_2 = \tilde{S}_1$, $R_2 = \tilde{R}_1^{-1}$.

Finally, to prove $R_1 = R_2$, note that if $L = S_2 R_2$ then $L = R_2 R_2^{-1} S_2 R_2 = R_2 (R_2^T S_2 R_2)$. We claim that $R_2^T S_2 R_2$ is symmetric and positive definite. If we can prove that, then $L = R_2 (R_2^T S_2 R_2)$ is of the form $L = R_1 S_1$, and uniqueness of R_1 and S_1 requires $R_1 = R_2$, $S_1 = R_2^T S_2 R_2$. Symmetry is easy. We have $(R_2^T S_2 R_2)^T = R_2^T S_2^T (R_2^T)^T = R_2^T S_2 R_2$. For positive definiteness let $\vec{v} \in V$, $\vec{v} \neq \vec{0}$. Then $R_2(\vec{v}) \neq 0$ because R_2^{-1} exists. Then $R_2(\vec{v}) \cdot S_2[R_2(\vec{v})] > 0$ because S_2 is positive definite. Then $\vec{v} \cdot R_2^T S_2 [R_2(\vec{v})] > 0$, or $\vec{v} \cdot [(R_2^T S_2 R_2)(\vec{v})] > 0$. Thus $R_2^T S_2 R_2$ is positive definite.

We shall not need the PDT when $\det L < 0$ or L^{-1} fails to exist. For these cases, see Halmos, p. 170.

Representation Theorem for Orthogonal Operators

Definition PP-4 A linear operator R on Euclidean vector space V is "orthogonal" iff it preserves length; that is, $\|R(\vec{v})\| = \|\vec{v}\|$ for every $\vec{v} \in V$.

There are several other properties each of which completely characterizes orthogonal operators (i.e., R is orthogonal iff it is linear and has one of these properties). We list them in lemmas.

Lemma PP-10 R is orthogonal iff R is linear and preserves all inner products; that is, $R(\vec{u}) \cdot R(\vec{v}) = \vec{u} \cdot \vec{v}$ for all $\vec{u}, \vec{v} \in V$.

Proof:

Obviously if R preserves all inner products it preserves lengths. The interesting half is that if R preserves lengths then it preserves all inner products. To see this, we note that for any $\vec{u}$ and $\vec{v} \in V$, $\|\vec{u}+\vec{v}\|^2 = (\vec{u}+\vec{v}) \cdot (\vec{u}+\vec{v}) = \vec{u} \cdot \vec{u} + \vec{u} \cdot \vec{v} + \vec{v} \cdot \vec{u} + \vec{v} \cdot \vec{v}$. That is

$$\|\vec{u}+\vec{v}\|^2 = \|\vec{u}\|^2 + 2\vec{u} \cdot \vec{v} + \|\vec{v}\|^2. \tag{PP-15}$$

Applying this result to $R(\vec{u})$ and $R(\vec{v})$, and using $R(\vec{u}+\vec{v}) = R(\vec{u}) + R(\vec{v})$, gives

$$\|R(\vec{u}+\vec{v})\|^2 = \|R(\vec{u})\|^2 + 2R(\vec{u}) \cdot R(\vec{v}) + \|R(\vec{v})\|^2. \tag{PP-16}$$

If R preserves *all* lengths, then $\|R(\vec{u}+\vec{v})\| = \|\vec{u}+\vec{v}\|$, $\|R(\vec{u})\| = \|\vec{u}\|$, $\|R(\vec{v})\| = \|\vec{v}\|$, so subtracting (PP-15) from (PP-16) gives $R(\vec{u}) \cdot R(\vec{v}) = \vec{u} \cdot \vec{v}$.

Lemma PP-11 R is orthogonal iff $R^T = R^{-1}$. (Therefore if R is orthogonal R^{-1} exists.)

Proof:

If $R^T = R^{-1}$ then $R^T R = I$. Then for any $\vec{u}, \vec{v} \in V$, $\vec{u} \cdot R^T R(\vec{v}) = \vec{u} \cdot \vec{v}$ so $R(\vec{u}) \cdot R(\vec{v}) = \vec{u} \cdot \vec{v}$. Hence R is orthogonal. Conversely, suppose R orthogonal. Then for any $\vec{u}, \vec{v} \in V$, $R(\vec{u}) \cdot R(\vec{v}) = \vec{u} \cdot \vec{v}$, so $\vec{u} \cdot R^T R(\vec{v}) = \vec{u} \cdot \vec{v}$, so $\vec{u} \cdot [(R^T R)(\vec{v}) - \vec{v}] = 0$. Fix $\vec{v}$. Then we see that $R^T R(\vec{v}) - \vec{v}$ is orthogonal to every $\vec{u} \in V$, hence to itself. Hence it is $\vec{0}$, so $R^T R(\vec{v}) = \vec{v}$. Since this is true for every $\vec{v} \in V$, $R^T R = I$. Then by lemma PP-8, $R^T = R^{-1}$.

Corollary PP-6 If either $R^T R = I$ or $R R^T = I$, R is orthogonal.

Proof:

By lemma PP-8, either equation implies $R^T = R^{-1}$.

Corollary PP-7 If R is orthogonal, so is R^{-1}.

Proof:

By lemma PP-11, R^{-1} exists and is R^T. But $R^T (R^T)^T = R^T R = I$, so R^T satisfies the second equation in corollary PP-6. Hence R^T is orthogonal.

The easiest way to discover whether a linear $R : V \to V$ is orthogonal is to study its matrix R_{ij} relative to some orthonormal basis for V. Recall that we define this matrix by

$$R(\hat{x})_i = R_{ij} \hat{x}_j \tag{PP-17}$$

where $(\hat{x}_1, \ldots, \hat{x}_n)$ is the orthonormal basis for V which we are using. The matrix L_{ij} of an arbitrary linear operator L is defined in the same way,

$$L(\hat{x})_i = L_{ij} \hat{x}_j. \tag{PP-18}$$

The matrix of L^T relative to $(\hat{x}_1, \ldots, \hat{x}_n)$ is the transposed matrix L_{ji}. That is

$$L^T(\hat{x}_i) = L_{ij}\hat{x}_j \quad \text{or} \quad (L^T)_{ij} = L_{ji}. \tag{PP-19}$$

To see this, we note that $\hat{x}_k \cdot L(\hat{x}_i) = L_{ij}\hat{x}_k \cdot \hat{x}_j = L_{ij}\delta_{kj} = L_{ik}$, so

$$L_{ij} = L(\hat{x}_i) \cdot \hat{x}_j. \tag{PP-20}$$

Therefore $(L^T)_{ij} = L^T(\hat{x}_i) \cdot \hat{x}_j = \hat{x}_i \cdot L(\hat{x}_j) = L_{ji}$.

Now we have

Remark 6 If R is orthogonal, then its matrix relative to every orthonormal basis satisfies

$$R_{ij}R_{ik} = \delta_{jk} \qquad \text{("orthonormal columns")} \tag{PP-21}$$

and

$$R_{ji}R_{ki} = \delta_{jk} \qquad \text{("orthonormal rows")} \tag{PP-22}$$

Proof:

If R is orthogonal so is R^T. Therefore, it suffices to prove (PP-21). We have $R^T(\hat{x}_j) = (R^T)_{ji}\hat{x}_i = R_{ij}\hat{x}_i$. Then $RR^T(\hat{x}_j) = R_{ij}R(\hat{x}_i) = R_{ij}R_{ik}\hat{x}_k$. But if R is orthogonal, $RR^T = I$, so $RR^T(\hat{x}_j) = \delta_{jk}\hat{x}_k$. Thus $R_{ij}R_{ik}\hat{x}_k = \delta_{jk}\hat{x}_k$, or $(R_{ij}R_{ik} - \delta_{jk})\hat{x}_k = 0$. Fix j. Since $\hat{x}_1, \ldots, \hat{x}_n$ are linearly independent, we must have $R_{ij}R_{ik} - \delta_{jk} = 0$. This is (PP-21).

Remark 7 If $R \in L(V \to V)$ and there is one orthonormal basis for V relative to which the matrix of R satisfies either (PP-21) or (PP-22), then R is orthogonal.

Proof:

Suppose that relative to the particular basis $\hat{x}_1, \ldots, \hat{x}_n$, the matrix of R satisfies (PP-21). Multiply by $\hat{x}_k$ and add over k, using $R_{ik}\hat{x}_k = R(\hat{x}_i)$. The result is

$$R_{ij}R(\hat{x}_i) = \hat{x}_j.$$

By linearity of R, this means $R(R_{ij}\hat{x}_i) = \hat{x}_j$ or $R\left[R^T(\hat{x}_j)\right] = \hat{x}_j$, or $(RR^T)(\hat{x}_j) = \hat{x}_j$. But then for any scalars $v_1, \ldots, \vec{v}_n$ we have $(RR^T)(v_j\hat{x}_j) = v_j RR^T(\hat{x}_j) = v_j\hat{x}_j$. Hence $RR^T(\vec{v}) = \vec{v}$ for all $\vec{v} \in V$. Hence $RR^T = I$. Hence R is orthogonal.

Next suppose R satisfies (PP-22) rather than (PP-21). Then $(R^T)_{ij}(R^T)_{ik} = \delta_{jk}$, so the matrix of R^T satisfies (PP-21). Hence, $R^T = R^{-1}$ is orthogonal. Hence so is $(R^T)^T = (R^{-1})^{-1} = R$.

Corollary PP-8 If the rows of a square matrix are orthonormal, so are the columns. A matrix with this property is called orthogonal. What we have seen is that if R is orthogonal, so is its matrix relative to every orthonormal basis. And if the matrix of R relative to one orthonormal basis is orthogonal, R is orthogonal, so its matrix relative to every orthonormal basis is orthogonal.

Now we discuss some properties of orthogonal operators which will eventually tell us what those operators look like.

First, we introduce the notation in

Definition PP-5 The set of all orthogonal operators on Euclidean space V is written $\Theta(V)$. The set of all proper orthogonal operators (those with determinant $+\,1$) is written $\Theta^+(V)$.

Lemma PP-12 $\Theta(V)$ and $\Theta^+(V)$ are groups if multiplication is defined as operator composition.

Proof:

All we need to prove is that both sets are closed under the operations of taking inverses and of multiplying. But if $R \in \Theta(V)$, we know $R^{-1} \in \Theta(V)$ by remark PP-2. And $\det(R^{-1}) = (\det R)^{-1} = 1$ if $\det R = 1$ so if $R \in \Theta^+(V)$ then $R^{-1} \in \Theta^+(V)$. If R_1 and $R_2 \in \Theta(V)$, then $(R_1R_2)^T(R_1R_2) = R_2^T R_1^T R_1 R_2 = R_2^T I R_2 = R_2^T R_2 = I$, so $R_1R_2 \in \Theta(V)$. And if R_1, $R_2 \in \Theta^+(V)$ then $\det(R_1R_2) = (\det R_1)(\det R_2) = 1$, so $R_1R_2 \in \Theta^+(V)$.

Lemma PP-13 If $R \in \Theta(V)$, $\det R = +1$ or -1.

Proof:

$R^T R = I$ so $(\det R^T)(\det R) = \det I$. But $\det R^T = \det R$ and $\det I = 1$. Thus $(\det R)^2 = 1$.

Lemma PP-14 Suppose $R \in \Theta(V)$ and U is a subspace of V and $R(U) \subseteq U$. Then $R|U \in \Theta(U)$.

Proof:

By hypothesis, $R|U$ is a linear operator on U. Since $\|R(\vec{v})\| = \|\vec{v}\|$ for all $\vec{v} \in V$, this is certainly true for all $\vec{v} \in U$. Hence, $R|U$ satisfies definition PP-4

Lemma PP-15 Under the hypothesis of lemma PP-14, $R^T(U) \subseteq U$.

Proof:

Suppose $\vec{u} \in U$. Then there is a unique $\vec{u}\,' \in U$ such that $\vec{u} = (R|U)(\vec{u}\,')$ (because $(R|U)^{-1}$ exists as a linear operator on U). But $\vec{u}\,' \in U$, so $(R|U)(\vec{u}\,') = R(\vec{u}\,')$. Thus $\vec{u} = R(\vec{u}\,')$, and $R^{-1}(\vec{u}) = \vec{u}\,'$. Then $R^T(\vec{u}) = \vec{u}\,'$. Thus $R^T(\vec{u}) \in U$ if $\vec{u} \in U$. QED

Lemma PP-16 Under the hypotheses of lemma PP-14, $R(U^\perp) \subseteq U^\perp$.

Proof:

Suppose $\vec{w} \in U^\perp$. We want to show $R(\vec{w}) \in U^\perp$. That is, $\vec{u} \cdot R(\vec{w}) = 0$ for all $\vec{u} \in U$. But $\vec{u} \cdot R(\vec{w}) = R^T(\vec{u}) \cdot \vec{w}$. By lemma PP-5, $R^T(\vec{u}) \in U$. By hypothesis $\vec{w} \perp U$. Hence, $R^T(\vec{u}) \cdot \vec{w} = 0$, as required.

Lemma PP-17 If $R \in \Theta(V)$ and λ is an eigenvalue of R, $\lambda = \pm 1$.

Proof:

Suppose $R(\vec{v}) = \lambda\vec{v}$ and $\vec{v} \neq \vec{0}$. Then $\|\vec{v}\| = \|R(\vec{v})\| = \|\lambda\vec{v}\| = |\lambda| \|\vec{v}\|$ so $1 = |\lambda|$. QED.

Lemma PP-18 Let $R \in \Theta(V)$. Let $\hat{x}_1, \ldots \hat{x}_n$ be an orthonormal basis for V. Let R_{ij} be the matrix of R relative to this basis. Suppose λ is a complex zero of the n'th degree polynomial $\det(R_{ij} - \lambda\delta_{ij})$. Then $|\lambda|^2 = 1$.

Proof:

There is a complex n-tuple $(r_1, \ldots, r_n) \neq (0, \ldots, 0)$ such that $R_{ij}r_j = \lambda r_i$. Taking complex conjugates gives $R_{ij}r_j^* = \lambda^* r_i^*$, or $R_{ik}r_k^* = \lambda^* r_i^*$. Multiplying and summing over i gives

$$R_{ij}R_{ik}r_j r_k^* = \lambda\lambda^* r_i r_i^*.$$

But $R_{ij}R_{ik} = \delta_{jk}$, so $r_j r_j^* = |\lambda|^2 r_i r_i^*$. Since $r_i r_i^* > 0$, we have $|\lambda|^2 = 1$.

Lemma PP-19 In lemma PP-18, λ^* is also a zero of $\det(R_{ij} - \lambda\delta_{ij}) = 0$.

Proof:

Since $(R_{ij} - \lambda^*\delta_{ij})r_j^* = 0$ and $(r_1^*, \ldots, r_n^*) \neq (0, \ldots 0)$, $\det(R_{ij} - \lambda^*\delta_{ij}) = 0$.

Lemma PP-20 Suppose R, R_{ij} as in lemma PP-18. Suppose λ is not real and $(r_1, \ldots, r_n) \neq (0, \ldots, 0)$ is a sequence of n complex numbers such that $R_{ij}r_j = \lambda r_i$. Then $r_i r_i = 0$.

Proof:

$R_{ik}r_k = \lambda r_i$ so $(R_{ij}r_j)(R_{ik}r_k) = \lambda^2 r_i r_i$, or $R_{ij}R_{ik}r_j r_k = \lambda^2 r_i r_i$, or $\delta_{jk}r_j r_k = \lambda^2 r_i r_i$, or $r_j r_j = \lambda^2 r_i r_i$, or $(\lambda^2 - 1)r_i r_i = 0$. Since λ is not real, $\lambda^2 - 1 \neq 0$, so $r_i r_i = 0$.

Remark 8 Suppose $R \in \Theta(V)$ and V contains no eigenvectors of R. Then V contains two mutually orthogonal unit vectors $\vec{x}$ and $\vec{y}$ such that for some θ in $0 < \theta < \pi$,

$$\begin{cases} R(\hat{x}) = \hat{x}\cos\Theta + \hat{y}\sin\Theta \\ \\ R(\hat{y}) = \hat{x}\sin\Theta + \hat{y}\cos\Theta. \end{cases} \tag{PP-23}$$

Proof:

Choose an orthonormal basis $\hat{z}_1, \dots, \hat{z}_n$ for V, and let R_{ij} be the matrix of R relative to that basis. If $0 = \det(R_{ij} - \lambda\delta_{ij})$ has a real root, λ, it is an eigenvalue of R, so R has an eigenvector. Hence no root of $\det(R_{ij} - \lambda\delta_{ij})$ is real. By lemma PP-18, we can write any root as $e^{i\theta}$ where $0 < \theta < \pi$ or $\pi < \theta < 2\pi$. By lemma PP-19, we can find a root $e^{-i\theta}$ with $0 < \theta < \pi$. Then there is an n-tuple of complex numbers $(r_1, \dots, r_n) \neq (0, \dots, 0)$ such that

$$R_{jk} r_j = e^{-i\theta} r_k. \tag{PP-24}$$

We may normalize $(r_1, \dots, r_n)$ so that

$$r_j^* r_j = 2 \tag{PP-25}$$

and by lemma PP-20

$$r_j r_j = 0. \tag{PP-26}$$

Now write $r_j = x_j + i y_j$ where x_j and y_j are real.

Then separating the real and imaginary parts of (PP-24) gives

$$\begin{aligned} R_{jk} x_j &= x_k \cos\theta + y_k \sin\theta \\ \\ R_{jk} y_j &= -x_k \sin\theta + y_k \cos\theta. \end{aligned}$$

Separating the real and imaginary parts of (PP-26) gives

$$\begin{aligned} x_j x_j - y_j y_j &= 0 \\ \\ x_j y_j &= 0 \end{aligned}$$

and (PP-25) gives

$$x_j x_j + y_j y_j = 2.$$

From these equations, clearly

$$x_j x_j = y_j y_j = 1 \text{ and } x_j y_j = 0.$$

Now take $\hat{x} = x_j \hat{z}_j$ and $\hat{y} = y_j \hat{z}_j$, and we have the $\hat{x}, \hat{y}, \theta$ whose existence is asserted in the theorem.

Corollary PP-9 $\dim V \geq 2$.

Remark 9 Suppose $R \in \Theta(V)$ and V contains no eigenvectors of R. Then V has an orthonormal basis $\hat{x}_1, \hat{y}_1, \ldots \hat{x}_m, \hat{y}_m$ such that for each $j \in \{1 \ldots, m\}$ there is an Θ_j in $0 < \theta_j < \pi$ with

$$\left\{ \begin{array}{l} R(\hat{x}_j) = \hat{x}_{(j)} \cos \theta_{(j)} + \hat{y}_{(j)} \sin \theta_{(j)} \\ \\ R(\hat{y}_j) = -\hat{x}_{(j)} \sin \theta_{(j)} + \hat{y}_{(j)} \cos \theta_{(j)} \end{array} \right. . \tag{PP-27}$$

Proof:

By remark PP-4 we can find mutually orthogonal unit vectors $\hat{x}_1, \hat{y}_1$ and θ_1 in $0 < \theta_1 < \pi$ such that (PP-27) holds for $j = 1$. Then $R(sp\{\hat{x}_1, \hat{y}_1\}) \subseteq sp\{\hat{x}_1, \hat{y}_1\}$. Hence, by lemma PP-16, $R(\{\hat{x}_1, \hat{y}_1\}^\perp) \subseteq \{\hat{x}_1, \hat{y}_1\}^\perp$. Hence, by lemma PP-14, $R|\{\hat{x}_1, \hat{y}_1\}^\perp \in \Theta(\{\hat{x}_1, \hat{y}_1\}^\perp)$. Since R has no eigenvectors in V, it has none in $\{\hat{x}_1, \hat{y}_1\}^\perp$. Thus $R|\{\hat{x}_1, \hat{y}_1\}^\perp$ satisfies the hypotheses of remark PP-3, and we can repeat the argument, finding θ_2 in $0 < \theta_2 < \pi$ an mutually orthogonal unit vectors $\hat{x}_2, \hat{y}_2$ in $\{\hat{x}_1, \hat{y}_1\}^\perp$ such that (PP-4) holds for $j = 2$. We can proceed in this way as long as the dimension of the remaining space is > 0. (The argument is by induction on $\dim V$.) When the dimension of the remaining space is 0, we have an orthonormal basis for V with the required properties.

Corollary PP-10 $\dim V$ is even. That is, if $R \in \Theta(V)$ and $\dim V$ is odd, R has an eigenvector. (This is obvious from another point of view. If $\dim V$ is odd, so is the degree of the real polynomial $\det(R - \lambda I)$. Hence, it has at least one real zero).

Theorem PP-14 Suppose V is Euclidean vector space and $R \in \Theta(V)$. Then V has an orthonormal basis $\hat{x}_1, \hat{y}_1, \ldots, \hat{x}_m, \hat{y}_m, \hat{z}_1, \ldots \hat{z}_n, \hat{w}_1 \ldots, \hat{w}_p$ with these properties:

i) $R(\hat{w}_j) = -\hat{w}_j$

ii) $R(\hat{z}_j) = \hat{z}_j$

iii) for $j \in \{1, \ldots, m\}$, there is a θ_j in $0 < \theta_j < \pi$ such that equation (PP-4) holds.

Proof:

Let Z and W be the eigenspaces of V with eigenvalues $+1$ and -1. Let $\hat{z}_1, \ldots, \hat{z}_n$ be an orthonormal basis for Z and $\hat{w}_1, \ldots, \hat{w}_p$ an orthonormal basis for W. Then $R(Z+W) \subseteq Z+W$, so $R[(Z+W)^\perp] \subseteq (Z+W)^\perp$. Moreover, $(Z+W)^\perp$ contains no eigenvectors of R, since all their eigenvalues are $+1$ or -1, and these are already in Z or W. Thus remark PP-4 applied to $(Z+W)^\perp$. QED.

Corollary PP-11 $\det R = (-1)^p$.

Proof:

Relative to the basis in theorem PP-15, the matrix of R looks thus:

Each 2×2 diagonal block contributes a factor $\cos^2 \theta_j + \sin^2 \theta_j = 1$ to the determinant, whence the result.

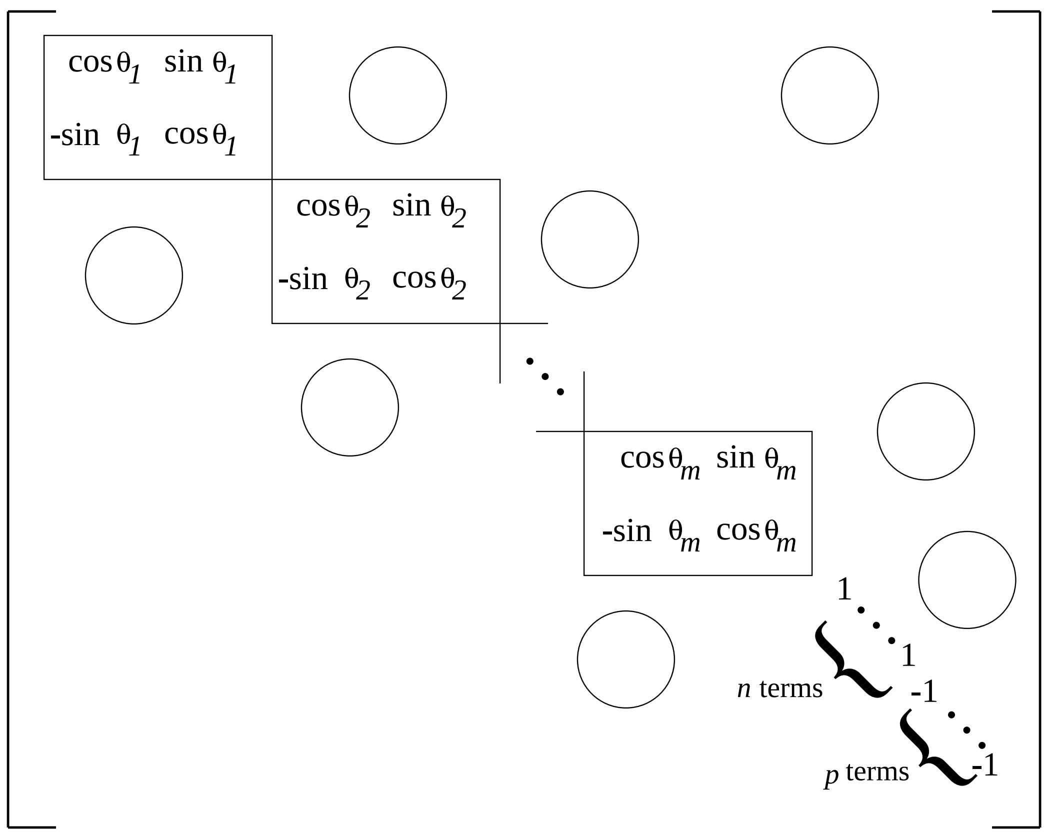


Figure PP-1: Matrix of R.

Corollary PP-12 If $R \in \Theta^+(V)$, then V has an orthonormal basis $\hat{x}_1, \hat{y}_1, \ldots, \hat{x}_m, \hat{y}_m, \hat{z}_1, \ldots, \hat{z}_n$ such that

$$\left\{\begin{array}{l} \textit{i)}\ R(\hat{z}_j) = \hat{z}_j \text{ for } j \in \{1, \ldots, n\}. \\ \\ \text{ii) For } j \in \{1, \ldots, m\} \text{ there is a } \theta_j \text{ in } 0 < \theta_j \leq \pi \text{ such that} \\ \\ \quad R(\hat{x}_j) = \hat{x}_{(j)} \cos\theta_{(j)} + \hat{y}_{(j)} \sin\theta_{(j)} \\ \\ \quad R(\hat{y}_j) = -\hat{x}_{(j)} \sin\theta_{(j)} + \hat{y}_{(j)} \cos\theta_{(j)}. \end{array}\right. \qquad \text{(PP-28)}$$

Proof:

Since $\det R = +1$, in theorem PP-15, p is even. Therefore, we may divide $\hat{w}_1, \ldots, \hat{w}_p$ into pairs. For any such pair, we have (for example)

$$\begin{aligned} R(\hat{w}_1) &= -\hat{w}_1 = \hat{w}_1 \cos\pi + \hat{w}_2 \sin\pi \\ R(\hat{w}_2) &= -\hat{w}_2 = -\hat{w}_1 \sin\pi + \hat{w}_2 \cos\pi. \end{aligned}$$

If we adjoin these pairs to the pairs $(\hat{x}_j, \hat{y}_j)$ given by theorem PP-15, and take the corresponding angles as $\theta = \pi$, we have the result of the corollary.

For obvious reasons, a linear operator on V which looks like (PP-27) in some orthonormal basis is called a "rotation." Corollary PP-12 says that every proper orthogonal operator is a rotation. If $\dim V = 3$, the eigenvector $\hat{z}_1$ is called the axis of the rotation, and the angle $\dot{\theta}_1$ in $0 < \theta_1 \leq \pi$ is the angle of the rotation. This terminology breaks down only when $m = 0$ and $n = 3$ in (PP-27). In that case, $R = I$. The angle of rotation of I is 0, but its axis of rotation obviously cannot be defined.

For $\dim V = 3$, corollary PP-12 is called Euler's theorem. If $R \in L(V \to V)$ and an orthonormal basis $\hat{x}_1, \ldots, \hat{x}_n$ for V can be found such that $R(\hat{x}_1) = -\hat{x}_1$, $R(\hat{x}_j) = \hat{x}_j$ for $j \geq 2$ then R is said to be a reflection along $\hat{x}_1$. Obviously $\det R = -1$ for a reflection, so a reflection is an "improper orthogonal" operator. Also, clearly a reflection has the property $R^{-1} = R$, or $R^2 = I$.

If R_1 is a reflection and R_2 is any other improper orthogonal operator, then $R = R_1 R_2$ is orthogonal and $\det R = (\det) R_1)(\det R_2) = (-1)^2 = 1$, so R is proper orthogonal. But then $R_1 R = R_1^2 R_2 = R_2$, so $R_2 = R_1 R$. That is, every improper orthogonal operator is the product of a rotation and a reflection.

Part I

Tensors over Euclidean Vector Spaces

Chapter 1

Multilinear Mappings

1.1 Definition of multilinear mappings

$$\mathcal{M}(V_1 \times \ldots \times V_q \to W).$$

$V_1, \ldots, V_q, W$ are vector spaces over a single field F, and $M \in F(V_1 \times \ldots \times V_q \to W)$. If M satisfies any of the following three equivalent conditions, M is called a multilinear mapping from $V_1 \times \ldots \times V_q$ to W.

Definition 1.1.6 For any $(\vec{v}_1, \ldots, \vec{v}_q) \in V_1 \times \ldots \times V_q$ and any $p \in \{1, \ldots, q\}$,

$$M(\vec{v}_1, \ldots, \vec{v}_{p-1}, \cdot, \vec{v}_{p+1}, \ldots, \vec{v}_q) \in \mathcal{L}(V_p \to W).$$

(That is, $M(\vec{v}_1, \ldots, \vec{v}_q)$ depends linearly on each of the separate vectors $\vec{v}_1, \ldots, \vec{v}_q$ if the others are held fixed.)

Definition 1.1.7 If $a \in F$, $p \in \{1, \ldots, q\}$, $\vec{x}_p$ and $\vec{y}_p \in V_p$, and $(\vec{v}_1, \ldots, \vec{v}_q) \in V_1 \times \ldots \times V_q$ then

i) $M(\vec{v}_1, \ldots, \vec{v}_{p-1}, a\vec{v}_p, \vec{v}_{p+1}, \ldots, \vec{v}_q) = aM(\vec{v}_1, \ldots, \vec{v}_{p-1}, \vec{v}_p, \vec{v}_{p+n}, \ldots, \vec{v}_q)$. Also

ii) $M(\vec{v}_1, \ldots, \vec{v}_{p-1}, \vec{x}_p + \vec{y}_p, \vec{v}_{p+1}, \ldots, \vec{v}_q) = M(\vec{v}_1, \ldots, \vec{v}_{p-1}, \vec{x}_p, \vec{v}_{p+1}, \ldots, \vec{v}_q)$
$+ M(\vec{v}_1, \ldots, \vec{v}_{p-1}, \vec{y}_p, \vec{v}_{P+1}, \ldots, \vec{v}_q)$.

Definition 1.1.8 For each $p \in \{1, \ldots, q\}$, suppose $a^1_{(p)}, \ldots, a^{n_p}_{(p)} \in F$ and $\vec{v}_1^{(p)}, \ldots, \vec{v}_{n_p}^{(p)} \in V_p$. Then $M\left(a^{i_1}_{(1)}\vec{v}^{(1)}_{i_1}, \ldots, a^{i_q}_{(q)}\vec{v}^{(q)}_{i_q}\right) = a^{i_1}_{(1)} \ldots a^{i_q}_{(q)} M\left(\vec{v}^{(1)}_{i_1}, \ldots, \vec{v}^{(q)}_{iq}\right)$.

Definition 1.1.9 $\mathcal{M}(V_1 \times \ldots \times V_q \to W)$ denotes the set of all multilinear mappings from $V_1 \times \ldots \times V_q$ to W.

1.2 Examples of multilinear mappings

In these examples, $V_1, \ldots, V_q$ are Euclidean vector spaces, and the field F is $\mathcal{R}$. The vector space W need not be Euclidean (i.e. need not have a dot product) but it must be real (i.e. its field is $\mathcal{R}$).

Example 1.2.1 For each p between 1 and q, let $\vec{u}_p$ be a fixed vector in V_p. Let $\vec{w}_0 \in W$. For any $(\vec{v}_1, \ldots, \vec{v}_q) \in V_1 \times \ldots \times V_q$, define

$$M(\vec{v}_1, \ldots, \vec{v}_q) = (\vec{u}_1 \cdot \vec{v}_1) \cdots (\vec{u}_q \cdot \vec{v}_q)\vec{w}_0. \tag{1.2.1}$$

To give the next example, some terminology is useful. Suppose that for each $p \in \{1, \ldots, q\}$ $B_p = \left(\vec{b}_1^{(p)}, \ldots, \vec{b}_{n_p}^{(p)}\right)$ is an ordered basis for V_p. Then the sequence of ordered bases $(B_1, \ldots, B_q)$ is called a "basis sequence" for $V_1 \times \ldots \times V_q$. Let $B_p^D = (\vec{b}^1_{(p)}, \ldots, \vec{b}^{n_p}_{(p)})$ be the ordered basis for V_p which is dual to B_p. Then $(B_1^D, \ldots, B_q^D)$ is "the basis sequence dual to $(B_1, \ldots, B_q)$."

Example 1.2.2 Let $(B_1, \ldots, B_q)$ be a basis sequence for $V_1 \times \ldots \times V_q$. Let $\vec{w}_{i_1 \ldots i_q}$ be any q'th order, $n_1 \times \ldots \times n_q$ array of vectors from W. For any $(\vec{V}_1, \ldots, \vec{V}_q) \in V_1 \times \ldots \times V_q$ define

$$M\left(\vec{v}_{(1)}, \ldots, \vec{v}_{(q)}\right) = \left(\vec{b}^{i_1}_{(1)} \cdot \vec{v}_{(1)}\right) \ldots \left(\vec{b}^{i_q}_{(q)} \cdot \vec{v}_{(q)}\right) \vec{w}_{i_1 \ldots i_q}. \tag{1.2.2}$$

Then

$$M \in \mathcal{M}\left(V_1 \times \ldots \times V_q \to W\right) \tag{1.2.3}$$

and also

$$M\left(\vec{b}^{(1)}_{j_1}, \ldots \vec{b}^{(1)}_{j_q}\right) = \vec{w}_{j_1 \ldots j_q}. \tag{1.2.4}$$

The multilinearity of M is almost obvious; (1.2.2) is a sum of $n_1 n_2 \dots n_q$ terms like (1.2.1). Each of these terms is linear in each $\vec{v_p}$ if $\vec{v}_1, \dots, \vec{v}_{p-1}, \vec{v}_{p+1}, \dots \vec{v}_q$ are fixed. To prove (1.2.4) we let $\vec{v}_p = \vec{b}^{(p)}_{j_p}$ in the definition (1.2.2), obtaining $M(\vec{b}^{(1)}_{j_1}, \dots, \vec{b}^{(q)}_{j_q}) = (\vec{b}^{i_1}_{(1)} \cdot \vec{b}^{(1)}_{j_1}) \dots (\vec{b}^{i_q}_{(q)} \cdot \vec{b}^{(q)}_{j_q}) \vec{w}_{i_1 \dots i_q} = \delta^{i_1}{}_{j_1} \dots \delta^{i_q}{}_{j_q} \vec{w}_{i_1 \dots i_q} = \vec{w}_{j_1 \dots j_q}$. There are no other examples. Every member of $\mathcal{M}(V_1 \times \dots \times V_q \to W)$ is like example (1.2.2). We have

Remark 1.2.10 Suppose $V_1, \dots, V_q$ are Euclidean vector spaces and W is a real vector space. Suppose $\dim V_1 = n_1, \dots, \dim V_q = n_q$. Suppose $(B_1, \dots, B_q)$ is a basis sequence for $V_1 \times \dots \times V_q$, and $\vec{w}_{i_1 \dots i_q}$ is an $n_1 \times n_2 \times \dots \times n_q$ array of vectors from W. Then $\exists 1$ $M \in \mathcal{M}(V_1 \times \dots \times V_q \to W) \ni M(\vec{b}^{(1)}_{i_1}, \dots, \vec{b}^{(q)}_{i_q}) = \vec{w}_{i_1 \dots i_q}$.

Proof:

Example 1.2.2 shows at least one such M. To establish uniqueness suppose M does satisfy (1.2.3) and (1.2.4). Let $(\vec{b}^1_{(p)}, \dots, \vec{b}^{n_p}_{(p)}) = B^D_p$ be the ordered basis for V_p which is dual to $B_p = (\vec{b}^{(p)}_1, \dots, \vec{b}^{(p)}_{n_p})$. For any $\vec{v}_{(p)} \in V_p$, we have $\vec{v}_{(p)} = v^{i_p}_{(p)} \vec{b}^{(p)}_{i_p}$ where $\vec{v}^{i_p}_{(p)} = v_{(p)} \cdot \vec{b}^{i_p}_{(p)}$. Then

$$
\begin{aligned}
M(\vec{v}_{(1)}, \dots, \vec{v}_{(q)}) &= M(v^{i_1}_{(1)} \vec{b}^{(1)}_{i_1}, \dots, v^{i_q}_{(q)} \vec{b}^{(q)}_{i_q}) \quad \text{(by definition 1.1.8)} \\
&= b^{i_1}_{(1)} \dots v^{i_q}_{(q)} M(\vec{b}^{(1)}_{i_1}, \dots, \vec{b}^{(q)}_{i_q}) \quad \text{(by hypothesis)} \\
&= v^{i_1}_{(1)} \dots v^{i_q}_{(q)} \vec{w}_{i_1 \dots i_q} \\
&= \left(\vec{b}^{i_1}_{(1)} \cdot \vec{v}_{(1)} \right) \dots \left(\vec{b}^{i_q}_{(q)} \cdot \vec{v}_{(q)} \right) \vec{w}_{i_1 \dots i_q}.
\end{aligned}
$$

In other words, M is the function defined by (1.2.2).

It follows that every $M \in \mathcal{M}(V_1 \times \dots \times V_q \to W)$ has the form (1.2.2).

1.3 Elementary properties of multilinear mappings

Remark 1.3.11 If one of $\vec{v}_1, \ldots, \vec{v}_q$ is $\vec{0}$ and if $M \in \mathcal{M}(V_1 \times \ldots \times V_q \to W)$ then $M(\vec{v}_1, \ldots, \vec{v}_q) = \vec{0}$, the zero in W.

Proof:

We suppose $\vec{v}_1 = \vec{0}$. Then $M(\vec{0}, \vec{v}_2, \ldots, \vec{v}_q) = M(0\vec{0}, \vec{v}_2, \ldots, \vec{v}_q) = 0M(\vec{0}, \vec{v}_2, \ldots, \vec{v}_q) = \vec{0}$, where we have used definition (1.1.7).

Remark 1.3.12 $\mathcal{M}(V_1 \times \ldots \times V_q \to W)$ is a vector space over F if we use the *ad* and *sc* defined in example 3, page D-12, or in equations (14.1), (14.2).

Proof:

Refer to "subspaces" on page D-15. We know from example 3, page D-12, that $F(V_1 \times \ldots \times V_q \to W)$ is a vector space over F. Obviously, $\mathcal{M}(V_1 \times \ldots \times V_q \to W)$, is a subset of $F(V_1 \times \ldots \times Vq) \to W)$. Therefore, it suffices to prove that if $a \in F$ and M, $N \in \mathcal{M}(V_1 \times \ldots \times Vq \to W)$, then aM and $M + N \in \mathcal{M}(V_1 \times \ldots \times V_q \to W)$. That is, if M and N are multilinear, so are aM and $M + N$. This is immediate from definition 1.1.6 because we know it for linear functions. Alternatively, we can use definition 1.1.8 and write everything out. We will do this for $M + N$. We have

$$\begin{aligned}
& (M+N)\left(a_{(1)}^{i_1}\vec{v}_{i_1}^{(1)}, \ldots, a_{(q)}^{i_q}\vec{v}_{i_q}^{(q)}\right) \\
= \ & M\left(a_{(1)}^{i_1}\,\vec{v}_{i_1}^{(1)}, \ldots a_{(q)}^{i_q}\,\vec{v}_{i_q}^{(q)}\right) + N\left(a_{(1)}^{i_1}\vec{v}_{i_1}^{(1)}, \ldots, a_{(q)}^{i_q}\vec{v}_{i_q}^{(q)}\right) \\
& \qquad \text{(by definition of } M+N\text{)} \\
= \ & a_{(1)}^{i_1} \ldots a_{(q)}^{i_q} M\left(\vec{v}_{i_1}^{(1)}, \ldots, \vec{v}_{i_q}^{(q)}\right) + a_{(1)}^{i_1} \ldots a_{(q)}^{i_q} N\left(\vec{v}_{i_1}^{(1)}, \ldots, \vec{v}_{i_q}^{(q)}\right) \\
& \qquad \text{(because } M \text{ and } N \text{ are multilinear)} \\
= \ & a_{(1)}^{i_1} \ldots a_{(a)}^{i_q}\left[M\left(\vec{v}_{i_1}^{(1)}, \ldots, \vec{v}_{i_q}^{(q)}\right) + N\left(\vec{v}_{i_1}^{(1)}, \cdots, \vec{v}_{i_q}^{(q)}\right)\right] \\
& \qquad \text{because } W \text{ is a vector space over } F
\end{aligned}$$

$$= \quad a^{i_1}_{(1)} \dots a^{i_q}_{(q)} \left[(M+N) \left(\vec{v}^{(1)}_{i_1}, \dots, \vec{v}^{(q)}_{i_q} \right) \right]$$

by definition of $M+N$.

Comparing the two ends of this chain of equalities, we have a proof that $M+N$ is multilinear if M and N are, and if $M+N$ is defined by (14.1).

1.4 Permuting a multilinear mapping

Definition 1.4.10 Suppose $(\vec{v}_1, \dots, \vec{v}_q) \in V_1 \times \dots \times V_q$ and $1 \le r < s \le q$. The q-tuple $(\vec{v}_1, \dots, \vec{v}_{r-1}, \vec{v}_s, \vec{v}_{r+1} \dots, \vec{v}_{s-1}, \vec{v}_r, \vec{v}_{s+1}, \dots \vec{v}_q)$ will be abbreviated $(\vec{v}_1, \dots \vec{v}_q)_{(rs)}$. It is simply $(\vec{v}_1, \dots, \vec{v}_q)$ with $\vec{v}_r$ and $\vec{v}_s$ interchanged. Note that if $V_r = V_s$, then $(\vec{v}_1, \dots, \vec{v}_q)_{(rs)} \in V_1 \times \dots \times V_q$.

Definition 1.4.11 Suppose $M \in \mathcal{M}(V_1 \times \dots \times V_q \to W)$ and $1 \le r < s \le q$. Suppose $V_r = V_s$. Then we write $(rs)M$ for the member of $F(V_1 \times \dots \times V_q \to W)$ defined by requiring for any $(\vec{v}_1, \dots, \vec{v}_q) \in V_1 \times \dots \times V_q$ that

$$[(rs)M] (\vec{v}_1, \dots, \vec{v}_q) = M (\vec{v}_1, \dots, \vec{v}_q)_{(rs)} . \tag{1.4.1}$$

That is

$$[(rs)M] (\vec{v}_1, \dots, \vec{v}_{r-1}, \vec{v}_r, \vec{v}_{r+1}, \dots, \vec{v}_{s-1}, \vec{v}_s, \vec{s}_{s+1}, \dots, \vec{v}_q)$$

$$= M (\vec{v}_1, \dots, \vec{v}_{r-1}, \vec{v}_s, \vec{v}_{r+1}, \dots, \vec{v}_{s-1}, \vec{v}_r, \vec{s}_{s+1}, \dots, \vec{v}_q) . \tag{1.4.2}$$

Remark 1.4.13 $(rs)M \in \mathcal{M}(V_1 \times \dots \times V_q \to W)$ if $V_r = V_s$.

Proof:

Obvious from (1.4.2) and definition (1.1.6).

Remark 1.4.14 (rs) is a linear operator on $\mathcal{M}(V_1 \times \dots \times V_q \to W)$ if $V_r = V_s$.

Proof:

By remark (1.4.13), $(rs) : \mathcal{M}(V_1 \times \ldots \times V_q \to W) \to \mathcal{M}(V_1 \times \ldots \times V_q \to W)$. It remains to prove that (rs) is linear. Suppose $a_1, \ldots, a_N$ are scalars and $M_1, \ldots M_N \in \mathcal{M}(V_1 \times \ldots \times V_q \to W)$. We want to prove that $(rs)(a_j M_j) = a_j[(rs)M_j]$. Let $(\vec{v}_1, \ldots, \vec{v}_q \in V_1 \times \ldots \times V_q)$. Then

$$[(rs)(a_j M_j)]\,(\vec{v}_1, \ldots, \vec{v}_q) = (a_j M_j)(\vec{v}_1, \ldots, \vec{v}_q)_{(rs)} = a_j \left[M_j(\vec{v}_1, \ldots, \vec{v}_q)_{(rs)}\right]$$

$$= a_j \left\{ \Big[(rs)M_j\Big](\vec{v}_1, \ldots, \vec{v}_q) \right\} = \left\{ a_j \Big[(rs)M_j\Big] \right\}(\vec{v}_1, \ldots, \vec{v}_q).$$

Since this is true for any $(\vec{v}_1, \ldots, \vec{v}_q) \in V_1 \times \ldots \times V_q$, we have the required result.

Definition 1.4.12 Suppose M, r, s as in definition (1.4.2). If $(rs)M = M$, M is "symmetric" under (rs). If $(rs)M = -M$, M is "antisymmetric" under (rs).

Remark 1.4.15 If $V_r = V_s$, the members of $\mathcal{M}(V_1 \times \ldots \times V_q \to W)$ which are symmetric under (rs) constitute a subspace of $\mathcal{M}(V_1 \times \ldots \times V_q \to W)$. The same holds for the members antisymmetric under (rs).

Proof:

It suffices to show that if a is any scalar and M, N are symmetric (or antisymmetric) under (rs), so are aM and $M+N$. We give the proof for antisymmetry. In that case, by hypothesis,

$$(rs)M = -M \quad , \quad (rs)N = -N.$$

But (rs) is linear, so $(rs)[aM] = a[(rs)M] = a(-M) = -aM$, and $(rs)(M + N) = (rs)M + (rs)N = -M - N = -(M + N)$. (Note that the set of members of $\mathcal{M}(V_1 \times \ldots \times V_q \to W)$ which are symmetric (anti-symmetric) under (rs) is simply the eigenspace of (rs) with eigenvalue $1(-1)$. Of course, it is a subspace of $\mathcal{M}(V_1 \times \ldots \times V_q \to W)$.)

Definition 1.4.13 Suppose $V_1 = \cdots = V_q = V$. Suppose $M \in \mathcal{M}(\times^q V \to W)$ and $\pi \in S_q$. Then πM is defined as that member of $\mathcal{F}(\times^q V \to W)$ such that for any $\vec{v}_1, \ldots \vec{v}_q \in V$

$$(\pi M)(\vec{v}_1, \ldots, \vec{v}_q) = M\left(\vec{v}_{\pi(1)}, \ldots, \vec{v}_{\pi(q)}\right).$$

Remark 1.4.16 If $V_1 = \cdots = V_q = V$ then $\pi M \in \mathcal{M}(\times^q V \to W)$, and in fact π is a linear operator on $\mathcal{M}(\times^q V \to W)$. Moreover if σ, $\pi \in S_q$ and $M \in \mathcal{M}(\times^q V \to W)$ then

$$\sigma(\pi M) = (\sigma\pi)M. \tag{1.4.3}$$

Proof:

Except for (1.4.3), the proof is like that of remark (1.4.14). To prove equation (1.4.3), let $\vec{v}_1, \ldots, \vec{v}_q \in V$ and let $\vec{w}_1 = \vec{v}_{\sigma(1)}, \ldots, \vec{w}_q = \vec{v}_{\sigma(q)}$. Then

$$\begin{aligned}
\left[\sigma(\pi M)\right](\vec{v}_1, \ldots, \vec{v}_q) &= (\pi M)\left(\vec{v}_{\sigma(1)}, \ldots, \vec{v}_{\sigma(q)}\right) \\
&= (\pi M)(\vec{w}_1, \ldots, \vec{w}_q) = M\left(\vec{w}_{\pi(1)}, \ldots, \vec{w}_{\pi(q)}\right) \\
&= M\left(\vec{v}_{\sigma[\pi(1)]}, \ldots, \vec{v}_{\sigma[\pi(q)]}\right) \\
&= M\left(\vec{v}_{(\sigma\pi)(1)}, \ldots, \vec{v}_{(\sigma\pi)(q)}\right) \\
&= [(\sigma\pi)M](\vec{v}_1, \ldots, \vec{v}_q)
\end{aligned}$$

Since this is true for all $\vec{v}_1, \ldots, \vec{v}_q \in V$, we have (1.4.3).

Definition 1.4.14 Suppose $V_1 = \cdots = V_q = V$ and $M \in \mathcal{M} \in (\times^q V \to W)$. M is totally symmetric (antisymmetric) if $\pi M = M$ $(\pi M = (\operatorname{sgn}\pi)M\ )$ for every $\pi \in S_q$.

Remark 1.4.17 M is totally symmetric (antisymmetric) iff $(rs)M = M(= -M)$ for every transposition (rs).

Proof:

> $\Longrightarrow$ is obvious. For $\Longleftarrow$, recall that any $\pi \in S_q$ is a product of transpositions. By remark 1.4.16 the effects of these transpositions on M can be calculated one after the other. If all leave M unchanged, so does π. If all multiply M by -1, then πM is M multiplied by $\operatorname{sgn} \pi$ $(= \pm 1$ for even/odd $\pi)$.

Remark 1.4.18 The totally symmetric (antisymmetric) members of $\mathcal{M}(\times^q V \to W)$ constitute a subspace of $\mathcal{M}(\times^q V \to W)$.

Proof:

> We give the proof for the totally symmetric case. Suppose $M_1, \ldots, M_N$ are totally antisymmetric members of $\mathcal{M}(\times^q V \to W)$ and $a_1, \ldots, a_M$ are scalars. We want to show that $a_i M_i$ is totally antisymmetric. But for any $\pi \in S_q$, π is linear operator on $\mathcal{M}(\times^q V \to W)$, so $\pi(a_i M_i) = a_i(\pi M_i) = a_i[(\operatorname{sgn} \pi) M_i] = (\operatorname{sgn} \pi)(a_i M_i)$. That is, $a_i M_i$ is totally antisymmetric.

Chapter 2

Definition of Tensors over Euclidean Vector Spaces

If $V_1, \ldots, V_q$ are Euclidean vector spaces, then $\mathcal{M}(V_1 \times \ldots \times V_q) \to \mathcal{R}$ is called the "tensor product of $V_1, \ldots, V_q$". It is written $V_1 \otimes \ldots \otimes V_q$. Thus

$$V_1 \otimes \ldots \otimes V_q := \mathcal{M}(V_1 \times \ldots \times V_q) \to \mathcal{R}. \tag{2.0.1}$$

The multilinear functionals which are the members of $V_1 \otimes \ldots \otimes V_q$ are called the tensors of order q over $V_1 \times \ldots \times V_q$.

If $V_1 = \ldots = V_q = V$, then $V_1 \otimes \ldots \otimes V_q = \overbrace{V \otimes \ldots \otimes V}^{q \text{ times}}$, and this is abbreviated $\otimes^q V$. Its members are called the tensors of order q over V. By convention, tensors of order zero are scalars; that is

$$\otimes^\circ V = \mathcal{R}. \tag{2.0.2}$$

For tensors of order 1, the notation requires some comment. When $q = 1$, multilinearity is simply linearity. That is $\mathcal{M}(V \to \mathcal{R}) = \mathcal{L}(V \to \mathcal{R})$. Thus the tensors of order 1 over a vector space V are simply the linear functional on V. This would not be a problem, except that for $q = 1$, (2.0.1) reduces to $V_1 := \mathcal{M}(V_1 \to \mathcal{R})$, which says

$$V = \mathcal{L}(V \to \mathcal{R}). \tag{2.0.3}$$

It is at this point that we require V to be a Euclidean vector space. If it is, then $\vec{v} \in V$ can be safely confused with the linear functional $\lambda_{\vec{v}} \in \mathcal{L}(V \to \mathcal{R})$ defined by (21.2). Tensors of order 1 over V are simply the vectors in V.

When V is not a Euclidean vector space, (2.0.3) is false. (The situation is very subtle. If V is not Euclidean but is finite dimensional, every basis B for V generates an isomorphism between V and $\mathcal{L}(V \to \mathcal{R})$, namely $\vec{b}_i \mapsto c_B^i$. But each of these isomorphisms depends on the basis B chosen to produce it. There is no "natural", basis-independent isomorphism which can be used to identify vectors with linear functionals. See MacLane and Birkhoff, Algebra, MacMillan 1967, p. 237 for a more detailed discussion.) Tensors over vector spaces which are not Euclidean are constructed in a more complicated fashion than (2.0.1). We will not need this generality, and tensors over Euclidean spaces have some very useful properties for continuum mechanics. It is for this reason that we consider only tensors over Euclidean spaces. See M. Marcus's book, on reserve, for a treatment of tensors over arbitrary vector spaces.

Chapter 3

Alternating Tensors, Determinants, Orientation, and n-dimensional Right-handedness

In this section we discuss an application of tensors which some of you may have seen in your linear algebra courses.

Let V be an n-dimensional Euclidean vector space. The set of totally antisymmetric tensors of order q over V is written $\Lambda^q V$. It is a subspace of $\otimes^q V$. (See remark 1.4.18.) We want to study $\Lambda^n V$. The members of $\Lambda^n V$ are called "alternating tensors over V."

3.1 Structure of $\Lambda^n V$

We want to understand what the alternating tensors over V are like. It will turn out that there are not very many of them. In fact $\dim \Lambda^n V = 1$. We want to prove this and to find a basis "vector" in the vector space $\Lambda^n V$.

First we note a technique for verifying that $A \in \Lambda^n V$. We need to prove only that $A \in \otimes^n V$ and that $(rs)A = -A$ for every transposition $(rs) \in S_n$. If this latter fact is true, then clearly remark 2.0.1 tells us $A \in \Lambda^n V$.

A second property of an $A \in \otimes^n V$ which is equivalent to membership in $\Lambda^n V$ is that for any $(\vec{v}_1, \ldots, \vec{v}_n) \in \times^n V$ and any $\pi \in S_n$,

$$A\left(\vec{v}_{\pi(1)}, \ldots, \vec{v}_{\pi(n)}\right) = (\operatorname{sgn} \pi) A\left(\vec{v}_1, \ldots \vec{v}_n\right). \tag{3.1.1}$$

This is obvious from the definition of πA and $\Lambda^n V$. It implies the first step in our attempt to understand alternating tensors:

Lemma 3.1.21 If $A \in \Lambda^n V$ and two of $\vec{v}_1, \ldots \vec{v}_n$ are equal, then $A(\vec{v}_1, \ldots, \vec{v}_n) = 0$.

Proof:

> If $\vec{v}_r = \vec{v}_s$, interchanging them will not change $A(\vec{v}_1, \ldots, \vec{v}_n)$. But $\operatorname{sgn}(rs) = -1$, so (3.1.1) says interchanging them changes the sign of $A(\vec{v}_1, \ldots, \vec{v}_n)$. Hence $A(\vec{v}_1, \ldots, \vec{v}_n) = 0$.

Equation (3.1.1) and lemma 3.1.21 are equivalent to the single equation

$$A\left(\vec{v}_{i_1}, \ldots, \vec{v}_{i_n}\right) = \varepsilon_{i_1 \ldots i_n} A\left(v_1, \ldots, \vec{v}_n\right). \tag{3.1.2}$$

For if two of $i_1, \ldots, i_n$ are equal, both sides of (3.1.2) vanish. If $i_1, \ldots, i_n$ are all different, there is a permutation $\pi \in S_n$ such that $i_1, = \pi(1), \ldots, i_n = \pi(n)$. Then (3.1.2) reduces to equation (3.1.1).

Lemma 3.1.21 can be extended to

Remark 3.1.19 If $A \in \Lambda^n V$ and $\vec{v}_1, \ldots, \vec{v}_n$ are linearly dependent, then $A(\vec{v}_1, \ldots, \vec{v}_n) = 0$.

Proof:

> By hypothesis, there are scalars $c_1, \ldots, c_n$, not *all* zero, such that $c_i \vec{v}_i = \vec{0}$. If $\vec{v}_1 = \vec{0}$, then $A(\vec{v}_1, \ldots, \vec{v}_n) = 0$ by remark 1.3.11. If $\vec{v}_1 \neq 0$ then at least one of $c_2, \ldots, c_n$ is nonzero. Let c_p be the last scalar which is nonzero. Then $c_1 \vec{v}_1 + \ldots + c_p \vec{v}_p = \vec{0}$ and we can divide by c_p to write
>
> $$\vec{v}_p = \sum_{i=1}^{p-1} a_i \vec{v}_i$$

where

$$a_i = -\frac{c_i}{c_p}.$$

Then

$$\begin{aligned}
A(\vec{v}_1, \ldots, \vec{v}_n) &= A\left(\vec{v}_1, \ldots, \vec{v}_{p-1}, \sum_{i=1}^{p-1} a_i \vec{v}_i, \vec{v}_{p+1}, \ldots, \vec{v}_n\right) \\
&= \sum_{i=1}^{p-1} a_i A(\vec{v}_1, \ldots \vec{v}_{p-1}, \vec{v}_i, \vec{v}_{p+1}, \ldots, \vec{v}_n) \\
&= 0 \qquad \text{by lemma 3.1.21.}
\end{aligned}$$

Remark 3.1.19 would be rather dull if $A \in \Lambda^n V$ implied $A = 0$. Therefore at this point we pause to construct a nonzero member of $\Lambda^n V$. Let $B = (\vec{b}_1, \ldots, \vec{b}_n)$ be any basis for V. Let $B^D = (\vec{b}^1, \ldots, \vec{b}^n)$ be its dual basis. Define $A_B \in \mathcal{F}(\times^n V \to \mathcal{R})$ by requiring

$$A_B(\vec{v}_1, \ldots, \vec{v}_n) = \left(\vec{v}_1 \cdot \vec{b}^{i_1}\right) \ldots \left(\vec{v}_n \cdot \vec{b}^{i_n}\right) \varepsilon_{i_1 \ldots i_n} \tag{3.1.3}$$

for any $(\vec{v}_1, \ldots \vec{v}_n) \in \times^n V$. From definition (1.1.6) it is clear that A_B is multilinear, so $A_B \in \otimes^n V$, because the values of A_B are scalars. Furthermore, $A_B\left(\vec{b}_1, \ldots, \vec{b}_n\right) = \left(\vec{b}_1 \cdot \vec{b}^{i_1}\right) \ldots \left(\vec{b}_n \cdot \vec{b}^{i_n}\right) \varepsilon_{i_1 \ldots i_n} = \delta_1{}^{i_1} \ldots \delta_n{}^{i_n} \varepsilon_{i_1 \ldots i_n} = \varepsilon_{12 \ldots n} = 1$. That is, if $B = \left(\vec{b}_1, \ldots, \vec{b}_n\right)$,

$$A_B\left(\vec{b}_1, \ldots \vec{b}_n\right) = 1 \tag{3.1.4}$$

when A_B is defined by (3.1.3). Thus $A_B \neq 0$. We claim also that $A_B \in \Lambda^n V$. To show this, is suffices to show that $(rs)A_B = -A_B$ for any transposition (rs). The argument is the same for all (rs), so we consider (12). We have

$$\begin{aligned}
[(12)A_B](\vec{v}_1, \ldots, \vec{v}_n) &= A_B(\vec{v}_2, \vec{v}_1, \vec{v}_3, \ldots, \vec{v}_n) \\
&= \left(\vec{v}_2 \cdot \vec{b}^{i_1}\right)\left(\vec{v}_1 \cdot \vec{b}^{i_2}\right)\left(\vec{v}_3 \cdot \vec{b}^{i_3}\right) \ldots \left(\vec{v}_n \cdot \vec{b}^{i_n}\right) \varepsilon_{i_1 \ldots i_n}.
\end{aligned}$$

If we relabel the summation index i_1 as i_2, and relabel i_2 as i_1, this is

$$[(12)A_B](\vec{v}_1, \ldots, \vec{v}_n) = \left(\vec{v}_2 \cdot \vec{b}^{i_2}\right)\left(\vec{v}_1 \cdot \vec{b}^{i_1}\right)\left(\vec{v}_3 \cdot \vec{b}^{i_3}\right) \ldots \left(\vec{v}_n \cdot \vec{b}^{i_n}\right) \varepsilon_{i_2 i_1 i_3 \ldots i_n}$$

$$= \left(\vec{v}_1 \cdot \vec{b}^{i_1}\right) \dots \left(\vec{v}_n \cdot \vec{b}^{i_n}\right) \varepsilon_{i_2\, i_1\, i_3 \dots i_n}$$

$$= -\left(\vec{v}_1 \cdot \vec{b}^{i_1}\right) \dots \left(\vec{v}_n \cdot \vec{b}^{i_n}\right) \varepsilon_{i_1\, i_2\, i_3 \dots i_n}$$

$$= -A_B \left(\vec{v}_1, \dots, \vec{v}_n\right).$$

Since these equations hold for any $(\vec{v}_1, \dots, \vec{v}_n) \in \times^n V$, we have $(12)A_B = -A_B$. QED

Now for every ordered basis B in V we have managed to construct a nonzero member of $\Lambda^n V$, namely A_B. It looks as if $\Lambda^n V$ must be rather large. In fact it is not. We can now easily prove

$$\dim \Lambda^n V = 1 \quad \text{if} \quad n \dim V. \tag{3.1.5}$$

To see this, let $B = (\vec{b}_1, \dots, \vec{b}_n)$ be a fixed ordered basis for V. We will prove that if $A \in \Lambda^n V$ then A is a scalar multiple of A_B. Thus, $\{A_B\}$ is a basis for $\Lambda^n V$. The argument goes thus. Let $A \in \Lambda^n V$. For any $(\vec{v}_1, \dots, \vec{v}_n) \in \times^n V$ we can write $\vec{v}_p = v_p{}^i \vec{b}_i$ where $v_p{}^i = \vec{v}_p \cdot \vec{b}^i$, $B^D = (\vec{b}^1, \dots, \vec{b}^n)$ being the basis dual to B. Then

$$\begin{aligned}
A\left(\vec{v}_1, \dots, \vec{v}_n\right) &= A\left(v_1{}^{i_1} \vec{b}_{i_1}, \dots, v_n{}^{i_n} \vec{b}_{i_n}\right) \\
&= v_1{}^{i_1} \dots v_n{}^{i_n} A\left(\vec{b}_{i_1}, \dots, \vec{b}_{i_n}\right) \\
&= v_1{}^{i_1} \dots v_n{}^{i_n} \varepsilon_{i_1 \dots i_n} A\left(\vec{b}_1, \dots, \vec{b}_n\right) \qquad \text{by} \quad (3.1.2) \\
&= A\left(\vec{b}_1, \dots, \vec{b}_n\right) \left(\vec{v}_1 \cdot \vec{b}^{i_1}\right) \dots \left(\vec{v}_n \cdot \vec{b}^{i_n}\right) \varepsilon_{i_1 \dots i_n} \\
&= A\left(\vec{b}_1, \dots, \vec{b}_n\right) A_B\left(\vec{v}_1, \dots, \vec{v}_n\right).
\end{aligned}$$

Since this is true for all $\vec{v}_1, \dots, \vec{v}_n \in V$, we have for ordered basis $B = (\vec{b}_1, \dots \vec{b}_n)$ in V

and for any $A \in \Lambda^n V$

$$A = A\left(\vec{b}_1, \ldots, \vec{b}_n\right) A_B. \tag{3.1.6}$$

Of course, $A\left(\vec{b}_1, \ldots, \vec{b}_n\right) \in \mathcal{R}$, so (3.1.6) shows $A \in sp\{A_B\}$. This proves (3.1.5).

Equation 3.1.6 permits us to prove the following test for linear independence of a sequence of n vectors in an n-dimensional space.

Remark 3.1.20 Suppose $A \in \Lambda^n V$ and $A \neq 0$, and $\dim V = n$. Suppose $\vec{v}_1, \ldots, \vec{v}_n \in V$. Then $\vec{v}_1, \ldots, \vec{v}_n$ are linearly dependent if and only if $A(\vec{v}_1, \ldots, \vec{v}_n) = 0$.

Proof:

> $\Longrightarrow$ is remark 3.1.2. To prove $\Longleftarrow$, we want to prove that if $A(\vec{v}_1, \ldots, \vec{v}_n) = 0$ then $\vec{v}_1, \ldots, \vec{v}_n$ are linearly dependent. If they are not, they are a basis B for V. Then (3.1.6) says $A = A(\vec{v}_1, \ldots, \vec{v}_n)A_B$, and since $A \neq 0$, we must have $A(\vec{v}_1, \ldots, \vec{v}_n) \neq 0$, contrary to our hypothesis.

3.2 Determinants of linear operators

The fact that $\dim \Lambda^n V = 1$ makes possible a simple and elegant definition of the determinant $\det L$ of a linear operator $L \in \mathcal{L}(V \to V)$. Choose any nonzero $A \in \Lambda^n V$. Define $A^{[L]} \in F(\times^n V \to \mathcal{R})$ by

$$A^{[L]}\left(\vec{v}_1, \ldots, \vec{v}_n\right) = A\left[L(\vec{v}_1), \ldots, L(\vec{v}_n)\right], \quad \forall\left(\vec{v}_1, \ldots, \vec{v}_n\right) \in \times^n V. \tag{3.2.1}$$

Since L is linear and A is multilinear, definition (1.1.6) shows that $A^{[L]}$ is multilinear. It is obvious that $A^{[L]}$ is totally antisymmetric if A is. Thus $A^{[L]} \in \Lambda^n V$. Since $A \neq 0$, A is a basis for $\Lambda^n V$, and there is a scalar $k_{A,L}$ such that

$$A^{[L]} = k_{A,L}A. \tag{3.2.2}$$

From the definition (3.2.1), if $a \in \mathcal{R}$, $(aA)^{[L]} = aA^{[L]}$. Thus by (3.2.2), $(aA)^{[L]} = k_{A,L}aA$. But, applying the general result (3.2.2) to aA instead of A, we have $(aA)^{[L]} = k_{aA,L}(aA)$. Thus

$$k_{aA,L}(aA) = (aA)^{[L]} = aA^{[L]} = ak_{A,L}A = k_{A,L}(aA).$$

Since $A \neq 0$, $k_{aA,L} = k_{A,L}$. But every member of $\Lambda^n V$ can be written aA for some $a \in \mathcal{R}$. Therefore $k_{A,L}$ has the same value for all nonzero $A \in \Lambda^n V$. It is independent of A, and depends only on L. Thus for any $L \in \mathcal{L}(V \to V)$ there is a real number k_L such that for any nonzero $A \in \Lambda^n V$,

$$A^{[L]} = k_L A. \tag{3.2.3}$$

This equation is obviously true also for $A = 0$, and hence for *all* A in $\Lambda^n V$. The determinant of L, $\det L$, is *defined* to be k_L. Thus, by the definition of $\det L$,

$$A^{[L]} = (\det L)\, A \tag{3.2.4}$$

for any $A \in \Lambda^n V$. Returning to the definition (3.2.1) of $A^{[L]}$, we have

$$A\left[L\left(\vec{v}_1\right), \ldots, L\left(\vec{v}_n\right)\right] = (\det L)\, A\left(\vec{v}_1, \ldots, \vec{v}_n\right) \tag{3.2.5}$$

for any $A \in \Lambda^n V$ and any $(\vec{v}_1, \ldots, \vec{v}_n) \in \times^n V$.

Now suppose $(\vec{v}_1, \ldots, \vec{v}_n)$ is any ordered basis for V, and $L_i{}^j$ is the matrix of L relative to this basis. Then $L(\vec{v}_i) = L_i{}^j \vec{v}_j$ so

$$\begin{aligned}
A\left[L\left(\vec{v}_1\right), \ldots, L\left(\vec{v}_n\right)\right] &= A\left[L_1{}^{j_1} \vec{v}_{j_1}, \ldots, L_n{}^{j_n} \vec{v}_{jn}\right] \\
&= L_1{}^{j_1} \ldots L_n{}^{j_n} A\left(\vec{v}_1, \ldots, \vec{v}_n\right) \text{ because } A \text{ is multilinear} \\
&= L_1{}^{j_1} \ldots L_n{}^{j_n} \varepsilon_{j_1} \ldots j_n A\left(\vec{v}_{j_1}, \ldots, \vec{v}_{j_n}\right) \text{ because} \\
&\qquad A \text{ is totally antisymmetric (3.1.1)} \\
&= \left(\det L_i{}^j\right) A\left(\vec{v}_1, \ldots, \vec{v}_n\right)
\end{aligned}$$

where $\det L_i{}^j$ is the determinant of the $n \times n$ matrix $L_i{}^j$. Comparing this with (3.2.5), we see that

$$\det L = \det L_i{}^j. \tag{3.2.6}$$

The determinant of a linear operator is the determinant of its matrix relative to any basis. In deducing (3.2.6) from (3.2.5) we have used the fact that $A(\vec{v}_1, \ldots, \vec{v}_n) \neq 0$. This fact follows from remark (3.1.6).

Some properties of $\det L$ are easy to deduce from (3.2.5). For example, if K and L are both linear operators on V,

$$\det(K \circ L) = (\det K)(\det L). \tag{3.2.7}$$

The proof is simple. From (3.2.5),

$$\begin{aligned} A\left[(K \circ L)(\vec{v}_1), \ldots, (K \circ L)(v_n)\right] &= \det(K \circ L)\, A(\vec{v}_1, \ldots, \vec{v}_n) \\ &= A\left[K(L(\vec{v}_1)), \ldots, K(L(\vec{v}_n))\right] \\ &= (\det K)\, A\left[L(\vec{v}_1), \ldots, L(\vec{v}_n)\right] \\ &= (\det K)(\det L)\, A(\vec{v}_1, \ldots, \vec{v}_n). \end{aligned}$$

These equations are true for any $A \in \Lambda^n V$ and any $\{\vec{v}_1, \ldots, \vec{v}_n\} \subseteq V$. If we choose $A \neq 0$ and $(\vec{v}_1, \ldots, \vec{v}_n)$ an ordered basis, then $A(\vec{v}_1, \ldots, \vec{v}_n) \neq 0$, and we can cancel it, obtaining (3.2.7).

This choice of A and $(\vec{v}_1, \ldots, \vec{v}_n)$ also shows that if we set $L = I_V$ in (3.2.5) ($I_V :=$ identity mapping of V onto V) then

$$\det I_V = 1. \tag{3.2.8}$$

As another application of (3.2.5), let $A \neq 0$ and $\{\vec{v}_1, \ldots, \vec{v}_n\}$ be a basis for V. Then, as remarked in theorem (18.2), L is an isomorphism $\Longleftrightarrow \{L(\vec{v}_1), \ldots, L(\vec{v}_n)\}$ is linearly independent. From remark 3.1.20 this is true $\Longleftrightarrow A\left[L(\vec{v}_1), \ldots, L(\vec{v}_n)\right] \neq 0$. From (3.2.5), since $A(\vec{v}_1, \cdots \vec{v}_n) \neq 0$, $A[L(\vec{v}_1), \cdots, L(\vec{v}_r)] \neq 0 \Leftrightarrow \det L \neq 0$. Thus a linear operator is invertible, and hence an isomorphism, iff its determinant is nonzero.

If $\det L \neq 0$, then L^{-1} exists, and $L \circ L^{-1} = I_V$, so $(\det L)(\det L^{-1}) = \det(L \circ L^{-1}) = \det I_V = 1$. Thus

$$\det\left(L^{-1}\right) = (\det L)^{-1}. \tag{3.2.9}$$

Finally we claim

$$\det L^T = \det L. \tag{3.2.10}$$

Recall that L^T is the unique member of $\mathcal{L}(V \to V) \ni$

$$L(\vec{u}) \cdot \vec{w} = L^T(\vec{w}) \cdot \vec{u}, \quad \forall \vec{u}, \vec{w} \in V.$$

Let $(\hat{x}_1, \ldots, \hat{x}_n)$ be an ordered orthonormal basis for V. Relative to this basis B, the matrix of L^T is

$$\begin{aligned}
\left(L^T\right)_i{}^j &= c_B^j\left[L^T\left(\vec{b}_i\right)\right] = \hat{x}_j \cdot \left[L^T(\hat{x}_i)\right] = \hat{x}_i \cdot L(\hat{x}_j) \\
&= c_B^i\left[L(x_j)\right] = L_j{}^i, \qquad \text{and} \\
\det L^T &= \det\left(L^T\right)_i{}^j = \left(L^T\right)_1{}^{j_1} \ldots \left(L^T\right)_n{}^{j_n} \varepsilon_{j_1 \ldots j_n} \\
&= L_{j_1}{}^1 \ldots L_{j_n}{}^n \varepsilon_{j_1 \ldots j_n} \\
&= \sum_{\pi \in S_n} L_{\pi(1)}{}^1 \ldots L_{\pi(n)}{}^n \left(\operatorname{sgn} \pi\right).
\end{aligned}$$

Now for any $\pi \in S_n$,

$$L_i{}^j \in \left\{L_{\pi(1)}{}^1, \ldots, L_{\pi(n)}{}^n\right\}$$

$$\iff \quad i = \pi(j) \iff j = \pi^{-1}(i) \iff L_i{}^j \in \left\{L_1{}^{\pi^{-1}(1)}, \ldots, L_n{}^{\pi^{-1}(n)}\right\}.$$

Hence,

$$\left\{L_{\pi(1)}{}^1, \ldots, L_{\pi(n)}{}^n\right\} = \left\{L_1^{\pi^{-1}(1)}, \ldots, L_n^{\pi^{-1}(n)}\right\}$$

so,

$$L_{\pi(1)}{}^{1} \dots L_{\pi(n)}{}^{n} = L_1{}^{\pi^{-1}(1)} \dots L_n{}^{\pi^{-1}(n)}.$$

Moreover, $\operatorname{sgn} \pi = \operatorname{sgn} \pi^{-1}$. Hence

$$\det L^T = \sum_{\pi \in S_n} L_1{}^{\pi^{-1}(1)}, \dots, L_n{}^{\pi^{-1}(n)} \left(\operatorname{sgn} \pi^{-1} \right). \tag{3.2.11}$$

Now suppose $f : S_n \to W$ where W is any vector space. We claim

$$\sum_{\pi \in S_n} f\left(\pi^{-1}\right) = \sum_{\pi \in S_n} f(\pi). \tag{3.2.12}$$

The reason is that as π runs over S_n so does π^{-1}. The mapping $\pi \to \pi^{-1}$ is a bijection of S_n to itself. But from (3.2.11) and (3.2.12),

$$\begin{aligned} \det L^T &= \sum_{\pi \in S_n} L_1{}^{\pi(1)} \dots L_n{}^{\pi(n)} \operatorname{sgn} \pi \\ &= L_1{}^{i_1} \dots L_n{}^{i_n} \varepsilon_{i_1 \dots i_n} = \det L. \end{aligned}$$

If $L \in \Theta(V)$ (see page D-25) then $L^{-1} = L^T$ so $L \circ L^T = I_V$. Then $(\det L)(\det L^T) = \det I_V = 1$, so $(\det L)^2 = 1$. Thus

$$\det L = \pm 1 \quad \text{if} \quad V \in \Theta(V).$$

3.3 Orientation

Two ordered orthonormal bases for V, $(\hat{x}_1, \dots, \hat{x}_n)$ and $(\hat{x}'_1, \dots, \hat{x}'_n)$, are said to have the same orientation iff one basis can be obtained from the other by a continuous rigid motion. That is, there must be for each t in $0 \le t \le 1$ an ordered orthonormal basis $(\hat{x}_1(t), \dots, \hat{x}_n(t))$ with these properties:

i) $\hat{x}_i(0) = \hat{x}_i$

ii) $\hat{x}_i(1) = \hat{x}'_i$

iii) $\hat{x}_i(t)$ depends continuously on t.

Since $(\hat{x}_1(t), \ldots, \hat{x}_n(t)$ is orthonormal, there is an $L_t \in \Theta(V) \ni \hat{x}_i(t) = L_t(\hat{x}_i)$. From i) and iii) above,

i)′ $L_0 = I_V$

iii)′ L_t depends continuously on t (i.e., its components relative to any basis depend continously on t. For example, its components relative to $(\hat{x}_1, \ldots, \hat{x}_n)$ are $L_i{}^j = \hat{x}_j \cdot L_t(\hat{x}_i) = \hat{x}_j \cdot \hat{x}_i(t)$, and these are continuous in t.)

Then $\det L_0 = 1$ and $\det L_t$ depends continuously on t. But $L_t \in \Theta(V)$ for all t, so $\det L_t = \pm 1$ for all t. Hence $\det L_t = 1$ and $\det L_1 = 1$. Thus,

Remark 3.3.21 If two ordered *o.n.* bases for V have the same orientation, the orthogonal operators L which maps one basis onto the other has $\det L = 1$.

The converse of remark 3.3.21 is also true, as is clear from the comment on page D-26 and from Theorem PP-14. We will not prove that comment here. It is obvious when $\dim V = 2$, and for $\dim V = 3$ it is Euler's theorem.

If two ordered orthonormal bases do not have the same orientation, we say they have opposite orientations. The orthogonal operator L which maps one basis onto the other has $\det L = -1$. Given any ordered orthonormal basis $B = (\hat{x}_1, \ldots, \hat{x}_n)$, if two other ordered orthonormal bases are oriented oppositely to B, they have the same orientation. Proof:

> Let the bases be B' and B''. Let L' and L'' be the orthogonal operators which map B onto B' and B''. Then $L'' \circ (L')^{-1}$ is the orthogonal operator which maps B' onto B'', and $\det[L'' \circ (L')^{-1}] = (\det L'') \det((L')^{-1}) = (\det L'')(\det L')^{-1} = (-1)(-1) = +1$.

Thus, we can divide the ordered orthonormal bases for V into two oppositely oriented classes. Within each class, all bases have the same orientation. We call these two classes the "orientation classes" for V. Choosing an orientation for V amounts to choosing one of the two orientation classes for V.

It turns out that orientations are easily described in terms of alternating tensors.

Definition 3.3.15 We call $A \in \Lambda^n V$ "unimodular" if

$$|A(\hat{x}_1, \ldots, \hat{x}_n)| = 1 \tag{3.3.1}$$

for every ordered orthonormal basis $(\hat{x}_1, \ldots, \hat{x}_n)$ of V.

If A and A' are both unimodular, neither is 0, so there is a real nonzero a such that $A' = aA$. Applying (3.3.1) to both A and A' for one particular $(\hat{x}_1, \ldots, \hat{x}_n)$ shows $a = \pm 1$. Thus

$$A' = \pm A \tag{3.3.2}$$

if A and A' are unimodular. Hence

Remark 3.3.22 $\Lambda^n V$ contains at most two unimodular members. Does it contain any? To see that the answer is yes, we prove

Remark 3.3.23 Let $B = (\hat{x}_1, \ldots \hat{x}_n)$ and $B' = (\hat{x}'_1, \ldots \hat{x}'_n)$ be any two ordered orthonormal bases for V. Let A_B be defined by (3.1.1). Then $A_B(B')$ (an abbreviation for $A_B(\hat{x}'_1, \ldots, \hat{x}'_n)$) is $+1$ or -1 according as B and B' have this same or opposite orientations. Therefore, A_B and $-A_B$ are unimodular.

Proof:

> There is an $L \in \Theta(V)$ such that $\hat{x}'_i = L(\hat{x}_i)$. Then $A_B(\hat{x}'_1, \ldots, \hat{x}'_n) = A_B[L(\hat{x}_1), \ldots, L(\hat{x}_n)]$ $= (\det L) A_B(\hat{x}_1, \ldots, \hat{x}_n)$. But $\det L = \pm 1$, and $A_B(B) = +1$ by (3.1.3). QED.

If A and $-A$ are the two unimodular members of $\Lambda^n V$, we see from remark 3.3.23 that A is positive on one of the two orientation classes of V, and $-A$ is positive on the other orintation class. Choosing an orientation class for V amounts simply to choosing one of the two unimodular members of $\Lambda^n V$. Therefore we accept

Definition 3.3.16 An oriented Euclidean vector space is an ordered pair (V, A) in which V is Euclidean vector space and A is one of the two unimodular alternating tensors over V. A is called "the" alternating tensor of the oriented space. An ordered orthonormal basis $(\hat{x}_1, \ldots, \hat{x}_n)$ is "positively" or "negatively" oriented according as $A(\hat{x}_1 \ldots, \hat{x}_n)$ is $+1$ or -1.

Chapter 4

Tensor Products

4.1 Definition of a tensor product

Definition 4.1.17 Suppose $U_1, \ldots, U_p, V_1, \ldots, V_q$ are Euclidean vector spaces and $P \in U_1 \otimes \ldots \otimes U_p$ and $Q \in V_1 \otimes \ldots \otimes V_q$. Define PQ as the member of $\mathcal{F}(U_1 \times \ldots \times U_p \times V_1 \times \ldots \times V_q \to \mathcal{R})$ which assigns to any $(p+q)$-tuple $\vec{u}_1, \ldots, \vec{u}_p, \vec{v}_1, \ldots \vec{v}_q) \in U_1 \times \ldots \times U_p \times V_1 \times \ldots \times V_q$ the real number

$$PQ\left(\vec{u}_1, \ldots, \vec{u}_p, \vec{v}_1, \ldots, \vec{v}_q\right) = P\left(\vec{u}_1, \ldots, \vec{u}_p\right) Q\left(\vec{v}_1, \ldots, \vec{v}_q\right). \tag{4.1.1}$$

By definition (1.1.6) it is clear that PQ is multilinear, so

$$PQ \in U_1 \otimes \ldots \otimes U_p \otimes V_1 \otimes \ldots \otimes V_q. \tag{4.1.2}$$

PQ is called the "tensor product" of P and Q. Note that if P were in $\mathcal{M}(U_1 \times \ldots \times U_p \to W)$ and Q were in $\mathcal{M}(V_1 \times \ldots \times V_q \to W')$ with $W \neq \mathcal{R}$, $W' \neq \mathcal{R}$, then PQ could not be defined, because the product on the right of (4.1.1) would not be defined. Tensors can be multiplied by one another, but in general multilinear mappings cannot. The tensor product PQ is sometimes written $P \otimes Q$.

4.2 Properties of tensor products

Remark 4.2.24 Commutativity can fail. PQ need not equal QP. In fact, these two functions usually have different domains, namely $U_1 \times \ldots \times U_p \times V_1 \times \ldots \times V_q$ and $V_1 \times \ldots \times V_q \times U_1 \times \ldots \times U_p$. Even if their domains are the same, it is unusual to have $PQ = QP$. For example, suppose $p = q = 1$ and $U_1 = V_1 = V$. Then P and Q are vectors in V, say $P = \vec{u}$, $Q = \vec{v}$. For any $\vec{x}$ and $\vec{y} \in V$ we have $P(\vec{x}) = \vec{u} \cdot \vec{x}$, $Q(\vec{y}) = \vec{v} \cdot \vec{y}$ so $(PQ)(\vec{x}, \vec{y}) = P(\vec{x})\,Q(\vec{y}) = (\vec{u} \cdot \vec{x})(\vec{v} \cdot \vec{y})$. Similarly

$$(QP)(\vec{x}, \vec{y}) = Q(\vec{x})\,P(\vec{y}) = (\vec{v} \cdot \vec{x})(\vec{u} \cdot \vec{y}).$$

If $PQ = QP$, then $(PQ)(\vec{x}, \vec{y}) = (QP)(\vec{x}, \vec{y})$ for all $\vec{x}, \vec{y} \in V$, so

$$(\vec{u} \cdot \vec{x})(\vec{v} \cdot \vec{y}) = (\vec{v} \cdot \vec{x})(\vec{u} \cdot \vec{y}).$$

Then $\vec{x} \cdot [\vec{u}\,(\vec{v} \cdot \vec{y})] = \vec{x} \cdot [\vec{v}\,(\vec{u} \cdot \vec{y})]$ so

$$\vec{x} \cdot [\vec{u}\,(\vec{v} \cdot \vec{y}) - \vec{v}\,(\vec{u} \cdot \vec{y})] = 0. \tag{4.2.1}$$

If we fix $\vec{y}$ to be any particular vector in V, then (4.2.1) holds for all $\vec{x} \in V$. Therefore

$$\vec{u}\,(\vec{v} \cdot \vec{y}) - \vec{v}\,(\vec{u} \cdot \vec{y}) = \vec{0}. \tag{4.2.2}$$

This holds for every $\vec{y} \in V$, so it holds for $\vec{y} = \vec{u}$. If $\vec{u} \neq \vec{0}$ then $\vec{u} \cdot \vec{u} \neq 0$ so $\vec{u}(\vec{v} \cdot \vec{u}) - (\vec{u} \cdot \vec{u})\vec{v} = \vec{0}$ shows that $\vec{u}$ and $\vec{v}$ are linearly dependent. And if $\vec{u} = \vec{0}$, then of course $\vec{u}$ and $\vec{v}$ are also linearly dependent. Thus $PQ = QP$ implies $\vec{u}$ and $\vec{v}$ are linearly dependent. The converse is obvious.

Remark 4.2.25 There are no divisors of zero. That is if $P \in U_1 \otimes \cdots \otimes U_p$ and $Q \in V_1 \otimes \cdots \otimes V_p$ then

$$PQ = 0 \Longrightarrow P = 0 \text{ or } Q = 0. \tag{4.2.3}$$

To prove this, suppose $PQ = 0$ and $Q \neq 0$. We will show $P = 0$. Choose $(\vec{v}_1 \cdots, \vec{v}_q) \in V_1 \times \cdots \times V_q$ so that $Q(\vec{v}_1, \cdots, \vec{v}_q) \neq 0$ (by hypothesis $Q \neq 0$ so this is possible). Then for any $(\vec{u}_1, \cdots, \vec{u}_p) \in U_1 \times \cdots \times U_p$, $PQ = 0$ implies $P(\vec{u}_1, \cdots, \vec{u}_p)Q(\vec{v}_1, \cdots, \vec{v}_q) = 0$. Cancelling $Q(\vec{v}_1, \cdots, \vec{v}_q)$, which is $\neq 0$, gives $P(\vec{u}_1, \cdots, \vec{u}_p) = 0$. Since $\vec{u}_1, \cdots, \vec{u}_p$ were arbitrary, $P = 0$.

Corollary 4.2.13 $P_1 \cdots P_s = 0 \Longrightarrow$ one of $P_1, \cdots, P_s = 0$. (Induction on s gives proof).

Remark 4.2.26 Associative law. If $P \in U_1 \otimes \cdots \otimes U_p$, $Q \in V_1 \otimes \cdots \otimes V_q$, and $R \in W_1 \otimes \cdots \otimes W_r$ then

$$(PQ)\,R = P\,(QR)\,. \tag{4.2.4}$$

Hence, we can write PQR without ambiguity.

Proof:

For any $\tilde{u} = (\vec{u}_1, \cdots, \vec{u}_p) \in U_1 \times \cdots \times U_p$, any $\tilde{v} = (\vec{v}_1, \cdots, \vec{v}_q) \in V_1 \times \cdots \times V_q$, and any $\tilde{w} = (\vec{w}_1, \cdots, \vec{w}_r) \in W_1 \times \cdots \times W_r$, we have

$$\begin{aligned}
[(PQ)R]\,(\tilde{u}, \tilde{v}, \tilde{w}) &:= [(PQ)(\tilde{u}, \tilde{v})]\,R(\tilde{w}) \\
&:= P(\tilde{u})\,[(Q\tilde{v})R(\tilde{w})] \;\; \text{(rule of arithmetic in } \mathcal{R}) \\
&:= P(\tilde{u})\,[(QR)(\tilde{v}, \tilde{w})] \\
&:= [P(QR)]\,(\tilde{u}, \tilde{v}, \tilde{w}).
\end{aligned}$$

Remark 4.2.27 Right and left distributive laws. Suppose $a_1, \cdots, a_m \in \mathcal{R}$, $P_1, \cdots, P_m \in U_1 \otimes \cdots \otimes U_p$, and $Q \in V_1 \otimes \cdots \otimes V_q$. Then

$$(a_i P_i)Q = a_i(P_i Q) \text{ and } Q(a_i P_i) = a_i(QP_i). \tag{4.2.5}$$

Proof:

For any $\tilde{u} = (\vec{u}_1, \cdots, \vec{u}_p) \in U_1 \times \cdots \times U_p$ and any $\tilde{v} = (\vec{v}_1, \cdots, \vec{v}_q) \in V_1 \times \cdots \times V_q$ we have

$$\begin{aligned}
[(a_i P_i)Q]\,(\tilde{u}, \tilde{v}) &= [(a_i P_i)(\tilde{u})]\,Q(\tilde{v}) \\
&= \{a_i\,[P_i(\tilde{u})]\}\,Q(\tilde{v}) \\
&= a_i\,[P_i(\tilde{u})Q(\tilde{v})] \;\; \text{(real arithmetic)} \\
&= a_i\,[(P_i Q)(\tilde{u}, \tilde{v})] = [a_i(P_i Q)]\,(\tilde{u}, \tilde{v}).
\end{aligned}$$

The second of equations (4.2.5) is proved in the same way.

Remark 4.2.28 Multilinearity. For $i \in \{1, \cdots, s\}$, suppose $P_i \in U_1^{(i)} \otimes \cdots \otimes U_{p_i}^{(i)}$. Then $P_1 P_2 \cdots P_s$ is separately linear in each of $P_1, P_2, \cdots, P_s$.

Proof:

> To prove linearity in P_r, write $P = P_1 \cdots P_{r-1}$, $Q = P_r$, $R = P_{r+1} \cdots P_s$. (If $r = 1$, $P = 1$; if $r = s$, $R = 1$.) Then we want to show that PQR depends linearly on Q. We take $Q_1, \cdots, Q_m \in U_1^{(r)} \otimes \cdots \otimes U_{p_r}^{(r)}$ and $a_1, \cdots, a_m \in \mathcal{R}$. Then $P(a_i Q_i)R =$ (by 4.2.2) $P\left[(a_i Q_i)R\right]$ (by 4.2.5) $= P\left[a_i(Q_i R)\right] =$ (by 4.2.5) $= a_i\left[P(Q_i R)\right] =$ (by remark 4.2.2) $= a_i(PQ_i R)$.

4.3 Polyads

Definition 4.3.18 If $(\vec{v}_1, \cdots, \vec{v}_q) \in V_1 \times \cdots \times V_q$, where $V_1, \cdots, V_q$ are Euclidean spaces, then the tensor product $\vec{v}_1 \vec{v}_2 \cdots \vec{v}_q$ is called a polyad of order q. It is a member of $V_1 \otimes \cdots \otimes V_q$. It is a "dyad" if $q = 2$, a "triad" if $q = 3$, a "tetrad" if $q = 4$, a "pentad" if $q = 5$. By the definition 4.1.17 of a tensor product, for any $(\vec{x}_1, \cdots, \vec{x}_q) \in V_1 \times \cdots \times V_q$ we have

$$(\vec{v}_1 \cdots \vec{v}_q)(\vec{x}_1, \cdots, \vec{x}_q) = (\vec{v}_1 \cdot \vec{x}_1) \cdots (\vec{v}_q \cdot \vec{x}_q). \tag{4.3.1}$$

By remark (4.2.28), $\vec{v}_1 \vec{v}_2 \cdots \vec{v}_q$ depends linearly on each of $\vec{v}_1, \cdots, \vec{v}_q$ when the others are fixed. We define a mapping

a) $\quad \otimes : V_1 \times, \cdots, \times V_q \to V_1 \otimes \cdots \otimes V_q$

by requiring that for any $(\vec{v}_1, \cdots, \vec{v}_q) \in V_1 \times \cdots \times V_q$ we have

b) $\quad \otimes(\vec{v}_1, \cdots, \vec{v}_q) = \vec{v}_1 \vec{v}_2 \cdots \vec{v}_q.$

Then $\otimes$ is multilinear. That is

c) $\quad \otimes \in \mathcal{M}\left(V_1 \times \cdots \times V_q\right) \longrightarrow V_1 \otimes \cdots \otimes V_q\right). \qquad (4.3.2)$

The tensor product PQ is sometimes written $P \otimes Q$, and $\vec{v}_1 \vec{v}_2 \cdots \vec{v}_q$ is written $\vec{v}_1 \otimes \vec{v}_2 \otimes \cdots \otimes \vec{v}_q$. This notation makes (4.3.2b) look thus: $\otimes(\vec{v}_1, \cdots, \vec{v}_q) = \vec{v}_1 \otimes \cdots \otimes \vec{v}_q$. We will usually avoid these extra $\otimes$'s.

Remark 4.3.29 $\vec{u}_1\vec{u}_2\cdots\vec{u}_q = 0$ iff at least one of $\vec{u}_1, \vec{u}_2, \cdots, \vec{u}_q$ is 0.

Proof:

> $\Longleftarrow$ follows from remark 1.3.11 and the multilinearity of $\otimes$ in (4.3.2). $\Longrightarrow$ follows by induction on q from remark (4.2.25).

Remark 4.3.30 Suppose for each $p \in \{1, \cdots, q\}$ that $\vec{u}_p, \vec{v}_p \in V_p$ and $a_p \in \mathcal{R}$, and $\vec{v}_p = a_{(p)}\vec{u}_{(p)}$. Suppose also $a_1 \cdots a_q = 1$. Then $\vec{u}_1 \cdots \vec{u}_q = \vec{v}_1 \cdots \vec{v}_q$.

Proof:

> Since $\otimes$ is multilinear, $\vec{v}_1\vec{v}_2\cdots\vec{v}_q = (a_1\vec{u}_1)(a_2\vec{u}_2)\cdots(a_{(q)}\vec{u}_{(q)}) = \left[a_1a_2\cdots a_{(q)}\right]$ $\vec{u}_1\vec{u}_2\cdots\vec{u}_q = \vec{u}_1, \vec{u}_2\cdots\vec{u}_q$.

Remark 4.3.31 Suppose that for each $p \in \{1, \cdots, q\}$, we have $\vec{u}_p, \vec{v}_p \in V_p$. Suppose also that $\vec{u}_1 \cdots \vec{u}_q = \vec{v}_1 \cdots \vec{v}_q \neq 0$. Then $\exists a_1, \cdots, a_q \in \mathcal{R}$ such that $a_1a_2\cdots a_q = 1$ and for each $p \in \{1, \cdots, q\}$ we have $\vec{v}_p = a_{(p)}\vec{u}_{(p)}$.

Proof:

By remark 4.3.29, no $\vec{u}_p$ or $\vec{v}_p$ is $\vec{0}$. By hypothesis, for any $(\vec{x}_1, \cdots, \vec{x}_q) \in V_1 \times \cdots \times V_q$ we have

$$(\vec{u}_1 \cdot \vec{x}_1)\cdots\left(\vec{u}_{(q)} \cdot \vec{x}_{(q)}\right) = (\vec{v}_1 \cdot \vec{x}_1)\cdots\left(\vec{v}_{(q)} \cdot \vec{x}_{(q)}\right). \tag{4.3.3}$$

Choose any $p \in \{1, \cdots q\}$, and take $\vec{x}_r = \vec{v}_r$ if $r \neq p$. Then for any $\vec{x}_p \in V_p$ (4.3.3) implies

$$\vec{v}_{(p)} \cdot \vec{x}_{(p)} = \frac{(\vec{u}_1 \cdot \vec{v}_1)\cdots(\vec{u}_{(p-1)} \cdot \vec{v}_{(p-1)})(\vec{u}_{(p)} \cdot \vec{x}_{(p)})(\vec{u}_{(p+1)} \cdot \vec{v}_{(p+1)})\cdots(\vec{u}_{(q)} \cdot \vec{v}_{(q)})}{(\vec{v}_1 \cdot \vec{v}_1)\cdots(\vec{v}_{(p-1)} \cdot \vec{v}_{(p-1)})(\vec{v}_{(p+1)} \cdot \vec{v}_{(p+1)})\cdots(\vec{v}_{(q)} \cdot \vec{v}_{(q)})}$$

or

$$\vec{v}_{(p)} \cdot \vec{x}_{(p)} = a_{(p)}\vec{u}_{(p)} \cdot \vec{x}_{(p)} \tag{4.3.4}$$

where

$$a_{(p)} = \frac{(\vec{u}_1 \cdot \vec{v}_1)\cdots(\vec{u}_{(p-1)} \cdot \vec{v}_{(p-1)})(\vec{u}_{(p+1)} \cdot \vec{v}_{(p+1)})\cdots(\vec{u}_{(q)} \cdot \vec{v}_{(q)})}{(\vec{v}_1 \cdot \vec{v}_1)\cdots(\vec{v}_{(p-1)} \cdot \vec{v}_{(p-1)})(\vec{v}_{(p+1)} \cdot \vec{v}_{(p+1)})\cdots(\vec{v}_{(q)} \cdot \vec{v}_{(q)})}. \tag{4.3.5}$$

Since (4.3.4) is true for all $\vec{x}_p \in V_p$, it follows that

$$\vec{v}_p = a_{(p)}\vec{u}_{(p)}.$$

Then $\vec{v}_1 \cdots \vec{v}_q = (a_1 \cdots a_q)(\vec{u}_1 \cdots \vec{u}_q)$. But $\vec{v}_1 \cdots \vec{v}_q = \vec{u}_1 \cdots \vec{u}_q$ by hypothesis, so $(a_1 \cdots a_q - 1)(\vec{u}_1 \cdots \vec{u}_q) = 0$. Since $\vec{u}_1 \cdots \vec{u}_q \neq 0$, $a_1 \cdots a_q - 1 = 0$, or $a_1 \cdots a_q = 1$ (see facts about $\vec{0}$, p/s).

Chapter 5

Polyad Bases and Tensor Components

5.1 Polyad bases

It is true (see exercise 4) that some tensors of order ≥ 2 are not polyads. However, every tensor is a sum of polyads. We have

Theorem 5.1.15 For $p \in \{1, \ldots, q\}$, suppose V_p is a Euclidean space and $B_p = (\vec{b}_1^{(p)}, \cdots, \vec{b}_{n_p}^{(p)})$ is an ordered basis for V_p. Then the $n_1 n_2 \cdots n_q$ polyads $\vec{b}_{i_1}^{(1)} \vec{b}_{i_2}^{(2)} \cdots \vec{b}_{i_q}^{(q)}$ are a basis for $V_1 \otimes \cdots \otimes V_q$.

Proof:

First these polyads are linearly independent. Suppose that $S^{i_1 \cdots i_q}$ is an $n_1 \times n_2 \cdots \times n_q$ dimensional array of scalars such that

$$S^{i_1 \cdots i_q} \vec{b}_{i_1}^{(1)} \cdots \vec{b}_{i_q}^{(q)} = 0. \tag{5.1.1}$$

We want to prove $S^{i_1 \cdots i_q} = 0$. Equation (5.1.1) is equivalent to the assertion that for any $(\vec{x}_1, \cdots, \vec{x}_q) \in V_1 \times \cdots \times V_q$,

$$S^{i_1 \cdots i_q} \left(\vec{b}_{i_1}^{(1)} \cdot \vec{x}_1\right) \cdots \left(\vec{b}_{i_q}^{(q)} \cdot \vec{x}_q\right) = 0. \tag{5.1.2}$$

Let $B_p^D = (\vec{b}_{(p)}^1, \cdots, \vec{b}_{(p)}^{n_p})$ be the dual basis for B_p. Choose a particular q-tuple of integers $(j_1, \cdots, j_q)$ and set $\vec{x}_1 = \vec{b}_{(1)}^{j_1}, \cdots, \vec{x}_q = \vec{b}_{(q)}^{j_q}$ in (5.1.2). The result is

$$S^{i_1 \cdots i_q} \left(\vec{b}_{i_1}^{(1)} \cdot \vec{b}_{(1)}^{j_1}\right) = \left(\vec{b}_{i_q}^{(q)} \cdot \vec{b}_{(q)}^{j_q}\right) = 0.$$

But for dual bases, $(\vec{b}_i \cdot \vec{b}^j) = \delta_i{}^j$, so

$$S^{i_1 \cdots i_q} \delta_{i_1}{}^{j_1} = \delta_i{}^{j_q} = 0.$$

Hence, $S^{j_1 \cdots i_q} = 0$, as we hoped.

Second, the $n_1 n_2 \cdots n_q$ polyads $\vec{b}^{(1)}_{i_1} \cdots \vec{b}^{(q)}_{i_q}$ span $V_1 \otimes \cdots \otimes V_q$. For let $T \in V_1 \otimes \cdots \otimes V_q$. Define

$$T^{i_1 \cdots i_q} := T\left(\vec{b}^{i_1}_{(1)}, \cdots, \vec{b}^{i_q}_{(q)}\right) \tag{5.1.3}$$

and

$$\Pi := T^{i_1 \cdots i_q} \vec{b}^{(1)}_{i_1} \cdots \vec{b}^{(q)}_{i_q}. \tag{5.1.4}$$

Clearly Π is a linear combination of the polyads $\vec{b}^{(i)}_{i_1} \cdots \vec{b}^{(q)}_{i_q}$. We will show that $T = \Pi$. Since $(B_1^D, \cdots, B_q^D)$ is a basis sequence for $V_1 \times \cdots \times V_q$, by remark 1.2.10 it suffices to show that

$$\Pi\left(\vec{b}^{j_1}_{(1)}, \cdots, \vec{b}^{j_q}_{(q)}\right) = T\left(\vec{b}^{j_1}_{(1)}, \cdots, \vec{b}^{j_q}_{(q)}\right).$$

But

$$\begin{aligned}
\Pi\left(\vec{b}^{j_1}_{(1)}, \cdots, \vec{b}^{j_q}_{(q)}\right) &= T^{i_1 \cdots i_q}\left[\left(\vec{b}^{(1)}_{i_1} \cdots \vec{b}^{(q)}_{i_q}\right)\left(\vec{b}^{(j_1)}_{(1)}, \cdots, \vec{b}^{(q)}_{i_q}\right)\right] \\
&= T^{i_1 \cdots i_q}\left(\vec{b}^{(1)}_{i_1} \cdot \vec{b}^{j_1}_{(1)}\right) \cdots \left(\vec{b}^{(q)}_{i_q} \cdot \vec{b}^{j_q}_{(q)}\right) = T^{i_1 \cdots i_q} \delta_{i_1}{}^{j_1} \cdots \delta_{i_q}{}^{j_q} \\
&= T^{j_1 \cdots i_q} = T\left(\vec{b}^{j_1}_{(1)}, \cdots, \vec{b}^{j_q}_{(q)}\right).
\end{aligned}$$

QED.

Corollary 5.1.14 If $T \in V_1 \otimes \cdots \otimes V_q$ and $p \in \{1, \cdots, q\}$ then T is a sum of $n_1 \cdots n_q / n_p$ polyads, where $n_1 = \dim V_1, \cdots, n_q = \dim V_q$.

Proof:

From $T = \Pi$ and (5.1.4) we can write

$$T = \vec{b}^{(1)}_{i_1} \cdots \vec{b}^{(p-1)}_{i_{p-1}} \vec{v}^{i_1 \cdots i_{p-1} i_{p+1} \cdots i_q} \vec{b}^{(p+1)}_{i_{p+1}} \cdots \vec{b}^{(q)}_{i_q}$$

where

$$\vec{v}^{i_1 \cdots i_{p-1} i_{p+1} \cdots i_q}_{(p)} = T^{i_1 \cdots i_q} \vec{b}^{(p)}_{i_p} \in V_p.$$

Corollary 5.1.15 Suppose $V_1, \cdots, V_q$ are Euclidean spaces and W is a real vector space, not necessarily Euclidean. Suppose $L : V_1 \otimes \cdots \otimes V_q \to W$ and $M : V_1 \otimes \cdots \otimes V_q \to W$ are linear, and $L(T) = M(T)$ for every polyad T. Then $L = M$.

Proof:

For an arbitrary $T \in V_1 \otimes \cdots \otimes V_q$, write $T = \sum_{j=1}^N T_j$ where the T_j are polyads. Then $L(T) = \sum_{j=1}^N L(T_j) = \sum_{j=1}^N M(T_j) = M(T)$ because L and M are linear.

5.2 Components of a tensor relative to a basis sequence: Definition

Definition 5.2.19 For each $p \in \{1, \cdots, q\}$, suppose V_p is a Euclidean space and $B_p = (\vec{b}_1^{(p)}, \cdots, \vec{b}_{n_p}^{(p)})$ is an ordered basis for V_p, with dual basis $B_p^D = (\vec{b}_{(p)}^1, \cdots, \vec{b}_{(p)}^{n_p})$. Suppose $T \in V_1 \otimes \cdots \otimes V_q$. Then T has 2^q arrays of components relative to the basis sequence $(B_1, \cdots, B_q)$ for $V_1 \times \cdots \times V_q$. Each array has dimension $n_1 \times \cdots \times n_q$. The arrays are as follows:

$$\begin{cases} T^{i_1 \cdots i_q} = T\left(\vec{b}_{(1)}^{i_1}, \cdots, \vec{b}_{(q)}^{i_q}\right) \\ \\ T^{i_1 \cdots i_{q-1}}{}_{i_q} = T\left(\vec{b}_{(1)}{}^{i_1}, \cdots, \vec{b}_{(q-1)}{}^{i_{q-1}}, \vec{b}_{i_q}^{(q)}\right) \\ \\ \vdots \\ \\ T_{i_1 \cdots i_{q-1}}{}^{i_q} = T\left(\vec{b}_{i_1}^{(1)}, \cdots, \vec{b}_{i_{q-1}}{}^{q-1}, \vec{b}_{(q)}{}^{i_q}\right) \\ \\ T_{i_1 \cdots i_q} = T\left(\vec{b}_{i_1}^{(1)}, \cdots, \vec{b}_{i_q}^{(q)}\right) \end{cases} \tag{5.2.1}$$

An array is said to be "covariant" in its subscripts and "contravariant" in its superscripts. The first array in (5.2.1) is the array of "contravariant components of T relative to $(B_1, \cdots, B_q)$." The last array is the array of "covariant components relative to $(B_1, \cdots, B_q)$." The other $2^q - 2$ arrays are arrays of "mixed components relative to $(B_1, \cdots, B_q)$." Note that each array in the list (5.2.1) is the array of contravariant components of T relative to a suitable basis sequence. For example, $T^{i_1 \cdots i_{q-1}}{}_{iq}$ is the contravariant array relative to the basis sequence $(B_1, \cdots, B_{q-1}, B_q^D)$, and $T_{i_1 \cdots i_q}$ is the contravariant array relative to $(B_1^D, B_2^D, \cdots, B_q^D)$. Similarly, each array is the covariant array relative to some basis sequence.

From (5.1.3) and (5.1.4) and the fact that $T = \Pi$ in those equations, we conclude

$$T = T^{i_1 \cdots i_q} \vec{b}_{i_1}^{(1)} \vec{b}_{i_2}^{(2)} \cdots \vec{b}_{i_q}^{(q)}. \tag{5.2.2}$$

That is, if $T \in V_1 \otimes \cdots \otimes V_q$, then the coefficients required to express T as a linear combination of the basis polyads from the basis sequence $(B_1, \cdots, B_q)$ are precisely the contravariant components of T relative to that basis sequence. By considering all 2^q basis sequences, $(B_1, \cdots, B_q)$, $(B_1, \cdots, B_{q-1}, B_q^D), \cdots, (B_1^D, \cdots, B_q^D)$ we obtain from (5.2.2) the 2^q equations

$$\begin{cases} T = T^{i_1 \cdots i_q} \vec{b}_{i_1}^{(1)} \cdots \vec{b}_{i_q}^{(q)} \\ \\ T = T^{i_1 \cdots i_{q-1}}{}_{i_q} \vec{b}_{i_1}^{(1)} \cdots \vec{b}_{i_{q-1}}^{q-1} \vec{b}_{(q)}^{i_q} \\ \\ \vdots \\ \\ T = T_{i_1 \cdots i_{q-1}}{}^{i_q} = \vec{b}_{(1)}^{i_1} \cdots \vec{b}_{(q-1)}^{i_{q-1}}, \vec{b}_{i_q}^{(q)} \\ \\ T = T_{i_1 \cdots i_q} \vec{b}_{(1)}^{i_1} \cdots \vec{b}_{(q)}^{i_q}, \end{cases} \tag{5.2.3}$$

Note 5.2.1 Suppose $T = \tilde{T}^{i_1 \cdots i_q} \vec{b}_{i_1}^{(1)} \cdots \vec{b}_{i_q}^{(q)}$ for some array $\tilde{T}^{i_1 \cdots i_q}$. Since the polyads $\vec{b}_{i_1}^{(1)} \cdots \vec{b}_{i_q}^{(q)}$ are a basis for $V_1 \otimes \cdots \otimes V_q$, therefore $\tilde{T}^{i_1 \cdots i_q} = T^{i_1 \cdots i_q}$. By using others of the 2^q basis sequences one can reach the same conclusion for any of the 2^q equations (5.2.3).

Note 5.2.2 If S and $T \in V_1 \otimes \cdots \otimes V_q$ and there is one basis sequence for $V_1 \times \cdots \times V_q$ relative to which one of the 2^q component arrays of S is the same as the corresponding component array of T (e.g. $T_{i_1}{}^{i_2}{}_{i_3 \cdots i_q} = S_i{}^{i_2}\, i_3 \cdots i_q$) then $S = T$.

Note 5.2.3 The notation has been chosen to permit the restricted index conventions. A double index is summed only when it is a superscript at one appearance and a subscript at the other. The equation $a^i = b^i$ holds for all possible values of i, but $a^i = b_i$ should not occur unless we introduce orthonormal bases.

Note 5.2.4 If each of the bases $B_1, \cdots, B_q$ is orthonormal, then $B_p = B_p^D$, and all 2^q component arrays (5.2.3) are identical, and all 2^q equations (5.2.3) are the same. One uses the non-restricted index convention.

5.3 Changing bases

Suppose $T \in V_1 \otimes \cdots \otimes V_q$. If we know the array of contravariant components of T relative to one basis sequence $(B_1, \cdots, B_q)$, we know T. Therefore, we ought to be able to calculate the contravariant components of T relative to any other basis sequence $(\tilde{B}_1, \cdots, \tilde{B}_q)$. The procedure is this. Let

$$B_p = \left(\vec{b}_1^{(p)}, \cdots, \vec{b}_{n_p}^{(p)}\right), \quad \tilde{B}_p = \left(\tilde{\vec{b}}_1^{(p)}, \cdots, \tilde{\vec{b}}_{n_p}^{(p)}\right)$$

$$B_p^D \left(= \vec{b}^1_{(p)}, \cdots, \vec{b}^{n_p}_{(p)}\right), \quad \tilde{B}_p^D = \left(\tilde{\vec{b}}^1_{(p)}, \cdots, \tilde{\vec{b}}^{n_p}_{(p)}\right).$$

The array of contravariant components of T relative to $(\tilde{B}_1, \cdots, \tilde{B}_q)$ is

$$\begin{aligned}
\tilde{T}^{j_1 \cdots j_q} &= T\left(\tilde{\vec{b}}^{j_1}_{(1)}, \cdots, \tilde{\vec{b}}^{j_q}_{(q)}\right) \\
&= \left(T^{i_1 \cdots i_q} \vec{b}_{i_1}^{(1)} \cdots \vec{b}_{i_q}^{(q)}\right)\left(\tilde{\vec{b}}^{j_1}_{(1)}, \cdots, \tilde{\vec{b}}^{j_q}_{(q)}\right) \\
&= T^{i_1 \cdots i_q}\left[\left(\vec{b}_{i_1}^{(1)} \cdots \vec{b}_{i_q}^{(q)}\right)\left(\tilde{\vec{b}}^{j_1}_{(1)}, \cdots, \tilde{\vec{b}}^{j_q}_{(q)}\right)\right] \\
&= T^{i_1 \cdots i_q}\left(\vec{b}_{i_1}^{(1)} \cdot \vec{b}^{j_1}_{(1)}\right) \cdots \left(\tilde{\vec{b}}_{i_q}^{(q)} \cdot \tilde{\vec{b}}^{j_q}_{(q)}\right).
\end{aligned}$$

Thus

$$\tilde{T}^{j_1 \cdots j_q} = T^{i_1 \cdots i_q}\left(\vec{b}_{i_1}^{(1)} \cdot \tilde{\vec{b}}^{j_1}_{(1)}\right) \cdots \left(\vec{b}_{i_q}^{(q)} \cdot \tilde{\vec{b}}^{j_q}_{(q)}\right). \tag{5.3.1}$$

If we replace some of the bases in $(B_1, \cdots, B_q)$ and $(\tilde{B}_1, \cdots, \tilde{B}_q)$ by their duals, we can immediately obtain (5.3.1) with some of the j's as subscripts and some of the i's as subscripts on the components and superscript on the $\vec{b}$'s.

For example,

$$\begin{aligned}\tilde{T}_{j_1}{}^{j_2\cdots j_q} &= T^{i_1\cdots i_{q-1}}{}_{i_q}\left(\vec{b}^{(1)}_{i_1}\cdot\tilde{\vec{b}}^{(1)}_{j_1}\right)\left(\vec{b}^{(2)}_{i_2}\cdot\tilde{\vec{b}}^{j_2}_{(2)}\right)\cdots\\ &\cdots\left(\vec{b}^{(q-1)}_{i_{q-1}}\cdot\tilde{\vec{b}}^{j_{q-1}}_{(q-1)}\right)\left(\vec{b}^{i_q}_{(q)}\cdot\tilde{\vec{b}}^{j_q}_{(q)}\right)\end{aligned} \tag{5.3.2}$$

As an interesting special case of (5.3.1), we can take for $(\tilde{B}_1, \cdots, \tilde{B}_q)$ the basis sequence obtained from $(B_1, \cdots, B_q)$ by replacing some of the B_p by B_p^D. For example, if $\tilde{B}_1 = B_1^D$ and $\tilde{B}_p = B_p$ for $p \geq 2$, (5.3.1) becomes

$$T_{j_1}{}^{j_2\cdots j_q} = T^{i_1 j_2\cdots j_q} g^{(1)}_{i_1 j_1} \tag{5.3.3}$$

where $g^{(1)}_{ij} = \vec{b}^{(1)}_i \cdot \vec{b}^{(1)}_j$ is the covariant metric matrix of B_1. Similarly, if $(\tilde{B}_1, \cdots, \tilde{B}_q) = (B_1^D, \cdots, B_q^D)$, (5.3.1) becomes

$$T_{j_1\cdots j_q} = T^{i_1\cdots i_q} g^{(1)}_{i_1 j_1} \cdots g^{(q)}_{i_q j_q}. \tag{5.3.4}$$

Formulas like (5.3.3) and (5.3.4) (there are 4^q such) are said to raise or lower the indices of the component array of T.

5.4 Properties of component arrays

Remark 5.4.32 If $S, T \in V_1 \otimes \cdots \otimes V_q$ and $a, b \in \mathcal{R}$, then relative to any basis sequence for $V_1 \times \cdots \times V_q$,

$$(aS + bT)^{i_1\cdots i_q} = aS^{i_1\cdots i_q} + bT^{i_1\cdots i_q}.$$

Proof:

Immediate from definition of components, 5.2.19.

Remark 5.4.33 Suppose $P \in U_1 \otimes \cdots \otimes U_p$ and $Q \in V_1 \otimes \cdots \otimes V_q$. Then relative to any basis sequence for $U_1 \times \cdots \times U_p \times V_1 \times \cdots \times V_q$

$$(PQ)^{i_1 \cdots i_p j_1 \cdots j_q} = P^{i_1 \cdots i_p} Q^{j_1 \cdots j_q}.$$

Proof:

Immediate from definition of components, 5.2.19.

Remark 5.4.34 Suppose $T \in V_1 \otimes \cdots \otimes V_q$ and $\vec{v}^{(p)} \in V_p, p = 1, \cdots, q$. Taking components relative to any basis sequence gives

$$T\left(\vec{v}^{(1)}, \cdots, \vec{v}^{(q)}\right) = T^{i_1 \cdots i_q} v_{i_1}^{(1)} \cdots v_{i_q}^{(q)}.$$

Proof:

$\vec{v}^{(p)} = v_{i_p}^{(p)} \vec{b}_{(p)}^{i_p}$ so $T(\vec{v}^{(1)}, \cdots, \vec{v}^{(q)}) = T\left(v_{i_1}^{(1)} \vec{b}_{(1)}^{i_1}, \cdots, v_{i_q}^{(q)} \vec{b}_{(q)}^{i_q}\right)$
$= v_{i_1}^{(1)} \cdots v_{i_q}^{(q)} T\left(\vec{b}_{(1)}^{i_1}, \cdots, \vec{b}_{(q)}^{i_q}\right) = v_{i_1}^{(1)} \cdots v_{i_q}^{(q)} T^{i_1 \cdots i_q}$.

5.5 Symmetries of component arrays

If $V_1 = \cdots = V_q = V$, it is usual to consider only basis sequences $(B_1, \cdots, B_q)$ in which $B_1 = \cdots = B_q = B$. This is not necessary, but it does help if one wants to study the symmetries of $T \in V_1 \otimes \cdots \otimes V_q = \otimes^q V$ by looking at its arrays of components. If $B_1 = \cdots = B_q = B$, then one speaks of the arrays of components of T relative to the basis B.

If we do have $B_1 = \cdots = B_q = B = (\vec{b}_1, \cdots, \vec{b}_n)$, and if $T \in \otimes^q V$, then for any permutation $\pi \in S_q$ we have

$$\begin{aligned} (\pi T)^{i_1 \cdots i_q} &= (\pi T)\left(\vec{b}^{i_1}, \cdots, \vec{b}^{i_q}\right) = T\left(b^{i_{\pi(1)}}, \cdots, \vec{b}^{i_{\pi(q)}}\right) \\ &= T^{i_{\pi(1)} \cdots i_{\pi(q)}}. \end{aligned} \tag{5.5.1}$$

Here $B^D = (\vec{b}^1, \cdots, \vec{b}^n)$ is the basis dual to B. For mixed arrays of components, things are not quite so simple. For example, if (rs) is a transposition with $r < s$ then

$$\begin{aligned}
[(rs)T]^{i_1 \cdots i_r \cdots i_{s-1}\ i_s i_{s+1} \cdots i_q} &= [(rs)T]\left(\vec{b}^{i_1}, \cdots, \vec{b}^{i_{s-1}}, \vec{b}_{i_s}, \vec{b}^{i_{s+1}}, \cdots \vec{b}^{i_q}\right) \\
&= T\left(\vec{b}^{i_1}, \cdots, \vec{b}^{i_{r-1}}, \vec{b}_{i_s}, \vec{b}^{i_{r+1}}, \cdots, \vec{b}^{i_{s-1}}, \vec{b}^{i_r}, \vec{b}^{i_{s+1}}, \cdots \vec{b}^{i_q}\right) \\
&= T^{i_1 \cdots i_{r-1}}{}_{i_s}{}^{i_{r+1} \cdots i_{s-1} i_r i_{s+1} \cdots i_q}.
\end{aligned}$$

Whether $T \in \otimes^q V$ is symmetric or antisymmetric under (rs) can be tested by looking at one array of components of T relative to one basis, as long as the r and s indices in that array are both covariant or both contravariant. If one index is covariant and the other is contravariant, two different component arrays must be compared to test for symmetry. For example, if $T \in V \otimes V$ antisymmetric, then relative to any basis for V,

$$T^{ij} = -T^{ji}, \quad T^i{}_j = -T_j{}^i, \quad T_{ij} = -T_{ji}. \tag{5.5.2}$$

Proof:

$$T(\vec{b}^i, \vec{b}^j) = -T(\vec{b}^j, \vec{b}^i)\ ,\ T(\vec{b}^i, \vec{b}_j) = -T(\vec{b}_j, \vec{b}^i)\ ,\ T(\vec{b}_j, \vec{b}_i) = -T(\vec{b}_j, \vec{b}_i)\ .$$

Moreover, if any one of the three equations (5.5.2) is true relative to one basis B for V, then T is antisymmetric. For example, if $T^i{}_j = -T_j{}^i$ then T and $-(12)T$ take the same values on the basis sequence (B^D, B) for $V \times V$. Therefore, by remark 1.2.10, $T = -(12)T$.

5.6 Examples of component arrays

Example 5.6.3 Let $\mathcal{I}$ be the identity tensor on Euclidean vector space V. By definition, $\mathcal{I}$ is that member of $V \otimes V$ such that for any $\vec{u}, \vec{v} \in V$ we have

$$\mathcal{I}(\vec{u}, \vec{v}) = \vec{u} \cdot \vec{v}.$$

Let $B = (\vec{b}_1, \cdots, \vec{b}_n)$ be any ordered basis for V, with dual basis $B^D = (\vec{b}^1, \cdots, \vec{b}^n)$. Let $g_{ij} = \vec{b}_i \cdot \vec{b}_j$ and $g^{ij} = \vec{b}^i \cdot \vec{b}^j$ be the covariant and contravariant metric matrices relative to

B. Then

$$\begin{aligned}
\mathcal{I}^{ij} &= \mathcal{I}\left(\vec{b}^i, \vec{b}^j\right) = \vec{b}^i \cdot \vec{b}^j = g^{ij} \\
\mathcal{I}^i{}_j &= \mathcal{I}\left(\vec{b}^i, \vec{b}^j\right) = \vec{b}^i \cdot \vec{b}_j = \delta^i{}_j \\
\mathcal{I}_i{}^j &= \mathcal{I}\left(\vec{b}_i, \vec{b}^j\right) = \vec{b}_i \cdot \vec{b}^j = \delta_i{}^j \\
\mathcal{I}_{ij} &= \mathcal{I}\left(\vec{b}_i, \vec{b}_j\right) g_{ij}.
\end{aligned}$$

(The reason for calling $\mathcal{I}$ the "identity tensor" will appear later.) Relative to an orthonormal basis, any component array of $\mathcal{I}$ is δ_{ij}.

From (5.2.3) it follows that

$$\mathcal{I} = g^{ij}\vec{b}_i\vec{b}_j = \delta^i{}_j\vec{b}_i\vec{b}^j = \delta_i{}^j\vec{b}^i\vec{b}_j = g_{ij}\vec{b}^i\vec{b}^j.$$

Thus,

$$\mathcal{I} = g^{ij}\vec{b}_i\vec{b}_j = \vec{b}_i\vec{b}^i = \vec{b}^i\vec{b}_i = g_{ij}\vec{b}^i\vec{b}^j. \tag{5.6.1}$$

In particular, if $(\hat{x}_1, \cdots, \hat{x}_n)$ is an orthonormal basis for V,

$$\mathcal{I} = \hat{x}_i\hat{x}_i \tag{5.6.2}$$

Example 5.6.4 Let V, B, B^D be as in example 5.6.3. Let A be any alternating tensor over V (i.e. any member of $\Lambda^n V$). Then from (3.1.2)

$$A^{i_1\cdots i_n} = A\left(\vec{b}^{i_1}, \cdots, \vec{b}^{i_n}\right) = \varepsilon^{i_1\cdots i_n} A\left(\vec{b}^1, \cdots, \vec{b}^n\right). \tag{5.6.3}$$

$$A_{i_1\cdots i_n} = A\left(\vec{b}_{i_1}, \cdots \vec{b}_{i_n}\right) = \varepsilon_{i_1\cdots i_n} A\left(\vec{b}_1, \cdots \vec{b}_n\right). \tag{5.6.4}$$

It is *not* true that

$$A_{i_1}{}^{i_2\cdots i_n} = \varepsilon_{i_1}{}^{i_2\cdots i_n} A\left(\vec{b}_1, \vec{b}^2, \cdots, \vec{b}^n\right).$$

If (V, A) is an oriented Euclidean vector space, then A is unimodular. If $B = (\hat{x}_1, \cdots, \hat{x}_n)$ is an orthonormal basis for V then from (5.6.4)

$$A_{i_1\cdots i_n} = \varepsilon_{i_1\cdots i_n} \text{ if } (\hat{x}_1, \cdots, \hat{x}_n) \text{ is positively oriented.}$$

$$A_{i_1\cdots i_n} = -\varepsilon_{i_1\cdots i_n} \text{ if } (\hat{x}_1, \cdots, \hat{x}_n) \text{ is negatively oriented.}$$

Chapter 6

The Lifting Theorem

We must learn to perform a number of simple but useful operations on tensors of order q. Most of these operations will be easy to perform on polyads, so we express $T \in V_1 \otimes \cdots \otimes V_q$ as a sum of polyads and perform the operations on the individual polyads, adding the results. The procedure works quite well once we have overcome a difficulty which is best understood by considering an example.

Suppose $1 \leq r < s \leq q$ and $V_1, \cdots, V_q$ are Euclidean vector spaces and $V_r = V_s$. For any $T \in V_1 \otimes \cdots \otimes V_q$, we want to define the "trace on indices r and s", written $tr_{rs}T$. First, suppose P is a polyad in $V_1 \otimes \cdots \otimes V_q$, that is

$$P = \vec{u}_1 \cdots \vec{u}_r \cdots \vec{u}_s \cdots \vec{u}_q. \tag{6.0.1}$$

Then we define

$$\begin{aligned} tr_{rs}P &:= (\vec{u}_r \cdot \vec{u}_s)\, \vec{u}_1 \cdots \cancel{\vec{u}_r} \cdots \cancel{\vec{u}_s} \cdots \vec{u}_q \\ &:= (\vec{u}_r \cdot \vec{u}_s)\, \vec{u}_1 \cdots \vec{u}_{r-1}\vec{u}_{r+1} \cdots \vec{u}_{s-1}\vec{u}_{s+1} \cdots \vec{u}_q. \end{aligned} \tag{6.0.2}$$

If $T \in V_1 \otimes \cdots \otimes V_q$, we write

$$T = \sum_{i=1}^{m} P_i, \quad P_i \text{ a polyad} \tag{6.0.3}$$

and we define

$$tr_{rs}T = \sum_{i=1}^{m} tr_{rs}P_i. \tag{6.0.4}$$

There are two serious objections to this apparently reasonable procedure:

(i) Suppose $P = \vec{u}_1 \cdots \vec{u}_r \cdots \vec{u}_s \cdots \vec{u}_q = \vec{v}_1 \cdots \vec{v}_r \cdots \vec{v}_s \cdots \vec{v}_q$. Is it true that $(\vec{u}_r \cdot \vec{u}_s)\vec{u}_1 \cdots \cancel{\vec{u}_r} \cdots \cancel{\vec{u}_s} \cdots \vec{u}_q = (\vec{v}_r \cdot \vec{v}_s)\vec{v}_1 \cdots \cancel{\vec{v}_r} \cdots \cancel{\vec{v}_s} \cdots \vec{v}_q$? If not, $tr_{rs}P$ is not uniquely defined by (6.0.2).

(ii) Suppose $T = \sum_{i=1}^m P_i = \sum_{j=1}^n Q_j$ where P_i and Q_j are polyads. Is it true that $\sum_{i=1}^m tr_{rs}P_i = \sum_{j=1}^n tr_{rs}Q_j$? If not, $tr_{rs}T$ is not uniquely defined by (6.0.3).

To attack this problem, we return to the very first sloppiness, in (6.0.2). We hope that $tr_{rs}P$ will be a unique polyad in $V_1 \otimes \cdots \otimes \cancel{V_r} \otimes \cdots \otimes \cancel{V_s} \otimes \cdots \otimes V_q$, so that tr_{rs} maps the set of polyads in $V_1 \otimes \cdots \otimes V_q$ into the set of polyads in $V_1 \otimes \cdots \otimes \cancel{V_r} \otimes \cdots \otimes \cancel{V_s} \otimes \cdots \otimes V_q$. But difficulty i) above leads us to look carefully at (6.0.1) and (6.0.2), and to recognize that until we have provided some theory all we have really done in (6.0.2) is to show how to take an ordered q-tuple $(\vec{u}_1, \cdots, \vec{u}_q)$ and assign to it a polyad of order $q-2$ in $V_1 \otimes \cdots \otimes \cancel{V_r} \otimes \cdots \otimes \cancel{V_s} \otimes \cdots \otimes V_q$. The polyad assigned to $(\vec{u}_1, \cdots, \vec{u}_q)$ is

$$M(\vec{u}_1, \cdots, \vec{u}_q) = (\vec{u}_r \cdot \vec{u}_s)\,\vec{u}_1 \cdots \cancel{\vec{u}_r} \cdots \cancel{\vec{u}_s} \cdots \vec{u}_q. \tag{6.0.5}$$

This equation does clearly and unambiguously define a function $M \in \mathcal{F}(V_1 \times \cdots \times V_q \to V_1 \otimes \cdots \otimes \cancel{V_r} \otimes \cdots \otimes \cancel{V_s} \otimes \cdots \otimes V_q)$. In fact, from definition (1.1.6), property d' on page D-18 for dot products, and remark 4.2.28 for tensor products, it is clear that M is multilinear. Hence,

$$M \in \mathcal{M}(V_1 \times \cdots \times V_q \to V_1 \otimes \cdots \otimes \cancel{V_r} \otimes \cdots \otimes \cancel{V_s} \otimes \cdots \otimes V_q). \tag{6.0.6}$$

This M is all we really have. The process of constructing a linear mapping $tr_{rs} : V_1 \otimes \cdots \otimes V_q \to V_1 \otimes \cdots \otimes \cancel{V_r} \otimes \cdots \otimes \cancel{V_s} \otimes \cdots \otimes V_q$ from the M of (6.0.5) is a very general one. That such a construction is possible and unique is the intent of the "lifting theorem." This theorem holds for all tensors, not just those over Euclidean spaces, and it is sometimes (as in Marcus, Multilinear Algebra) taken as the definition of $V_1 \otimes \cdots \otimes V_q$.

A formal statement of the lifting theorem for Euclidean vector spaces is as follows:

Theorem 6.0.16 (Lifting theorem). Let $V_1, \cdots, V_q$ be Euclidean vector spaces and let W be any real vector space. Let $M \in \mathcal{M}(V_1 \times \cdots \times V_q \to W)$. Then there is exactly one

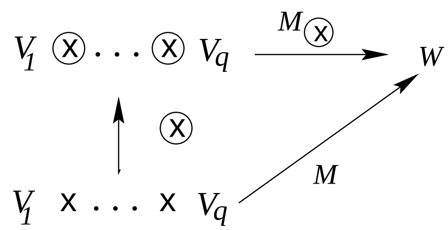


Figure 6.1:

$M_\otimes \in \mathcal{L}(V_1 \otimes \cdots \otimes V_q \to W)$ such that

$$M = M_\otimes \circ \otimes \tag{6.0.7}$$

where $\otimes \in \mathcal{M}(V_1 \times \cdots \times V_q \to V_1 \otimes \cdots \otimes V_q)$ is defined by $\otimes(\vec{v}_1, \cdots, \vec{v}_q) = \vec{v}_1 \cdots \vec{v}_q$.

Before proving the theorem, we discuss it and try to clarify its meaning. To make (6.0.7) less abstract we note that it is true $\Longleftrightarrow$

$$M\left(\vec{v}_1, \cdots, \vec{v}_q\right) = M_\otimes\left[\otimes\left(\vec{v}_1, \cdots, \vec{v}_q\right)\right], \quad \forall\left(\vec{v}_1, \cdots, \vec{v}_q\right) \in V_1 \times \cdots \times V_q.$$

This is equivalent to

$$M\left(\vec{v}_1, \cdots, \vec{v}_q\right) = M_\otimes\left(\vec{v}_1 \cdots \vec{v}_q\right). \tag{6.0.8}$$

In other words, the lifting theorem asserts the existence of exactly one linear mapping $M_\otimes : V_1 \otimes \cdots \otimes V_q \to W$ such that when $M_\otimes$ is applied to any polyad the result is the same as applying M to any q-tuple of vectors whose tensor product is that polyad.

A diagram of the three functions may clarify matters: M maps $V_1 \times \cdots \times V_q$ into W, and $\otimes$ maps $V_1 \times \cdots \times V_q$ into $V_1 \otimes \cdots \otimes V_q$. Both mappings are multilinear. The lifting theorem "lifts" M from $V_1 \times \cdots \times V_q$ to $V_1 \otimes \cdots \otimes V_q$ by producing a unique linear mapping $M_\otimes : V_1 \otimes \cdots \otimes V_q \to W$ which satisfies (6.0.7) or (6.0.8). These equations mean that if we start at any $(\vec{v}_1, \cdots, \vec{v}_q)$ in $(\vec{V}_1 \times \cdots \times \vec{V}_q)$, and go to W either via M or via $\otimes$ and $M_\otimes$, we will reach the same vector in W. Usually, when people draw the diagram as in Figure 6.2 they mean $f : U \to W$, $g : U \to V$, $h : V \to W$, and $f = h \circ g$.

Proof of the lifting theorem:

The proof is actually rather simple. Choose a basis sequence $(B_1, \cdots, B_q)$ for $V_1 \times \cdots \times V_q$, with $B_p = (\vec{b}_1^{(p)}, \cdots, \vec{b}_{n_p}^{(p)})$. Let $B_p^D = (\vec{b}^1_{(p)}, \cdots, \vec{b}^{n_p}_{(p)})$ be the basis for V_p dual to B_p. First

$$\begin{tikzcd} V \arrow[r, "h"] & W \\ U \arrow[u, "g"] \arrow[ur, "f"'] & \end{tikzcd}$$

Figure 6.2:

we will prove that $M_\otimes$ is unique, i.e. there can be no more than one $M_\otimes \in \mathcal{L}(V_1 \otimes \cdots \otimes V_q \to W)$ which satisfies (6.0.7) or (6.0.8). Suppose $M_\otimes$ is such a mapping. Let $T \in V_1 \otimes \cdots \otimes V_q$. By (5.2.2), $T = T^{i_1 \cdots i_q} \vec{b}^{(1)}_{i_1} \cdots \vec{b}^{(q)}_{i_q}$ so $M_\otimes(T) = T^{i_1 \cdots i_q} M_\otimes(\vec{b}^{(1)}_{i_1} \cdots \vec{b}^{(q)}_{i_q})$ because $M_\otimes$ is linear. Then $M_\otimes(T) = T^{i_1 \cdots i_q} M(\vec{b}^{(1)}_{i_1}, \cdots, \vec{b}^{(q)}_{i_q})$ from (6.0.8) with $\vec{v}_1 = \vec{b}^{(1)}_{i_1}, \cdots, \vec{v}_q = \vec{b}^{(q)}_{iq}$. Thus if such an $M_\otimes$ exists, then for any $T \in V_1 \otimes \cdots \otimes V_q$ we must have

$$M_\otimes(T) = T\left(\vec{b}^{i_1}_{(1)}, \cdots, \vec{b}^{i_q}_{(q)}\right) M\left(\vec{b}^{(1)}_{i_1}, \cdots, \vec{b}^{(q)}_{i_q}\right). \tag{6.0.9}$$

In other words, $M_\otimes$ is determined because $M_\otimes(T)$ is known for all $T \in V_1 \otimes \cdots \otimes V_q$.

It still remains to prove that there is an $M_\otimes \in \mathcal{L}(V_1 \otimes \cdots \otimes V_q \to W)$ which satisfies (6.0.7) and (6.0.8). Our uniqueness proof has given us an obvious candidate, namely the $M_\otimes : V_1 \otimes \cdots \otimes V_q \to W$ defined by (6.0.9). This $M_\otimes$ is certainly a well defined function mapping $V_1 \otimes \cdots \otimes V_q$ into W. By the definitions of ad and sc in $V_1 \otimes \cdots \otimes V_q$, $M_\otimes(T)$ depends linearly on T, i.e. $M_\otimes \in \mathcal{L}(V_1 \otimes \cdots \otimes V_q \to W)$. [For a formal proof, let $T_1, \cdots, T_n \in V_1 \otimes \cdots \otimes V_q$ and $a^1, \cdots, a^n \in \mathcal{R}$. Then we want to show $M_\otimes(a^j T_j) = a^j M_\otimes(T_j)$. From (6.0.9), and the definition of $a^j T_j$, and the rules of vector arithmetic in W, given on page D-11, we compute

$$\begin{aligned}
M_\otimes\left(a^j T_j\right) &= \left(a^j T_j\right)\left(\vec{b}^{i_1}_{(1)}, \cdots, \vec{b}^{i_q}_{(q)}\right) M\left(\vec{b}^{(1)}_{i_1}, \cdots, \vec{b}^{(q)}_{i_q}\right) \\
&= \left\{a^j \left[T_j\left(\vec{b}^{i_1}_{(1)}, \cdots, \vec{b}^{i_q}_{(q)}\right)\right]\right\} M\left(\vec{b}^{(1)}_{i_1}, \cdots, \vec{b}^{(q)}_{i_q}\right) \\
&= a^j \left[T_j\left(\vec{b}^{i_1}_{(1)}, \cdots, \vec{b}^{i_q}_{(q)}\right) M\left(\vec{b}^{(1)}_{i_1}, \cdots, \vec{b}^{(q)}_{i_q}\right)\right] \\
&= a^j M_\otimes(T_j).
\end{aligned}$$

QED.]

It remains only to prove that $M_\otimes$ satisfies (6.0.8). Let $(\vec{v}_1, \cdots, \vec{v}_q) \in V_1 \times \cdots \times V_q$. Then if $T = \vec{v}_1 \cdots \vec{v}_q$ we have $T\left(\vec{b}^{i_1}_{(1)}, \cdots, \vec{b}^{i_q}_{(q)}\right) = (\vec{v}_1 \cdot \vec{b}^{i_1}_{(1)}) \cdots (\vec{v}_q \cdot \vec{b}^{i_q}_{(q)}) = v_1^{i_1} \cdots v_{(q)}^{i_q}$, so $M_\otimes(\vec{v}_1 \cdots \vec{v}_q) = v_1^{i_1} \cdots v_{(q)}^{i_q} M(\vec{b}^{(1)}_{i_1}, \cdots, \vec{b}^{(q)}_{i_q})$. By hypothesis, M is multilinear, so $M_\otimes(\vec{v}_1 \cdots \vec{v}_q) = M(v_1^{i_1}\vec{b}^{(1)}_{i_1}, \cdots, v_q^{iq}\vec{b}^{(q)}_{i_q})$. But this is just (6.0.8). QED.

As an application of the lifting theorem, let us return to $tr_{rs}T$. The M defined in (6.0.5) will be the M of the lifting theorem. The vector space W in that theorem will be $W = V_1 \otimes \cdots \otimes \not{V}_r \otimes \cdots \otimes \not{V}_s \otimes \cdots \otimes V_q$. Then, $M_\otimes \in \mathcal{L}(V_1 \otimes \cdots \otimes V_q \to V_1 \otimes \cdots \otimes \not{V}_r \otimes \cdots \otimes \not{V}_s \otimes \cdots \otimes V_q)$ is given to us by the lifting theorem, and we define for any $T \in V_1 \otimes \cdots \otimes V_q$

$$tr_{rs}T := M_\otimes(T). \tag{6.0.10}$$

If a T is a polyad, $T = \vec{v}_1 \cdots \vec{v}_q$, then

$$\begin{aligned} tr_{rs}\left(\vec{v}_1 \cdots \vec{v}_q\right) &= M_\otimes\left(\vec{v}_1 \cdots \vec{v}_q\right) = M\left(\vec{v}_1, \cdots, \vec{v}_q\right) \\ &= \left(\vec{v}_r \cdot \vec{v}_s\right) \vec{v}_1 \cdots \not{\vec{v}}_r \cdots \not{\vec{v}}_s \cdots \vec{v}_q. \end{aligned} \tag{6.0.11}$$

Thus the lifting theorem answers both the objections on page 41.

We agree to define

$$tr_{sr}T := tr_{rs}T. \tag{6.0.12}$$

As an application of this "machinery," let us find the component arrays of $tr_{12}T$ relative to the basis sequence $(B_3, \cdots, B_q)$ for $V_3 \times \cdots \times V_q$. We have

$$T = T_{i_1}{}^{i_2 i_3 \cdots i_q} \vec{b}^{i_1}_{(1)} \vec{b}^{(2)}_{i_2} \vec{b}^{(3)}_{i_3} \cdots \vec{b}^{(q)}_{i_q}.$$

We must suppose $V_1 = V_2$ if $tr_{12}T$ is to be defined. Then we take $B_1 = B_2$, so $\vec{b}^i_{(1)} = \vec{b}^i_{(2)}$. Since tr_{12} is a linear mapping,

$$tr_{12}T = T_{i_1}{}^{i_2 \cdots i_q} tr_{12}\left[\vec{b}^{i_1}_{(1)} \vec{b}^{(2)}_{i_2} \vec{b}^{(3)}_{i_3} \cdots \vec{b}^{(q)}_{i_q}\right]. \tag{6.0.13}$$

By (6.0.12),

$$tr_{12}\left[\vec{b}^{i_1}_{(1)} \vec{b}^{(2)}_{i_2} \vec{b}^{(3)}_{i_3} \cdots \vec{b}^{(q)}_{i_q}\right] = \left(\vec{b}^{i_1}_{(1)} \cdot \vec{b}^{(2)}_{i_2}\right) \vec{b}^{(3)}_{i_3} \cdots \vec{b}^{(q)}_{i_q}.$$

We have agreed to take $B_1 = B_2$, so $\vec{b}^{i_1}_{(1)} \cdot \vec{b}^{(2)}_{i_2} = \delta^{i_1}{}_{i_2}$. Thus,

$$tr_{12}\left[\vec{b}^{i_1}_{(1)} \vec{b}^{(2)}_{i_2} \vec{b}^{(3)}_{i_3} \cdots \vec{b}^{(q)}_{i_q}\right] = \left(\delta^{i_1}{}_{i_2} \vec{b}^{(3)}_{i_3} \cdots \vec{b}^{(q)}_{i_q}\right). \tag{6.0.14}$$

Then, substituting this in (6.0.13),

$$tr_{12}T = T_j{}^{ji_3\cdots i_q}\vec{b}^{(3)}_{i_3}\cdots\vec{b}^{(q)}_{i_q},$$

so

$$(tr_{12}T)^{i_3\cdots i_q} = T_j{}^{ji_3\cdots i_q}. \tag{6.0.15}$$

Similarly

$$(tr_{12}T)^{i_3\cdots i_q} = T^j{}_j{}^{i_3\cdots i_q} \tag{6.0.16}$$

is proved from

$$T = T^{i_1}{}_{i_2}{}^{i_3\cdots i_q}\vec{b}^{(1)}_{i_1}\vec{b}^{i_2}_{(2)}\vec{b}^{(3)}_{i_3}\cdots\vec{b}^{(q)}_{i_q}.$$

Recall that for orthonormal bases we need not distinguish between superscripts and subscripts. Thus, if all bases are *orthonormal* and $B_1 = B_2$,

$$(tr_{12}T)_{i_3\cdots i_q} = T_{jji_3\cdots i_q}. \tag{6.0.17}$$

Chapter 7

Generalized Dot Products

7.1 Motivation and definition

Suppose U, V, and W are Euclidean vector spaces, and $\vec{u} \in U$, $\vec{x} \in V$, $\vec{y} \in V$, $\vec{w} \in W$. Then $\vec{u}\vec{x} \in U \otimes V$ and $\vec{y}\vec{w} \in V \otimes W$. What could be more natural than to define (following Willard Gibbs)

$$(\vec{u}\vec{x}) \cdot (\vec{y}\vec{w}) := \vec{u}\,(\vec{x} \cdot \vec{y})\,\vec{w}, \tag{7.1.1}$$

which is the same as $(\vec{u}\vec{x}) \cdot (\vec{y}\vec{w}) := (\vec{x} \cdot \vec{y})\,\vec{u}\vec{w}$? Then the definition can be extended so that if $P \in U \otimes V$ and $R \in V \otimes W$ we define $P \cdot R$ by writing

$$P = \sum_{i=1}^{m} P_i, \qquad R = \sum_{j=1}^{n} R_j \tag{7.1.2}$$

where P_i and R_j are dyads; then we define

$$P \cdot R := \sum_{i=1}^{m} \sum_{j=1}^{n} P_i \cdot R_j \tag{7.1.3}$$

where $P_i \cdot R_j$ is defined by (7.1.1). It is in $U \otimes W$, as is $P \cdot R$.

Does all this work? If $\vec{u}\vec{x} = \vec{u}'\vec{x}'$ and $\vec{y}\vec{w} = \vec{y}'\vec{w}'$, is it true that $(\vec{x} \cdot \vec{y})\,\vec{u}\vec{w} = (\vec{x}' \cdot \vec{y}')\,\vec{u}'\vec{w}'$? And if $P = \sum_{i=1}^{m} P_i = \sum_{i=1}^{m'} P_i'$ and $R = \sum_{j=1}^{n} R_j = \sum_{j=1}^{n'} R_j'$, is it true that $\sum_{i=1}^{m} \sum_{j=1}^{n} P_i \cdot R_j = \sum_{i=1}^{m'} \sum_{j=1}^{n'} P_i' \cdot R_j'$?

The lifting theorem will show that all of this works, and that it can even be generalized as follows: Suppose $U_1, \cdots, U_p$, $V_1, \cdots, V_q$, $W_1, \cdots, W_r$ are Euclidean spaces, and $\vec{u}_i \in U_i$,

$\vec{x}_j$ and $\vec{y}_j \in V_j$, $\vec{w}_r \in W_r$. We would like to define a generalized dot product of order q as

$$\begin{aligned}(\vec{u}_1 \cdots \vec{u}_p \vec{x}_1 \cdots \vec{x}_q) \quad \langle q \rangle \quad & (\vec{y}_1 \cdots \vec{y}_q \vec{w}_1 \cdots \vec{w}_r) \\ := \quad & (\vec{u}_1 \cdots \vec{u}_p)(\vec{x}_1 \cdot \vec{y}_1) \cdots (\vec{x}_q \cdot \vec{y}_q)(\vec{w}_1 \cdots \vec{w}_r) \\ = \quad & (\vec{x}_1 \cdot \vec{y}_1) \cdots (\vec{x}_q \cdot \vec{y}_q)(\vec{u}_1 \cdots \vec{u}_p \vec{w}_1 \cdots \vec{w}_r). \end{aligned} \tag{7.1.4}$$

Then for any $P \in U_1 \otimes \cdots \otimes U_p \otimes V_1 \otimes \cdots \otimes V_q$ and any $R \in V_1 \otimes \cdots \otimes V_q \otimes W_1 \otimes \cdots \otimes W_r$ we would like to define $P\langle q \rangle R \in U_1 \otimes \cdots \otimes U_p \otimes W_1 \otimes \cdots \otimes W_r$ by writing (7.1.2) and

$$P\langle q \rangle R := \sum_{i=1}^{m} \sum_{j=1}^{n} P_i \langle q \rangle R_j. \tag{7.1.5}$$

[Incidentally, note the mnemonic device that

$$\begin{aligned}&(\vec{u}_1 \cdots \vec{u}_p \vec{x}_1 \cdots \vec{x}_q)\langle q \rangle (\vec{y}_1 \cdots \vec{y}_q \vec{w}_1 \cdots \vec{w}_r) \\ &:= \vec{u}_1 \cdots \vec{u}_p \begin{pmatrix} \vec{x}_1 \\ \cdot \\ \vec{x}_q \end{pmatrix} \begin{matrix} \cdots \\ \\ \cdots \end{matrix} \begin{pmatrix} \vec{y}_1 \\ \cdot \\ \vec{y}_q \end{pmatrix} \vec{w}_1 \cdots \vec{w}_r \quad . \end{aligned} \tag{7.1.6}$$

Some authors define

$$\begin{aligned}(\vec{u}_1 \cdots \vec{u}_p \vec{x}_1 \cdots \vec{x}_q) \quad \langle q \rangle \quad & (\vec{y}_q \cdots \vec{y}_1 \cdots \vec{w}_1 \cdots \vec{w}_r) \\ := \quad & \begin{matrix} (\vec{x}_q \cdot \vec{y}_q) \\ \vdots \quad \vdots \\ (\vec{x}_2 \cdot \vec{y}_2) \\ \vec{u}_1 \cdots \vec{u}_p (\vec{x}_1 \cdot \vec{y}_1) \vec{w}_1 \cdots \vec{w}_r \end{matrix} \\ = \quad & (\vec{x}_q \cdot \vec{y}_q) \cdots (\vec{x}_1 \cdot \vec{y}_1) \vec{u}_1 \cdots \vec{u}_p \vec{w}_1 \cdots \vec{w}_r. \end{aligned}$$

De gustabus non disputandum est, but I think the first definition has more convenient properties.]

To prove that these definitions are unambiguous, we define

$$\begin{aligned} M \quad \in \quad & \mathcal{M}(U_1 \times \cdots \times U_p \times V_1 \times \cdots \times V_q \times V_1 \times \cdots \times V_q \times W_1 \times \cdots \times W_r \to \\ & U_1 \otimes \cdots \otimes U_p \otimes W_1 \otimes \cdots \otimes W_r) \end{aligned}$$

by requiring

$$M\left(\vec{u}_1,\cdots,\vec{u}_p,\vec{x}_1,\cdots,\vec{x}_q,\vec{y}_1,\cdots\vec{y}_q,\vec{w}_1,\cdots,\vec{w}_r\right):= \\ =\left(\vec{x}_1\cdot\vec{y}_1\right)\cdots\left(\vec{x}_q\cdot\vec{y}_q\right)\vec{u}_1\cdots\vec{u}_p\vec{w}_1\cdots\vec{w}_r. \tag{7.1.7}$$

Since M is multilinear, the lifting theorem provides a unique

$$\begin{aligned} M_{\otimes} \in\ & \mathcal{L}(U_1\otimes\cdots\otimes U_p\otimes V_1\otimes\cdots\otimes V_q\otimes V_1\otimes\cdots\otimes V_q\otimes W_1\otimes\cdots\otimes W\to \\ & U_1\otimes\cdots\otimes U_p\otimes W_1\otimes\cdots\otimes W_r)\end{aligned}$$

satisfying

$$\begin{aligned} & M_{\otimes}\left(\vec{u}_1\otimes\cdots\otimes\vec{u}_p\otimes\vec{x}_1\otimes\cdots\otimes\vec{x}_q\otimes\vec{y}_1\otimes\cdots\otimes\vec{y}_q\otimes\vec{w}_1\otimes\cdots\otimes\vec{w}_r\right) \\ = \quad & M\left(\vec{u}_1,\cdots,\vec{u}_p,\vec{x}_1,\cdots,\vec{x}_q,\vec{y}_1,\cdots,\vec{y}_q,\vec{w}_1,\cdots,\vec{w}_r\right) \\ = \quad & \left(\vec{x}_1\cdot\vec{y}_1\right)\cdots\left(\vec{x}_q\cdot\vec{y}_q\right)\vec{u}_1\cdots\vec{u}_p\vec{w}_1\cdots\vec{w}_r \end{aligned} \tag{7.1.8}$$

for all

$$\begin{aligned} & \left(\vec{u}_1,\cdots,\vec{u}_p,\vec{x}_1,\cdots,\vec{x}_q,\vec{y}_1,\cdots,\vec{y}_q,\vec{w}_1,\cdots,\vec{w}_r,\right) \\ & \quad \in U_1\times\cdots\times U_p\times V_1\times\cdots\times V_q\times V_1\times\cdots\times V_q\times V_1\times\cdots\times V_q\times W_1\times\cdots\times W_r. \end{aligned}$$

For any $P\in U_1\otimes\cdots\otimes U_p\otimes V_1\otimes\cdots\otimes V_q$ and any $R\in V_1\otimes\cdots\otimes V_q\otimes W_1\otimes\cdots\otimes W_r$, we define

$$P\langle q\rangle R := M_{\otimes}(PR). \tag{7.1.9}$$

Then if P and R are polyads, (7.1.8) says we do have (7.1.4).

We can calculate $P\langle q\rangle R$ for any tensors and P and R by writing them as sums of polyads and doing the arithmetic justified by

Remark 7.1.35 Suppose $P_1,\cdots,P_m\in U_1\otimes\cdots\otimes U_p\otimes V_1\otimes\cdots\otimes V_q$ and $R_1,\cdots,R_n\in V_1\otimes\cdots\otimes V_q\otimes W_1\otimes\cdots\otimes W_r$ and $a_1,\cdots,a_m,b_1,\cdots,b_n\in\mathcal{R}$. Then

$$\left(a_iP_i\right)\langle q\rangle\left(b_jR_j\right)=a_ib_j\left(P_i\circ R_j\right) \tag{7.1.10}$$

Proof:

$$
\begin{aligned}
(a_iP_i)\langle q\rangle(b_jR_j) &= M_{\otimes}\left[(a_iP_i)(b_jR_j)\right] \text{ by definition of } \langle q\rangle \\
&= M_{\otimes}\left[a_ib_j(PiRj)\right] \text{by multilinearity of the tensor product} \\
&= a_ib_jM_{\otimes}(P_iR_j) \text{ because } M_{\otimes} \text{ is linear} \\
&= a_ib_j(P_i\langle q\rangle R_j) \text{ by definition of } \langle q\rangle.
\end{aligned}
$$

Equation (7.1.10) is true for any P_i and R_j, but when P_i and R_j are polyads, (7.1.10) and (7.1.4) give a convenient way to calculate $P\langle q\rangle R$ for any tensors P and R.

7.2 Components of $P\langle q\rangle R$

Suppose $P \in U_1 \otimes \cdots \otimes U_p \otimes V_1 \otimes \cdots \otimes V_q$ and $R \in V_1 \otimes \cdots \otimes V_q \otimes W_1 \otimes \cdots \otimes W_r$. Suppose $(\vec{b}_1^{(1)}, \cdots, \vec{b}_{m_i}^{(i)})$ is an ordered basis for U_i, and $(\vec{\beta}_1^{(j)}, \cdots \vec{\beta}_{n_j}^{(j)})$ is an ordered basis for V_j, and $(\vec{6}_1^{(R)}, \cdots \vec{6}_{P_k}^{(R)})$ is an ordered basis for W_k. We would like to compute the components of $P\langle q\rangle R$ relative to the basis sequence for $U_1 \times \cdots \times U_p \times W_1 \times \cdots \times W_r$. We have

$$
\begin{aligned}
P &= P^{i_1\cdots i_p j_1\cdots j_q}\vec{b}_{i_1}^{(1)}\cdots\vec{b}_{i_p}^{(p)}\vec{\beta}_{j_1}^{(1)}\cdots\vec{\beta}_{j_q}^{(q)} \\
R &= R_{k_1\cdots k_q}{}^{l_1\cdots l_r}\vec{\beta}_{(1)}^{k_1}\cdots\vec{\beta}_{(q)}^{k_q}\vec{6}_{l_1}^{(1)}\cdots\vec{6}_{l_r}^{(r)}
\end{aligned}
$$

($\vec{\beta}_{(p)}^{k_p}$ are the dual basis vectors to $\vec{\beta}_{k_p}^{(p)}$). Then (7.1.4) gives

$$
\begin{aligned}
P\langle q\rangle R &= P^{i_1\cdots i_p j_1\cdots j_q}R_{k_1\cdots k_q}{}^{l_1\cdots l_r}\left(\vec{\beta}_{j_1}^{(1)}\cdot\vec{\beta}_{(1)}^{k_1}\right)\cdots\left(\vec{\beta}_{j_q}^{(q)}\cdot\vec{\beta}_{(q)}^{k_q}\right)\vec{b}_{i_1}^{(1)}\cdots\vec{b}_{i_p}^{(p)}\vec{6}_{l_1}^{(1)}\cdots 6_{l_r}^{(r)} \\
&= P^{i_1\cdots i_p j_1\cdots j_q}R_{k_1\cdots k_q}{}^{l_1\cdots l_r}\delta_{j_1}^{k_1}\cdots\delta_{j_q}^{k_q}\beta_{i_1}^{(1)}\cdots\vec{b}_{i_p}^{(p)}\vec{6}_{l_1}^{(1)}\cdots\vec{6}_{l_r}^{(r)}
\end{aligned}
$$

But $R_{k_1\cdots k_q}{}^{l_1\cdots l_r}\delta_{j_1}{}^{k_1}\cdots\delta_{j_q}{}^{k_q} = R_{j_1\cdots j_q}{}^{l_1\cdots l_r}$ so

$$
P\langle q\rangle R = \left(P^{i_1\cdots i_p j_1\cdots j_q}R_{j_1\cdots j_q}{}^{l_1\cdots l_r}\right)\vec{b}_{i_1}^{(1)}\cdots\vec{b}_{i_p}^{(p)}\vec{6}_{l_1}^{(1)}\cdots\vec{6}_{l_r}^{(r)}.
$$

Therefore, by note (5.2.1),

$$
(P\langle q\rangle R)^{i_1\cdots i_p l_1\cdots l_r} = P^{i_1\cdots i_p j_1\cdots j_q}R_{j_1\cdots j_q}{}^{l_1\cdots l_r}. \tag{7.2.1}
$$

It is important to remember that for (7.2.1) to hold, the same bases for $V_1, \cdots, V_q$ (and their dual bases) must be used to calculate the components of P and R.

Covariant or contravariant metric matrices can be used to raise or lower any index i or l or pair j in (7.2.1). The same collection of 2^{p+q+r} formulas can be obtained immediately from (7.2.1) by choosing to regard certain of the original bases for the U_i, V_j or W_k as dual bases, the duals becoming the original bases. For example,

$$(P\langle q\rangle R)_{i_1}{}^{i_2\cdots i_p l_1\cdots l_{r-1}}{}_{l_r} =$$
$$= P_{i_1}{}^{i_2\cdots i_p j_1}{}_{j_2\cdots j_g} R_{j_1}{}^{j_2\cdots j_q l_1\cdots l_{r-1}}{}_{l_r}$$

All the 2^{p+q+r} variants of (7.2.1) collapse into a single formula when all the bases are orthonormal.

7.3 Properties of the generalized dot product

Property 7.3.1 Bilinearity. For $i \in \{1, \cdots, m\}$ and $j \in \{1, \cdots, n\}$, suppose $P_i \in U_1 \otimes \cdots \otimes U_p \otimes V_1 \otimes \cdots \otimes V_q$ and $R_j \in V_1 \otimes \cdots \otimes V_q \otimes W_1 \otimes \cdots \otimes W_r$ and $a_i, b_j \in \mathcal{R}$. Then $(a_i P_i)\langle q\rangle(b_j R_j) = a_i b_j (P_i\langle q\rangle R_j)$.

This has already been proved as remark (7.1.35).

Property 7.3.2 Suppose $P \in U_1 \otimes \cdots \otimes U_p \otimes V_1 \otimes \cdots \otimes V_q$ and $R \in V_1 \otimes \cdots \otimes V_q \otimes W_1 \otimes \cdots \otimes W_q$. Then if $P = 0$ or $R = 0$, $P\langle q\rangle R = 0$.

Proof:

Suppose $P = 0$. Then $0P = P$ so $P\langle q\rangle R = (0P)\langle q\rangle R = 0(P \circ R) = 0$.

Note that the converse is false. We can have $P\langle q\rangle R = 0$ even though $P \neq 0$ and $R \neq 0$. For example $\vec{u}\langle 1\rangle\vec{v} = \vec{u} \cdot \vec{v}$ and this can vanish even though $\vec{u} \neq \vec{0}$ and $\vec{v} \neq \vec{0}$.

Definition 7.3.20 Since $\vec{u}\langle 1\rangle\vec{v} = \vec{u} \cdot \vec{v}$ for vectors, it is usual to write $P\langle 1\rangle R$ as $P \cdot R$, $P\langle 2\rangle R$ as $P : R$, $P\langle 3\rangle R$ as $P \mathbin{\dot{\cdot\cdot}} R$ and $P\langle 4\rangle R$ as $P :: R$. The notation $P \mathbin{\dot{::}} R$ for $P\langle 5\rangle R$ seems not to be used.

Property 7.3.3 $P\langle 0\rangle R = PR$.

Proof:

If P and R are polyads, this is obvious from (7.1.4) with $q = 0$. Both $P\langle 0\rangle R$ and PR depend linearly on R, so if $R = \sum R_j$ where the R_j are polyads, and if P is a polyad, we have $P\langle 0\rangle R = P\langle 0\rangle(\sum_j R_j) = \sum_j(P\langle 0\rangle R_j) = \sum_j(PR_j) = P(\sum_j R_j) = PR$. Therefore $P\langle 0\rangle R = PR$ if P is a polyad and R is any tensor. If $P = \sum_i P_i$ where the P_i are polyads, then for any tensor R, $P0R = (\sum_i P_i)\langle 0\rangle R = \sum_i(P_i\langle 0\rangle R) = \sum_i(P_iR) = (\sum_i P_i)R = PR$. Thus $P\langle 0\rangle R = PR$ for any tensors P and R.

Property 7.3.4 $P\langle q\rangle R$ and $R\langle q\rangle P$ need not be the same even if both are defined. This is obvious from property 7.3.3 and the corresponding result for tensor products. An example with $q = 1$ is this. Let $\vec{u}, \vec{v}, \vec{w}, \vec{x} \in V$. Let $P = \vec{u}\vec{v}$, $R = \vec{w}\vec{x}$. Then $P \cdot R = (\vec{v} \cdot \vec{w})\vec{u}\vec{x}$ and $R \cdot P = (\vec{x} \cdot \vec{u})\vec{w}\vec{v}$. If these two dyads are equal and not 0, remark (4.3.30) shows that $\vec{u}$ must be a multiple of $\vec{w}$ and $\vec{x}$ must be a multiple of $\vec{v}$.

Property 7.3.5 (Associativity). Suppose

$$
\begin{aligned}
&P \in U_1 \otimes \cdots \otimes U_p \otimes V_1 \otimes \cdots \otimes V_q \\
&R \in \qquad\qquad V_1 \otimes \cdots \otimes V_q \otimes W_1 \otimes \cdots \otimes W_r \otimes X_1 \otimes \cdots \otimes X_s \\
&T \in \qquad\qquad\qquad\qquad X_1 \otimes \cdots \otimes X_s \otimes Y_1 \otimes \cdots \otimes Y_t.
\end{aligned}
$$

Then

$$P\langle q\rangle\,(R\langle s\rangle T) = (P\langle q\rangle R)\,\langle s\rangle T. \tag{7.3.1}$$

Proof:

First we prove (7.3.1) when P, R, T are polyads. In that case (7.3.1) follows from (7.1.4) by simple computation, which we leave to the reader. That (7.3.1) holds for polyads can also be seen immediately from the mnemonic diagram (7.1.6). The generalized dot products $\langle q\rangle$ and $\langle s\rangle$ are formed by

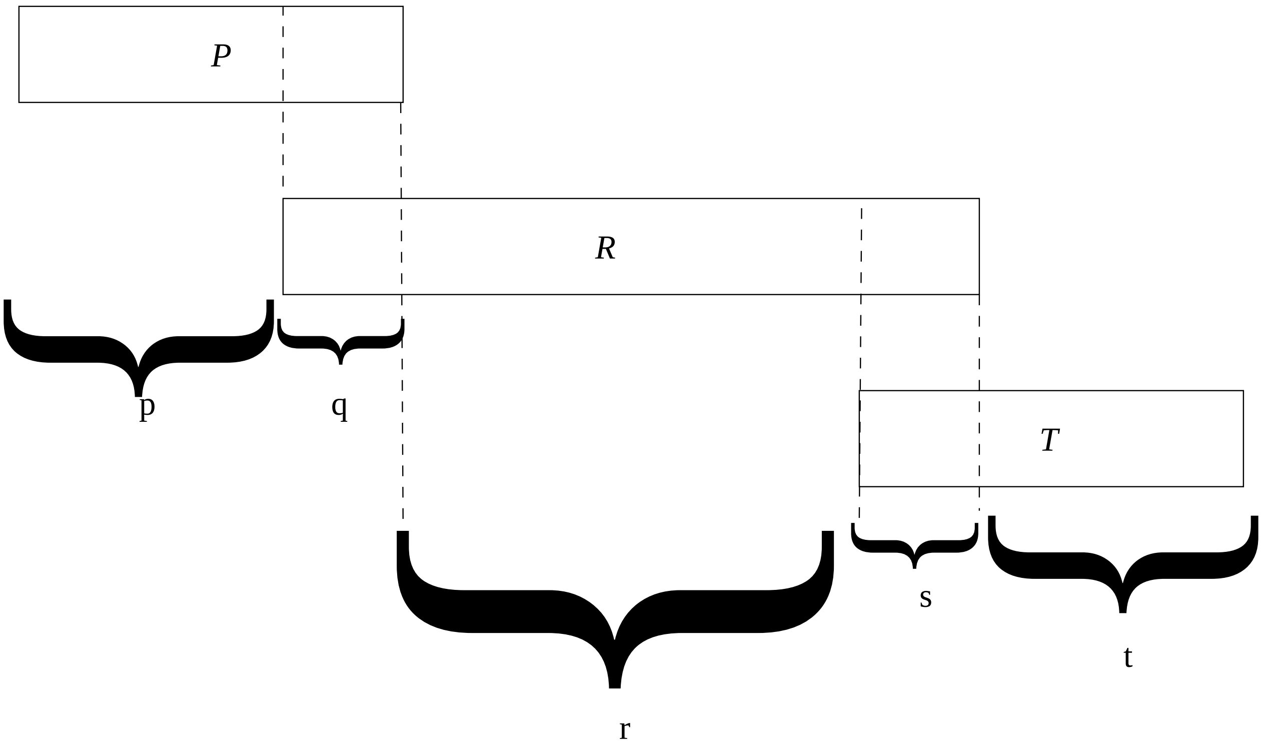


Figure 7.1:

dotting vertical pairs of vectors in the regions of overlap of P, R, T. It is clearly irrelevant to the outcome, whether we dot first the q vectors in P with those below them in R or the s vectors in R with those below them in T.

Next, suppose P, R and $T_1, \cdots, T_n$ are polyads and $T = \sum_i T_i$. Then $P\langle q\rangle(R\langle s\rangle T) = P\langle q\rangle\left[R\langle s\rangle(\sum_i T_i)\right] = P\langle q\rangle\left[\sum_i(R\langle s\rangle T_i)\right] = \sum_i\left[P\langle q\rangle(R\langle s\rangle T_i)\right] = \sum_i\left[P\langle q\rangle(R\langle s\rangle T_i)\right] = (P\langle q\rangle R)\langle s\rangle\left(\sum T_i\right) = (P\langle q\rangle R)\langle s\rangle T$. This proves (7.3.1) when P and R are polyads and T is any tensor.

Next suppose P, $R_1, \cdots, R_n$ are polyads and T is any tensor and $R = \sum_i R_i$. Then $P\langle q\rangle(R\langle s\rangle T) = P\langle q\rangle\left[(\sum_i R_i)\langle s\rangle T\right] = P\langle q\rangle\left[\sum_i(R_i\langle s\rangle T)\right] = \sum_i\left[P\langle q\rangle(R_i\langle s\rangle T)\right] = \sum_i\left[(P\langle q\rangle R_i)\langle s\rangle T\right] = \left[\sum_i(P\langle q\rangle R_i)\right]\langle s\rangle T = \left[P\langle q\rangle(E_i R_i)\right]\langle s\rangle T = (P\langle q\rangle R)\langle s\rangle T$. Thus (7.3.1) holds when P is a polyad and R and T are any tensors.

Finally, suppose $P_1, \cdots, P_n$ are polyads and R and T are any tensors, and $P = \sum_i P_i$, Then $P\langle q\rangle(R\langle s\rangle T) = (\sum_i P_i)\langle q\rangle(R\langle s\rangle T) = \sum_i\left[P_i\langle q\rangle(R\langle s\rangle T)\right] =$

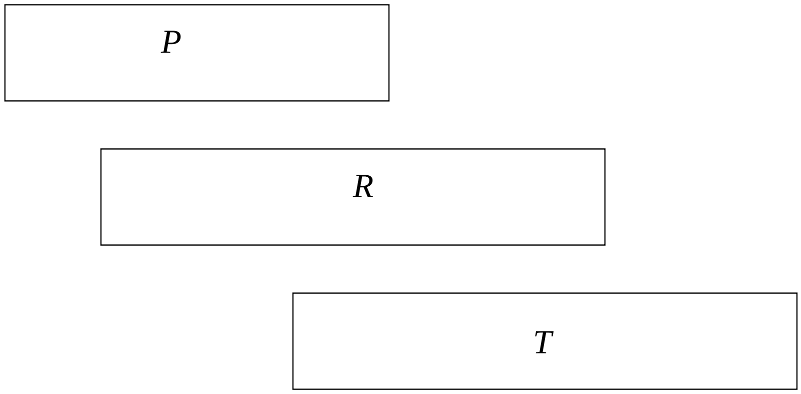


Figure 7.2:

$\sum_i [(P_i\langle q\rangle R)\langle s\rangle T] = [\sum_i (P_i\langle q\rangle R)]\,\langle s\rangle T = [(\sum_i P_i)\langle q\rangle R]\,\langle s\rangle T = (P\langle q\rangle R)\langle s\rangle T$. This proves (7.3.1) for any tensors P, R, T.

Figure (7.1) also shows how associativity can fail. If P and T overlap as in figure (7.2), there is trouble. The tensors $P\langle q\rangle(R\langle s\rangle T)$ and $(P\langle q\rangle R)\,\langle s\rangle T$ may be different even if both are defined (they may not be). As an example, suppose all tensors are over a single space V. Let $\vec u, \vec v, \vec w, \vec x, \vec y, \in V$. Let $P = \vec u\vec v$, $R = \vec w$, $T = \vec x\vec y$. Then we have the diagram

$$\begin{array}{rcl} P & = & \square\square \\ R & = & \;\;\,\square \\ T & = & \;\;\,\square\square \end{array}$$

which exhibits the fatal $P - T$ overlap shown in Figure (7.2). We have $P\cdot(R\cdot T) = \vec u\vec v\cdot[(\vec w\cdot\vec x)\vec y] = (\vec w\cdot\vec x)\,[(\vec u\vec v)\cdot\vec y] = (\vec w\cdot\vec x)(\vec v\cdot\vec y)\vec u$ and $(P\cdot R)\cdot T = [\vec u(\vec v\cdot\vec w)]\cdot(\vec x\vec y) = (\vec v\cdot\vec w)\,[\vec u\cdot(\vec x\vec y)] = (\vec v\cdot\vec w)(\vec u\cdot\vec x)\vec y$. If we choose $\vec x$ and $\vec y$ to be mutually perpendicular unit vectors, and take $\vec v = \vec y$, $\vec u = \vec w = \vec x$, we have $P\cdot(R\cdot T) = \vec u = \vec x$ and $(P\cdot R)\cdot T = \vec 0$.

Property 7.3.6 Suppose P, $R \in V_1\otimes\cdots\otimes V_q$. Then

$$P\langle q\rangle R = R\langle q\rangle P. \tag{7.3.2}$$

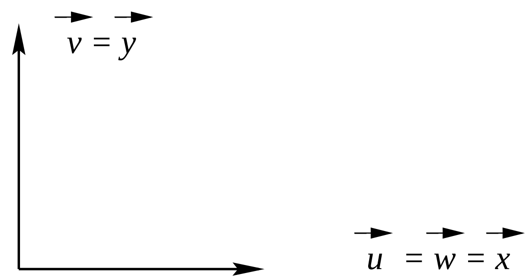


Figure 7.3:

Proof:

If $P = \vec{x}_1\vec{x}_2\cdots\vec{x}_q$ and $R = \vec{y}_1\vec{y}_2\cdots\vec{y}_q$ then $P\langle q\rangle R = (\vec{x}_1\cdot\vec{y}_1)\cdots(\vec{x}_q\cdot\vec{y}_q) = (\vec{y}_1\cdot\vec{x}_1)\cdots(\vec{y}_q\cdot\vec{x}_q) = R\langle q\rangle P$, so (7.3.2) is true of P and R are polyads. But both $R\langle q\rangle P$ and $P\langle q\rangle R$ are bilinear in P and R, so the proof of (7.3.2) for general tensors P and R in $V_1\otimes\cdots\otimes V_q$ is like the proof of (7.3.1).

Property 7.3.7 Suppose $P\in V_1\otimes\cdots\otimes V_q$ and $P\neq 0$. Then

$$P\langle q\rangle P > 0. \tag{7.3.3}$$

Proof:

For $i\in\{1,\cdots,q\}$, let $(\hat{x}_1^{(i)},\cdots,\hat{x}_{n_i}^{(i)})$ be an *orthonormal* basis for V_i. By (7.2.1),

$$P\langle q\rangle P = P_{i_1\cdots i_q}P_{i_1\cdots i_q}. \tag{7.3.4}$$

This is a sum of squares of real numbers. It is > 0 unless every term is 0, i.e. $P_{i_1\cdots i_q} = 0$. In that case $P = 0$.

Theorem 7.3.17 On the vector space $V_1\otimes\cdots\otimes V_q$, define the dot product $dp(P,R) = P\langle q\rangle R$ for all P, $R\in V_1\otimes\cdots\otimes V_q$. With this dot product, $V_1\otimes\cdots\otimes V_q$ is a Euclidean vector space.

Proof:

Properties (7.2.1), (7.3.2), (7.3.3) show that dp satisfies $a)$ $b)$, $c)$, $d)$ on page D-18.

Corollary 7.3.16 Let $(\hat{x}_1^{(i)}, \cdots, \hat{x}_{n_i}^{(i)})$ be an orthonormal basis for V_i, $i = 1, \cdots, q$. Then the $n_1 n_2 \cdots n_q$ polyads $\hat{x}_{i_1}^{(1)} \hat{x}_{i_2}^{(2)} \cdots \hat{x}_{i_q}^{(1)}$ are an orthonormal basis for $V_1 \otimes \cdots \otimes V_q$.

Proof:

By theorem 5.1.15 they are a basis. To prove them orthonormal, we note

$$(\hat{x}_{i_1}^{(1)} \cdots \hat{x}_{i_q}^{(q)}) \langle q \rangle (\hat{x}_{j_1}^{(1)} \cdots \hat{x}_{j_q}^{(q)}) = (\hat{x}_{i_1}^{(1)} \cdot \hat{x}_{j_1}^{(1)}) \cdots (\hat{x}_{i_q}^{(q)} \cdot \hat{x}_{j_q}^{(q)}) = \delta_{i_1 j_1} \delta_{i_2 j_2} \cdots \delta_{i_q j_q}.$$

Definition 7.3.21 If $P \in V_1 \otimes \cdots \otimes V_q$, define the "length" of P, $\|P\|$, in the way one usually does for vectors in a Euclidean space. That is, $\|P\|$ is the non-negative square root of $P\langle q \rangle P$.

Remark 7.3.36 Suppose $P \in U_1 \otimes \cdots \otimes U_p \otimes V_1 \otimes \cdots \otimes V_q$ and $R \in V_1 \otimes \cdots \otimes V_q \otimes W_1 \otimes \cdots \otimes W_r$. Then

$$\|P\langle q \rangle R\| \leq \|P\| \|R\|. \tag{7.3.5}$$

Proof:

Introduce an *orthonormal* basis for each space U_i , V_j and W_k. Take components relative to these bases. Abbreviate the p-tuple of integers $(i_1, \cdots i_p)$ as I, the q-tuple $(j_1, \cdots, j_q)$ as J, the r-tuple $(l_1, \cdots l_r)$ as L. Then (7.2.1) can be abbreviated

$$(P\langle q \rangle R)_{IL} = \sum_J P_{IJ} R_{JL}.$$

By the ordinary Schwarz inequality for real numbers,

$$\left(\sum_J P_{IJ} R_{JL} \right)^2 \leq \left(\sum_J P_{IJ}^2 \right) \left(\sum_K R_{KL}^2 \right)$$

where K stands for the q-tuple $(k_1, \cdots, k_q)$. (We have simply changed the summation index from J to K.) Summing over I and L gives

$$\sum_{I,L} \left(\sum_J P_{IJ} R_{JL} \right)^2 \leq \left(\sum_{I,J} P_{IJ}^2 \right) \left(\sum_{K,L} R_{KL}^2 \right),$$

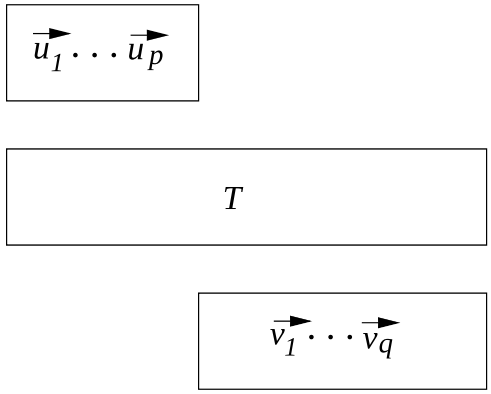


Figure 7.4:

or

$$\sum_{I,L} (P\langle q\rangle R)^2_{IL} \le \left(\sum_{I,J} P^2_{IJ}\right)\left(\sum_{K,L} R^2_{KL}\right).$$

By equation (7.3.4), this last inequality is

$$\|P\langle q\rangle R\|^2 \le \|P\|^2\|R\|^2.$$

Taking square roots gives (7.3.5).

7.4 Applications of the generalized dot product

Application 1: The value of a tensor.

Suppose $T \in U_1 \otimes \cdots \otimes U_p \otimes V_1 \otimes \cdots \otimes V_q$ and $(\vec{u}_1, \cdots, \vec{u}_p, \vec{v}_1, \cdots, \vec{v}_q) \in U_1 \times \cdots \times U_p \times V_1 \times \cdots \times V_q$. Then

$$T(\vec{u}_1, \cdots, \vec{u}_p, \vec{v}_1, \cdots, \vec{v}_q,) = (\vec{u}_1\vec{u}_2 \cdots \vec{u}_p)\langle p\rangle T\langle q\rangle (\vec{v}_1\vec{v}_2 \cdots \vec{v}_q). \tag{7.4.1}$$

Note that (7.3.1) applies, so no parentheses are needed to say whether $\langle p\rangle$ or $\langle q\rangle$ is calculated first. There is none of the overlap shown in figure 7.2. The diagram for (7.4.1) is shown in Figure 7.4. Therefore, at least (7.4.1) is unambiguous. To prove it true, suppose first that T is a polyad, $T = \vec{x}_1 \cdots \vec{x}_p\vec{y}_1 \cdots \vec{y}_q$. Then

$$(\vec{u}_1 \cdots \vec{u}_p) \quad \langle p\rangle \quad T\langle q\rangle (\vec{v}_1 \cdots \vec{v}_q) = [(\vec{u}_1 \cdots \vec{u}_p)\langle p\rangle T]\langle q\rangle (\vec{v}_1 \cdots \vec{v}_q)$$

$$\begin{aligned}
&= [(\vec{u}_1 \cdot \vec{x}_1) \cdots (\vec{u}_p \cdot \vec{x}_p)\, \vec{y}_1 \cdots \vec{y}_q]\, \langle q \rangle\, (\vec{v}_1 \cdots \vec{v}_q) \\
&= (\vec{u}_1 \cdot \vec{x}_1) \cdots (\vec{u}_p \cdot \vec{x}_p)\, [(\vec{y}_1 \cdots \vec{y}_q)\, \langle q \rangle\, (\vec{v}_1 \cdots \vec{v}_q)] \\
&= (\vec{x}_1 \cdot \vec{u}_1) \cdots (\vec{x}_p \cdot \vec{u}_p)\, (\vec{y}_1 \cdot \vec{v}_1) \cdots (\vec{y}_q \cdot \vec{v}_q) \\
&= T\,(\vec{u}, \cdots, \vec{u}_p, \vec{v}_1, \cdots, \vec{v}_q)\,.
\end{aligned}$$

This proves (7.4.1) when T is a polyad. But both sides of (7.4.1) depend linearly on T, so if (7.4.1) is true for polyads T, it is true when T is a sum of polyads. That is, (7.4.1) is true for any tensor $T \in U_1 \otimes \cdots \otimes U_p \otimes V_1 \otimes \cdots \otimes V_q$.

Corollary 7.4.17 Suppose P is a fixed tensor in $U_1 \otimes \cdots \otimes U_p \otimes V_1 \otimes \cdots \otimes V_q$, with the property that $P\langle q \rangle R = 0$ for every $R \in V_1 \otimes \cdots \otimes V_q \otimes W_1 \otimes \cdots \otimes W_r$. Then $P = 0$. (Also, if R is a fixed tensor in $V_1 \otimes \cdots \otimes V_q \otimes W_1 \otimes \cdots \otimes W_r$ with the property that $P\langle q \rangle R = 0$ for every $P \in U_1 \otimes \cdots \otimes U_p \otimes V_1 \otimes \cdots V_q$ then $R = 0$).

Proof:

Choose a fixed *nonzero* $T \in W_1 \otimes \cdots \otimes W_r$. For any $(\vec{v}_1, \cdots, \vec{v}_q) \in V_1 \times \cdots \times V_q$, we know by hypothesis that $P\langle q \rangle\, [(\vec{v}_1 \cdots \vec{v}_q)T] = 0$. Figure (7.3) works, so $[P\langle q \rangle(\vec{v}_1 \cdots \vec{v}_q)]T = 0$. By remark (4.2.25), $P\langle q \rangle(\vec{v}_1 \cdots \vec{v}_q) = 0$. Then for any $(\vec{u}_1, \cdots, \vec{u}_p) \in U_1 \times \cdots \times U_p$ we have $(\vec{u}_1 \cdots \vec{u}_p)\langle p \rangle[P\langle q \rangle(\vec{v}_1 \cdots \vec{v}_q)] = 0$. By equation (7.4.1), $P(\vec{u}_1, \cdots, \vec{u}_p, \vec{v}_1, \cdots, \vec{v}_q) = 0$. Since $(\vec{u}_1, \cdots, \vec{u}_p, \vec{v}_1, \cdots, \vec{v}_q)$ is an arbitrary member of $U_1 \times \cdots \times U_p \times V_1 \times \cdots \times V_q$, therefore $P = 0$. (The other half of the corollary, in parentheses, is proved similarly.

Corollary 7.4.18 Suppose P, $P' \in U_1 \otimes \cdots \otimes U_p \otimes V_1 \otimes \cdots \otimes V_q$ and for every $R \in V_1 \otimes \cdots \otimes V_q \otimes W_1 \otimes \cdots \otimes W_r$ we have $P\langle q \rangle R = P'\langle q \rangle R$. Then $P = P'$. (Also, if R, $R' \in V_1 \otimes \cdots \otimes V_q \otimes W_1 \otimes \cdots \otimes W_r$ and for every $P \in U_1 \otimes \cdots \otimes U_p \otimes V_1 \otimes \cdots \otimes V_q$ we have $P\langle q \rangle R = P\langle q \rangle R'$, then $R = R'$.)

Application 2: Vector cross product.

Suppose (V, A) is an oriented *three dimensional* Euclidean space whose alternating tensor is A. For any $\vec{u}$, $\vec{v} \in V$, we define their A-cross product as

$$\vec{u} \overset{A}{\times} \vec{v} = A\langle 2 \rangle\, (\vec{u}\vec{v})\,. \tag{7.4.2}$$

Usually we omit the A and write simply $\vec{u} \times \vec{v}$, but it should always be remembered that there are two oriented 3-spaces, (V, A) and $(V, -A)$, and so two ways to define $\vec{u} \times \vec{v}$.

We will now show that our cross product is the usual one. First, if a_i, $b_j \in \mathcal{R}$ and $\vec{u}_i$, $\vec{v}_j \in V$ for $i = 1, \cdots, m$ and $j = 1, \cdots, n$

$$(a_i \vec{u}_i) \times (b_j \vec{v}_j) = a_i b_j \left(\vec{u}_i \times \vec{v}_j\right). \tag{7.4.3}$$

For $(a_i \vec{u}_i) \times (b_j \vec{v}_j) = A\langle 2\rangle \left[(a_i \vec{u}_i)(b_j \vec{v}_j)\right] = A\langle 2\rangle \left[(a_i b_j \vec{u}_i \vec{v}_j)\right] = a_i b_j \left[A\langle 2\rangle (\vec{u}_i \vec{v}_j)\right] = a_i b_j (\vec{u}_i \times \vec{v}_j)$. Second, $\vec{w} \cdot (\vec{u} \times \vec{v}) = \vec{w} \cdot [A\langle 2\rangle (\vec{u}\vec{v})] = A(\vec{w}, \vec{u}, \vec{v}) = A(\vec{u}, \vec{v}, \vec{w})$ so

$$(\vec{u} \times \vec{v}) \cdot \vec{w} = A\left(\vec{u}, \vec{v}, \vec{w}\right). \tag{7.4.4}$$

Therefore, since A is totally antisymmetric,

$$\left\{ \begin{array}{l} (\vec{u} \times \vec{v}) \cdot \vec{w} = (\vec{v} \times \vec{w}) \cdot \vec{u} = (\vec{w} \times \vec{u}) \cdot \vec{v} \\ = -\left(\vec{v} \times \vec{u}\right) \cdot \vec{w} = -\left(\vec{u} \times \vec{w}\right) \cdot \vec{v} = -\left(\vec{w} \times \vec{v}\right) \cdot \vec{u}. \end{array} \right. \tag{7.4.5}$$

In particular, $(\vec{u} \times \vec{v}) \cdot \vec{w} = -(\vec{v} \times \vec{u}) \cdot \vec{w}$ for all $\vec{w}$, so

$$\vec{v} \times \vec{u} = -\vec{u} \times \vec{v}. \tag{7.4.6}$$

Therefore, $\vec{u} \times \vec{u} = \vec{0}$. Setting $\vec{w} = \vec{u}$ in $(\vec{u} \times \vec{v}) \cdot \vec{w} = (\vec{w} \times \vec{u}) \cdot \vec{v}$ gives $(\vec{u} \times \vec{v}) \cdot \vec{u} = (\vec{u} \times \vec{u}) \cdot \vec{v} = 0$. Thus, using (7.4.6),

$$(\vec{u} \times \vec{v}) \cdot \vec{u} = (\vec{u} \times \vec{v}) \cdot \vec{v} = 0. \tag{7.4.7}$$

Relative to any ordered basis $(\vec{b}_1, \vec{b}_2, \vec{b}_3)$ for V we have

$$(\vec{u} \times \vec{v})_i = A_{ijk} u^j v^k = A\left(\vec{b}_1, \vec{b}_2, \vec{b}_3\right) \varepsilon_{ijk} u^j v^k.$$

If the ordered basis is positively oriented and orthonormal, $A(\vec{b}_1, \vec{b}_2, \vec{b}_3) = 1$, so

$$(\vec{u} \times \vec{v})_i = \varepsilon_{ijk} u_j v_k. \tag{7.4.8}$$

If the ordered basis is negatively oriented and orthonormal, $A(\vec{b}_1, \vec{b}_2, \vec{b}_3) = -1$ so

$$(\vec{u} \times \vec{v})_i = -\varepsilon_{ijk} u_j v_k. \tag{7.4.9}$$

The notorious formula

$$\vec{u} \times (\vec{v} \times \vec{w}) = (\vec{u} \cdot \vec{w})\, \vec{v} - (\vec{u} \cdot \vec{v})\, \vec{w} \tag{7.4.10}$$

can be derived from (7.4.8) as follows: relative to any positively oriented orthonormal basis

$$\begin{aligned}
[\vec{u} \times (\vec{v} \times \vec{w})]_i &= \varepsilon_{ijk} u_j (\vec{v} \times \vec{w})_k \\
&= \varepsilon_{ijk} u_j \varepsilon_{klm} v_l w_m \\
&= \varepsilon_{ijk} \varepsilon_{klm} u_j v_l w_m \\
&= (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})\, u_j v_l w_m \quad \text{see page D-9} \\
&= u_j w_j v_i - u_j v_j w_i \\
&= (\vec{u} \cdot \vec{w})\, v_i - (\vec{u} \cdot \vec{v})\, w_i = [(\vec{u} \cdot \vec{w})\, \vec{v} - (\vec{u} \cdot \vec{v})\, \vec{w}]_i \,.
\end{aligned}$$

Since the vector on the left of (7.4.5) has the same components as the vector on the right, those two vectors are equal.

From (7.4.5) and (7.4.10) we have $(\vec{u}\times\vec{v})\cdot(\vec{u}\times\vec{v}) = \vec{u}\cdot[\vec{v} \times (\vec{u} \times \vec{v})] = \vec{u}\cdot[\|\vec{v}\|^2\vec{u} - (\vec{v} \cdot \vec{u})\vec{v}] = \|\vec{u}\|^2\|\vec{v}\|^2 - (\vec{u} \cdot \vec{v})^2$, so

$$\|\vec{u} \times \vec{v}\|^2 = \|\vec{u}\|^2\|\vec{v}\|^2 - (\vec{u} \times \vec{v})^2. \tag{7.4.11}$$

We have defined the angle θ between $\vec{u}$ and $\vec{v}$ as the θ in $0 \le \theta \le \pi$ such that $\vec{u} \cdot \vec{v} = \|\vec{u}\|\|\vec{v}\| \cos\theta$. With this definition of θ, (7.4.11) implies

$$\|\vec{u} \times \vec{v}\| = \|\vec{u}\|\|\vec{v}\| \sin\theta. \tag{7.4.12}$$

Equations (7.4.7) and (7.4.12) determine that $\vec{u} \times \vec{v}$ is $\perp\vec{u}$ and $\vec{v}$ and has length $\|\vec{u}\|\|\vec{v}\| \sin\theta$. This leaves two possibilities for a nonzero $\vec{u}\times\vec{v}$. The correct one is determined by the fact that

$$A\,(\vec{u}, \vec{v}, \vec{u} \times \vec{v}) > 0 \text{ if } \vec{u} \times \vec{v} \neq 0. \tag{7.4.13}$$

To prove (7.4.3), we note that $A(\vec{u}, \vec{v}, \vec{u} \times \vec{v}) = A(\vec{u} \times \vec{v}, \vec{u}, \vec{v}) = (\vec{u} \times \vec{v}) \cdot [A\langle 2\rangle(\vec{u}\vec{v})] = (\vec{u} \times \vec{v}) \cdot (\vec{u} \times \vec{v})$. Inequality (7.4.3) is usually paraphrased by saying that "$\vec{u}, \vec{v}, \vec{u} \times \vec{v}$ is a positively oriented ordered triple." If we orient the space we live in by choosing A so

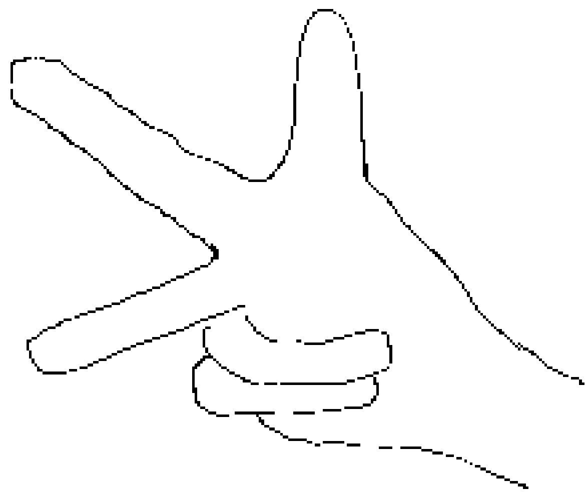


Figure 7.5:

that our right thumb, index finger and middle finger are positively oriented ordered triple of vectors when extended as in Figure (7.5), we will obtain the usual definition of $\vec{u} \times \vec{v}$ in terms of the right hand screw rule.

Application 3: Let $U_1, \cdots, U_p$, $V_1, \cdots, V_q$ be Euclidean spaces. Then

$$U_1 \otimes \cdots \otimes U_p \otimes V_1 \otimes \cdots \otimes V_q = (U_1 \otimes \cdots \otimes U_p) \otimes (V_1 \otimes \cdots \otimes V_q)\,. \tag{7.4.14}$$

Proof:

What is meant is that there is a natural isomorphism between the spaces on the two sides of (7.4.14). Before proving this, we note that the tensor product of $U_1 \otimes \cdots \otimes U_p$ and $V_1 \otimes \cdots \otimes V_q$ on the right in (7.4.14) is defined, because $U_1 \otimes \cdots \otimes U_p$ and $V_1 \otimes \cdots \otimes V_q$ are Euclidean spaces. Their dot products are $\langle p \rangle$ and $\langle q \rangle$. To construct the isomorphism whose existence is asserted by (7.4.14), let $T \in U_1 \otimes \cdots \otimes U_p \otimes V_1 \otimes \cdots \otimes V_q$ and define $\widetilde{T} : (U_1 \otimes \cdots \otimes U_p) \times (V_1 \otimes \cdots \otimes V_q) \to \mathcal{R}$ by requiring for any $P \in U_1 \otimes \cdots \otimes U_p$ and any $Q \in V_1 \otimes \cdots V_q$ that

$$\widetilde{T}(P, Q) = P\langle p \rangle T \langle q \rangle Q. \tag{7.4.15}$$

Then $\widetilde{T}$ is bilinear, so $\widetilde{T} \in (U_1 \otimes \cdots \otimes U_p) \otimes (V_1 \otimes \cdots \otimes V_q)$. From (7.4.15), $(\widetilde{a_j T_j}) = a_j \widetilde{T}_j$, so $T \mapsto \widetilde{T}$ is a linear mapping of $U_1 \otimes \cdots \otimes U_p \otimes V_1 \otimes \cdots \otimes V_q$ into $(U_1 \otimes \cdots \otimes U_p) \otimes (V_1 \otimes \cdots \otimes V_q)$. We claim this mapping is bijective. To see that it is injective, suppose T_1 , $T_2 \in U_1 \otimes \cdots \otimes U_p \otimes V_1 \otimes \cdots \otimes V_q$ and $\widetilde{T}_1 = \widetilde{T}_2$. Then for every $P \in U_1 \otimes \cdots \otimes U_p$ and $Q \in V_1 \otimes \cdots \otimes V_q$ we have $P\langle p\rangle T_1 \langle q\rangle Q = P\langle p\rangle T_2 \langle q\rangle Q$, so $P\langle p\rangle (T_1\langle q\rangle Q - T_2\langle q\rangle Q = 0$ so $P\langle p\rangle[(T_1 - T_2)\langle q\rangle Q] = 0$. Hence, $(T_1 - T_2)\langle q\rangle Q = 0$ for all $Q \in V_1 \otimes \cdots \otimes V_q$. Hence, $T_1 - T_2 = 0$. To see that $T \mapsto \widetilde{T}$ is a surjection, suppose $\Pi \in (U_1 \otimes \cdots \otimes U_p) \otimes (V_1 \otimes \cdots \otimes V_q)$. Define $T \in U_1 \otimes \cdots \otimes U_p \otimes V_1 \otimes \cdots \otimes V_q$ by requiring $T(\vec{u}_1, \cdots, \vec{u}_p, \vec{v}_1, \cdots \vec{v}_q) = \Pi(\vec{u}_1 \cdots \vec{u}_p, \vec{v}_1, \cdots \vec{v}_q)$. Then $P\langle p\rangle T\langle q\rangle Q = \Pi(P, Q)$ if P and Q are polyads. Hence $P\langle p\rangle T\langle q\rangle Q = \Pi(P, Q)$ for all $P \in U_1 \otimes \cdots \otimes U_p$ and $Q \in V_1 \otimes \cdots \otimes V_q$. By (7.4.15), $\Pi = \widetilde{T}$. Thus $T \mapsto \Pi$, and the mapping $T \mapsto \widetilde{T}$ is a surjection of $U_1 \otimes \cdots \otimes U_p \otimes V_1 \otimes \cdots \otimes V_q$ onto $(U_1 \otimes \cdots \otimes U_p) \otimes (V_1 \otimes \cdots \otimes V_q)$.

Application 4: Linear mappings as second order tensors.

Definition 7.4.22 Suppose U and V are Euclidean vector spaces and $L \in \mathcal{L}(U \to V)$. Define $\overleftrightarrow{L} \in \mathcal{F}(U \times V \to \mathcal{R})$ by

$$\overleftrightarrow{L}(\vec{u}, \vec{v}) = L(\vec{u}) \cdot \vec{v}. \tag{7.4.16}$$

Then $\overleftrightarrow{L}(\cdot, \vec{v})$ and $\overleftrightarrow{L}(\vec{u}, \cdot)$ are linear, so $\overleftrightarrow{L} \in \mathcal{M}(U \times V \to \mathcal{R})$.

That is,

$$\overleftrightarrow{L} \in U \otimes V. \tag{7.4.17}$$

Thus, (7.4.16) can be written in the equivalent form

$$\vec{u} \cdot \overleftrightarrow{L} \cdot \vec{v} = L(\vec{u}) \cdot \vec{v}. \tag{7.4.18}$$

For a fixed $\vec{u} \in U$, (7.4.18) is true for all $\vec{v} \in V$. Hence, since $\vec{u} \cdot \vec{L}$ and $L(\vec{u})$ are both in V, we have

$$\vec{u} \cdot \overleftrightarrow{L} = L(\vec{u}). \tag{7.4.19}$$

Let $\phi : \mathcal{L}(U \to V) \to U \otimes V$ be defined by $\phi(L) = \overleftrightarrow{L}$. If we know $\overleftrightarrow{L} = \phi(L)$, we can recover L from (7.4.19), so ϕ is an injection. We claim ϕ is also a surjection. That is, every $T \in U \otimes V$ is $\phi(L)$ for at least one $L \in \mathcal{L}(U \to V)$. Suppose T is given and define L by $L(\vec{u}) = \vec{u} \cdot T$. Then $\overleftrightarrow{L}(\vec{u}, \vec{v}) = \vec{u} \cdot \overleftrightarrow{L} \cdot \vec{v} = L(\vec{u}) \cdot \vec{v} = \vec{u} \cdot T \cdot \vec{v} = T(\vec{u}, \vec{v})$, so $T = \overleftrightarrow{L} = \phi(L)$.

We claim that $\phi : \mathcal{L}(U \to V) \to U \otimes V$ is also linear. If $L_1, \cdots, L_n \in \mathcal{L}(U \to V)$ and $a_1, \cdots, a_n \in \mathcal{R}$, we claim

$$\overleftrightarrow{(a_i L_i)} = a_i \overleftrightarrow{L}_i . \tag{7.4.20}$$

The proof is this: for any $\vec{u} \in U$ and $\vec{v} \in V$,

$$\begin{aligned}
\vec{u} \cdot \overleftrightarrow{(a_i L_i)} \cdot \vec{v} &= [(a_i L_i)(\vec{u})] \cdot \vec{v} = \{a_i [L_i(\vec{u})]\} \cdot \vec{v} \\
&= a_i [L_i(\vec{u}) \cdot \vec{v}] = a_i(\vec{u} \cdot \overleftrightarrow{L}_i \cdot \vec{v}) \\
&= \vec{u} \cdot \left(a_i \overleftrightarrow{L}_i\right) \cdot \vec{v}.
\end{aligned}$$

Thus $\overleftrightarrow{(a_i L_i)}(\vec{u}, \vec{v}) = (a_i \overleftrightarrow{L})(\vec{u}, \vec{v})$ for all $\vec{u} \in U$, $\vec{v} \in V$. Hence (7.4.20).

We have now shown that ϕ is a linear bijection between the two vector spaces $\mathcal{L}$ $(U \to V)$ and $U \otimes V$. We can use ϕ to regard any linear mapping $L : U \to V$ as a tensor $\overleftrightarrow{L} \in U \otimes V$ and vice-versa. The process of taking linear combinations can be done either to the tensors or to the linear mappings, so confusing tensors with linear mappings does no harm to linear combinations.

Linear mappings can also be "multiplied" by composition. If $K \in \mathcal{L}(U \to V)$ and $L \in \mathcal{L}(V \to W)$ then $L \circ K \in \mathcal{L}(U \to W)$. Our identification of tensors with linear mappings almost preserves this multiplication. We have

$$\overleftrightarrow{L \circ K} = \overleftrightarrow{K} \cdot \overleftrightarrow{L} . \tag{7.4.21}$$

The proof is simple. For any $\vec{u} \in U$,

$$\begin{aligned}
\vec{u} \cdot (\overleftrightarrow{L \circ K}) &= (L \circ K)(\vec{u}) = L[K(\vec{u})] \\
&= K(\vec{u}) \cdot \overleftrightarrow{L} = (\vec{u} \cdot \vec{K}) \cdot \overleftrightarrow{L} \\
&= \vec{u} \cdot (\vec{K} \cdot \overleftrightarrow{L}) \qquad \text{by (7.3.1).}
\end{aligned}$$

Since this is true for all $\vec{u} \in U$, corollary (7.4.18) gives (7.4.21).

If $I_U : U \to U$ is the identity operator on U (that is, $I_U(\vec{u}) = \vec{u}$ for all $\vec{u} \in U$) then for any $\vec{u}_1, \vec{u}_2 \in U$,

$$\overleftrightarrow{I}_U (\vec{u}_1, \vec{u}_2) = I_U(\vec{u}_1) \cdot \vec{u}_2 = \vec{u}_1 \cdot \vec{u}_2. \tag{7.4.22}$$

That is, $\overleftrightarrow{I}_U$ is what we have already called the identity tensor on U; see example 5.6.3. The origin of that name is now clear. And (7.4.22) can also be written

$$\vec{u}_1 \cdot \overleftrightarrow{I}_U \cdot \vec{u}_2 = \vec{u}_1 \cdot \vec{u}_2. \tag{7.4.23}$$

This is true for all $\vec{u}_1$ and $\vec{u}_2 \in U$, so we have for any $\vec{u} \in U$

$$\vec{u} \cdot \overleftrightarrow{I}_U = \overleftrightarrow{I}_U \cdot \vec{u} = \vec{u}. \tag{7.4.24}$$

If $L : U \to V$ then $L \circ I_U = I_V \circ L = L$. It follows then from (7.4.20) that

$$\overleftrightarrow{I}_U \cdot \overleftrightarrow{L} = \overleftrightarrow{L} \cdot \overleftrightarrow{I}_V = \overleftrightarrow{L}. \tag{7.4.25}$$

Actually, a more general result is true. If U, V, W_1, $\cdots$, W_q are Euclidean spaces and $T \in U \otimes W_1 \otimes \cdots \otimes W_q \otimes V$ then

$$\overleftrightarrow{I}_U \cdot T = T \cdot \overleftrightarrow{I}_V = T. \tag{7.4.26}$$

The proof follows the usual lines. Since all three expressions in (7.4.26) are linear in T, it suffices to prove (7.4.26) for polyads T. But for T a polyad, (7.4.26) follows immediately from associativity and (7.4.24).

Thus, $\overleftrightarrow{I}_U$ acts like a multiplicative identity. Then *some* second order tensors can have inverses.

Definition 7.4.23 If $L \in \mathcal{L}(U \to V)$ has an inverse $L^{-1} \in \mathcal{L}(V \to U)$, then we *define* $\left(\overleftrightarrow{L}\right)^{-1}$ to be $\overleftrightarrow{L^{-1}}$.

Thus, using (7.4.20), we conclude

$$\overleftrightarrow{L^{-1}} \cdot \overleftrightarrow{L} = \overleftrightarrow{I}_V, \quad \overleftrightarrow{L} \cdot \overleftrightarrow{L^{-1}} = \overleftrightarrow{I}_U. \tag{7.4.27}$$

Since the tensors in $U \otimes V$ and the linear mappings in $\mathcal{L}(U \to V)$ can be thought of as essentially the same objects, any concept defined for one can be defined immediately for the other. Definition 7.4.23 is an example of this process. Other examples are below.

Definition 7.4.24 If $\overleftrightarrow{L} \in V \otimes V$ then $\det \overleftrightarrow{L} := \det L$.

From definition 7.4.23 and 7.4.24 it is apparent that for any $\overleftrightarrow{L} \in V \otimes V$, $\overleftrightarrow{L}^{-1}$ exists iff $\det \overleftrightarrow{L} \neq 0$.

Definition 7.4.25 If $\overleftrightarrow{L} \in U \otimes V$ then $\left(\overleftrightarrow{L}\right)^T := \overleftrightarrow{L^T}$. Thus $\overleftrightarrow{L^T} \in V \otimes U$.

Here L^T is the transpose of $L : U \to V$, as defined on page D-24. From the definition of L^T we have $\left(\overleftrightarrow{L}\right)^T (\vec{v}, \vec{u}) = \overleftrightarrow{L^T} (\vec{v}, \vec{u}) = L^T(\vec{v}) \cdot \vec{u} = \vec{v} \cdot L(\vec{u}) = L(\vec{u}) \cdot \vec{v} = \overleftrightarrow{L} (\vec{u}, \vec{v})$. Thus

$$\overleftrightarrow{L^T} = (12) \overleftrightarrow{L} . \tag{7.4.28}$$

It is also true that for any $(\vec{u}, \vec{v}) \in U \times V$, $\vec{v} \cdot \overleftrightarrow{L^T} \cdot \vec{u} = \overleftrightarrow{L^T} (\vec{v} \cdot \vec{u}) = \overleftrightarrow{L} (\vec{u}, \vec{v}) = \vec{u} \cdot \overleftrightarrow{L} \cdot \vec{v}$. Thus $\vec{v} \cdot \left[\overleftrightarrow{L^T} \cdot \vec{u} - \vec{u} \cdot \overleftrightarrow{L}\right] = 0$. This is true for that for any $\vec{v} \in V$ so for all $\vec{u} \in U$

$$\overleftrightarrow{L^T} \cdot \vec{u} = \vec{u} \cdot \overleftrightarrow{L} . \tag{7.4.29}$$

Similarly, $\left[\vec{v} \cdot \overleftrightarrow{L^T} - \overleftrightarrow{L} \cdot \vec{v}\right] \cdot \vec{u} = 0$ for all $\vec{u} \in U$ so if $\vec{v} \in V$

$$\vec{v} \cdot \overleftrightarrow{L}^T = \overleftrightarrow{L} \cdot \vec{v}. \tag{7.4.30}$$

Application 5: Linear mappings as higher order tensors.

Suppose $U_1, \cdots, U_p, V_1, \cdots V_q$ are Euclidean vector spaces and $L \in \mathcal{L}(U_1 \otimes \cdots \otimes U_p \to V_1 \otimes \cdots V_q)$. Both $U_1 \otimes \cdots \otimes U_p$ and $V_1 \otimes \cdots \otimes V_q$ are Euclidean vectors spaces, with dot products $\langle p \rangle$ and $\langle q \rangle$. Therefore, L corresponds to a second order tensor $\overleftrightarrow{L}$ in $(U_1 \otimes \cdots \otimes U_p) \otimes (V_1 \otimes \cdots \otimes V_q)$, exactly as in (7.4.16). That is if $P \in U_1 \otimes \cdots \otimes U_p$ and $Q \in V_1 \otimes \cdots \otimes V_q$ then

$$\overleftrightarrow{L} (P, Q) = L(P)\langle q \rangle Q$$

or

$$P \langle q \rangle \overleftrightarrow{L} \langle q \rangle Q = L(P) \langle q \rangle Q. \tag{7.4.31}$$

Here $\langle p \rangle$ and $\langle q \rangle$ are to be interpreted as the dot products on $U_1 \otimes \cdots \otimes U_p$ and $V_1 \otimes \cdots \otimes V_q$, and $\overleftrightarrow{L}$ is a second order tensor in the space $(U_1 \otimes \cdots \otimes U_p) \otimes (V_1 \otimes \cdots \otimes V_q)$, which is the tensor product of two Euclidean spaces.

However, as we saw on page 61, $\overleftrightarrow{L}$ can also be interpreted as a tensor of order $p+q$ in $U_1 \otimes \cdots \otimes U_p \otimes V_1 \otimes \cdots \otimes V_q$. Equation (7.4.31) continues to hold, but now $\langle p \rangle$ and $\langle q \rangle$ are interpreted as generalized dot products, and $\overleftrightarrow{L} \in U_1 \otimes \cdots \otimes U_p \otimes V_1 \otimes \cdots \otimes V_q$. Therefore, the linear mappings $L \in \mathcal{L}(U_1 \otimes \cdots \otimes U_p \to V_1 \otimes \cdots \otimes V_q)$ can be viewed as tensors $\overleftrightarrow{L} \in U_1 \otimes \cdots \otimes U_p \otimes V_1 \otimes \cdots \otimes V_q$. If we know L, we find $\overleftrightarrow{L}$ from (7.4.31) by taking P and Q to be arbitrary polyads. If we know $\overleftrightarrow{L} \in U_1 \otimes \cdots \otimes U_p \otimes V_1 \otimes \cdots \otimes V_q$, then we know it in $(U_1 \otimes \cdots \otimes U_p) \otimes (V_1 \otimes \cdots \otimes V_q)$ so we find L by appealing to (7.4.19). That is, for any $P \in U_1 \otimes \cdots \otimes U_p$,

$$L(P) = P\langle p \rangle \overleftrightarrow{L} . \tag{7.4.32}$$

The generalization from application 4 to application 5 is particularly useful in continuum mechanics.

Chapter 8

How to Rotate Tensors (and why)

8.1 Tensor products of linear mappings

For $j \in \{1, \cdots, q\}$, suppose V_j and W_j are Euclidean vector spaces and $L_j \in \mathcal{L}(V_j \to W_j)$. We would like to define a linear mapping

$$L_1 \otimes \cdots \otimes L_q : V_1 \otimes \cdots \otimes V_q \to W_1 \otimes \cdots \otimes W_q$$

by requiring that for any polyad $\vec{v}_1 \cdots \vec{v}_q \in V_1 \otimes \cdots \otimes V_q$ we have

$$(L_1 \otimes \cdots \otimes L_q)\,(\vec{v}_1 \cdots \vec{v}_q) = L_1\,(\vec{v}_1) \cdots L_q\,(\vec{v}_q)\,. \tag{8.1.1}$$

The lifting theorem will show that there is exactly one such $L_1 \otimes \cdots \otimes L_q$. To find it, we invoke that theorem. We define

$$M : V_1 \times \cdots \times V_q \to W_1 \otimes \cdots \otimes W_q$$

by requiring for any $(\vec{v}_1, \cdots, \vec{v}_q) \in V_1 \times \cdots \times V_q$ that

$$M\,(\vec{v}_1, \cdots, \vec{v}_q) = L_1(\vec{v}_1) \cdots L_q(\vec{v}_q).$$

Since $L_1, \cdots, L_q$ are linear, M is multilinear, so we can use the lifting theorem with M as above and $W = W_1 \otimes \cdots \otimes W_q$. The picture is shown in 8.1. Then we define $L_1 \otimes \cdots \otimes L_q$ to be the $M_\otimes$ provided by the lifting theorem. With this choice, $L_1 \otimes \cdots \otimes L_q$ is linear and satisfies (8.1.1).

$$
\begin{tikzcd}
V_1 \otimes \cdots \otimes V_q \arrow[r, "M_{\otimes}"] & W_1 \otimes \cdots \otimes W_q \\
V_1 \times \cdots \times V_q \arrow[u, "\otimes"] \arrow[ur, "M"'] &
\end{tikzcd}
$$

Figure 8.1:

Remark 8.1.37 Suppose $\left(\vec{b}_1^{(j)}, \cdots, \vec{b}_{n_j}^{(j)}\right)$ is an ordered basis for V_j and $Q = Q^{i_1 \cdots i_q} \vec{b}_{i_1}^{(1)} \cdots \vec{b}_{i_q}^{(q)}$ $\in V_1 \otimes \cdots \otimes V_q$. Then

$$(L_1 \otimes \cdots \otimes L_q)(Q) = Q^{i_1 \cdots i_q} \left[L_1 \left(\vec{b}_{i_1}^{(1)}\right) \cdots L_q \left(\vec{b}_{i_q}^{(q)}\right)\right]. \tag{8.1.2}$$

Proof:

$L_1 \otimes \cdots \otimes L_q$ is linear, so $(L_1 \otimes \cdots \otimes L_q)(Q) = Q^{i_1 \cdots i_q} \left[(L_1 \otimes \cdots \otimes L_q) \left(\vec{b}_{i_1}^{(1)} \cdots \vec{b}_{i_q}^{(q)}\right)\right]$.
Now use (8.1.1).

Remark 8.1.38 Suppose for $j \in \{1, \cdots, q\}$ that U_j, V_j, W_j are Euclidean and $K_j \in \mathcal{L}(V_j \to W_j)$ and $L_j \in \mathcal{L}(U_j \to V_j)$. Then

$$(K_1 \otimes \cdots \otimes K_q) \circ (L_1 \otimes \cdots \otimes L_q) = (K_1 \circ L_1) \otimes \cdots \otimes (K_q \circ L_q). \tag{8.1.3}$$

Proof:

We want to show for every $P \in U_1 \otimes \cdots \otimes U_q$ that

$$[(K_1 \otimes \cdots \otimes K_q) \circ (L_1 \otimes \cdots \otimes L_q)](P) = [(K_1 \circ L_1) \otimes \cdots \otimes (K_q \circ L_q)](P).$$

By linearity it suffices to prove this when P is any polyad $\vec{u}_1 \cdots \vec{u}_q$. But

$$
\begin{aligned}
&[(K_1 \otimes \cdots \otimes K_q) \circ (L_1 \otimes \cdots \otimes L_q)](\vec{u}_1 \cdots \vec{u}_q) \\
&\quad = (K_1 \otimes \cdots \otimes K_q)[(L_1 \otimes \cdots \otimes L_q)(\vec{u}_1 \cdots \vec{u}_q)] \\
&\quad = (K_1 \otimes \cdots \otimes K_q)[L_1(\vec{u}_1) \cdots L_q(\vec{u}_q)] \\
&\quad = K_1[L_1(u_1)]\, K_2[L_2(u_2)] \cdots K_q[L_q(u_q)]\, [(K_1 \circ L_1)(\vec{u})]\, [(K_2 \circ L_2)(\vec{u}_2)] \\
&\qquad \cdots [(K_q \circ L_q)(\vec{u}_q)] \\
&\quad = [(K_1 \circ L_1) \otimes \cdots \otimes (K_q \circ L_q)](\vec{u}_1 \cdots \vec{u}_q).
\end{aligned}
$$

Another way to try to understand $L_1 \otimes \cdots \otimes L_q$ is to exhibit explicitly for any $P \in V_1 \otimes \cdots \otimes V_q$ the tensor $(L_1 \otimes \cdots \otimes L_q)(P) \in W_1 \otimes \cdots \otimes W_q$ by evaluating it at any $(\vec{w}_1, \cdots, \vec{w}_q) \in W_1 \times \cdots \times W_q$. We have

Remark 8.1.39 Suppose that for $j \in \{1, \cdots, q\}$ we have $L_j \in \mathcal{L}(V_j \to W_j)$ for Euclidean spaces V_j and W_j. Suppose $P \in V_1 \otimes \cdots \otimes V_q$ and $(\vec{w}_1, \cdots, \vec{w}_q) \in W_1 \times \cdots \times W_q$. Then

$$[(L_1 \otimes \cdots \otimes L_q)(P)]\,(\vec{w}_1, \cdots, \vec{w}_q) = P\left[L_1^T(\vec{w}_1), \cdots, L_q^T(\vec{w}_q)\right]. \tag{8.1.4}$$

Proof:

Both sides of this equation are linear in P, so it suffices to prove the equation when P is a polyad, $P = \vec{v}_1 \cdots \vec{v}_q$. In that case,

$$\begin{aligned}
[(L_1 \otimes \cdots \otimes L_q)(P)]\,(\vec{w}_1, \cdots, \vec{w}_q) &= [L_1(\vec{v}_1) \cdots L_q(\vec{v}_q)]\,(\vec{w}_1, \cdots, \vec{w}_q) \\
&= [L_1(\vec{v}_1) \cdot \vec{w}_1] \cdots [L_q(\vec{v}_q) \cdot \vec{w}_q] \\
&= \left[\vec{v}_1 \cdot L_1^T(\vec{w}_1)\right] \cdots \left[\vec{v}_q \cdot L_q^T(\vec{w}_q)\right] \\
&= P\left[L_1^T(\vec{w}_1), \cdots, L_q^T(\vec{w}_q)\right].
\end{aligned}$$

Finally, it will be useful to understand how permutations affect tensor products of mappings. To do this we need

Lemma 8.1.22 Suppose $V_1, \cdots, V_q$ are Euclidean spaces, $\vec{v}_j \in V_j$ for $j \in \{1, \cdots, q\}$, and $\pi \in S_q$. Then

$$\pi\,(\vec{v}_1 \cdots \vec{v}_q) = \vec{v}_{\pi^{-1}_{(1)}} \cdots \vec{v}_{\pi^{-1}_{(q)}}. \tag{8.1.5}$$

Proof:

From the definition (1.4.10),

$$[\pi\,(\vec{v}_1 \cdots \vec{v}_q)]\left(\vec{w}_{\pi^{-1}(1)}, \cdots, \vec{w}_{\pi^{-1}(q)}\right) = (\vec{v}_1 \cdots \vec{v}_q)\,(\vec{w}_1, \cdots, \vec{w}_q)$$

$$\begin{aligned}
&= (\vec{v}_1 \cdot \vec{w}_1) \cdots (\vec{v}_q \cdot \vec{w}_q) \\
&= \left(\vec{v}_{\pi^{-1}(1)} \cdot \vec{w}_{\pi^{-1}(1)}\right) \cdots \left(\vec{v}_{\pi^{-1}}(q) \cdot \vec{w}_{\pi^{-1}(q)}\right) \\
&= \left(\vec{v}_{\pi^{-1}(1)} \cdots \vec{v}_{\pi^{-1}(q)}\right)\left(\vec{w}_{\pi^{-1}(1)}, \cdots, \vec{w}_{\pi^{-1}(q)}\right)
\end{aligned}$$

for any $(\vec{w}_1, \cdots, \vec{w}_q) \in V_1 \times \cdots \times V_q$. QED.

Now we can prove

Remark 8.1.40 For $j \in \{1, \cdots, q\}$, suppose $L_j \in \mathcal{L}(V_j \to W_j)$, and suppose $\pi \in S_q$. Then

$$\pi \left(L_1 \otimes \cdots \otimes L_q\right) = \left(L_{\pi^{-1}(1)} \otimes \cdots \otimes L_{\pi^{-1}(q)}\right) \pi. \tag{8.1.6}$$

Proof:

The domain of $L_1 \otimes \cdots \otimes L_q$ is $V_1 \otimes \cdots \otimes V_q$. We want to prove that if $P \in V_1 \otimes \cdots \otimes V_q$ then $[\pi(L_1 \otimes \cdots \otimes L_q)](P) = [(L_{\pi^{-1}(1)} \otimes \cdots \otimes L_{\pi^{-1}(q)}\pi)](P)$, i.e., $\pi[(L_1 \otimes \cdots \otimes L_q)](P) = (L_{\pi^{-1}(1)} \otimes \cdots \otimes L_{\pi^{-1}(q)})(\pi P)$. Both sides of the last equation are linear in P, so it suffices to prove that equation when P is a polyad, say $P = \vec{v}_1 \cdots \vec{v}_q$. In this case

$$\pi \left[\left(L_1 \otimes \cdots \otimes L_q\right)\left(\vec{v}_1 \cdots \vec{v}_q\right)\right] = \pi \left[L_1(\vec{v}_1) \cdots L_q(\vec{v}_q)\right]$$

$$\begin{aligned}
&= L_{\pi^{-1}(1)}\left(\vec{v}_{\pi^{-1}(1)}\right) \cdots L_{\pi^{-1}(q)}\left(\vec{v}_{\pi^{-1}(q)}\right) \text{ by 8.1.5} \\
&= \left(L_{\pi^{-1}(1)} \otimes \cdots \otimes L_{\pi^{-1}(q)}\right)\left(\vec{v}_{\pi^{-1}(1)} \cdots \vec{v}_{\pi^{-1}(q)}\right) \\
&= \left(L_{\pi^{-1}(1)} \otimes \cdots \otimes L_{\pi^{-1}(q)}\right)\left[\pi\left(\vec{v}_1 \cdots \vec{v}_q\right)\right] \text{ by 8.1.5}
\end{aligned}$$

QED.

Special Case:

Suppose $V_1 = \cdots = V_q = V$ and $W_1 = \cdots = W_q = W$ and $L_1 = \cdots = L_q = L$. Then we write $L_1 \otimes \cdots \otimes L_q$ as $\otimes^q L$. Then

$$\otimes^q L \in \mathcal{L}\left(\otimes^q V \to \otimes^q W\right), \tag{8.1.7}$$

and $\otimes^q L$ is defined by requiring for all $(\vec{v}_1, \cdots, \vec{v}_q) \in x^q V$ that

$$\otimes^q L(\vec{v}_1 \cdots \vec{v}_q) = L(\vec{v}_1) \cdots L(\vec{v}_q). \tag{8.1.8}$$

If $P \in \otimes^q V$, (8.1.8) leads to the figure of speech that $(\otimes^q L)(P)$ is obtained by applying L to P. (Really $\otimes^q L$ is applied to P).

If $K \in \mathcal{L}(V \to W)$ and $L \in \mathcal{L}(U \to V)$ then from (8.1.3)

$$(\otimes^q K) \circ (\otimes^q L) = \otimes^q (K \circ L). \tag{8.1.9}$$

If $L \in \mathcal{L}(V \to W)$ and $P \in \otimes^q V$ and $(\vec{w}_1, \cdots, \vec{w}_q) \in x^q W$ then from (8.1.4)

$$[(\otimes^q L)(P)] (\vec{w}_1, \cdots, \vec{w}_q) = P \left[L^T(\vec{w}_1), \cdots, L^T(\vec{w}_q)\right]. \tag{8.1.10}$$

If $L \in \mathcal{L}(V \to W)$ and $\pi \in S_q$ then from (8.1.6)

$$\pi(\otimes^q L) = (\otimes^q L)\pi. \tag{8.1.11}$$

8.2 Applying rotations and reflections to tensors

Recall that $L \in \mathcal{L}(V \to V)$ is "orthogonal" if $L^T = L^{-1}$, and "proper orthogonal" if also $\det L = +1$ rather than -1. Recall the definitions on page D-25:

$$\begin{aligned} \Theta(V) &= \text{set of all orthogonal operators on } V \\ \Theta^+(V) &= \text{set of all proper orthogonal operators on } V. \end{aligned}$$

Recall that according to corollary PP-12, every proper orthogonal operator is a rotation. As remarked on page PP-30, if L is improper orthogonal ($\det L = -1$) then L is the product of a rotation and a reflection.

If L is an *orthogonal* operator on V and $Q \in \otimes^q V$, then two more facts about $(\otimes^q L)(Q)$ make it even more reasonable to think of that tensor as the result of applying L to Q. These facts are remark 8.2.41 and its corollary.

Remark 8.2.41 If $L \in \Theta(V)$, applying L to tensors over V preserves their generalized dot products. More precisely suppose p, q, r are any non-negative integers. Suppose $P \in \otimes^{p+q} V$, $R \in \otimes^{q+r} V$, and $L \in \Theta(V)$. Then

$$(\otimes^{p+r} L)(P\langle q\rangle R) = \left[(\otimes^{p+q} L)(P)\right] \langle q\rangle \left[(\otimes^{q+r} L)(R)\right]. \tag{8.2.1}$$

Proof:

Both sides of this equation are linear in P and R, so it suffices to prove the equation when P and R are polyads, say $P = \vec{u}_1 \cdots \vec{u}_p \vec{x}_1 \cdots \vec{x}_q$ and $R = \vec{y}_1 \cdots \vec{y}_q \vec{v}_1 \cdots \vec{v}_r$. Then $(\otimes^{p+r} L)(P\langle q \rangle R) = (\otimes^{p+r} L)\left[(\vec{x}_1 \cdot \vec{y}_1) \cdots (\vec{x}_q \cdot \vec{y}_q) \vec{u}_1 \cdots \vec{u}_p \vec{v}_1 \cdots \vec{v}_r\right]$
$= (\vec{x}_1 \cdot \vec{y}_1) \cdots (\vec{x}_q \cdot \vec{y}_q)(\otimes^{p+r} L)(\vec{u}_1 \cdots \vec{u}_p \vec{v}_1 \cdots \vec{v}_r)$
$= (\vec{x}_1 \cdot \vec{y}_1) \cdots (\vec{x}_q \cdot \vec{x}_q)\left[L(\vec{u}_1) \cdots L(\vec{u}_p) L(\vec{v}_1) \cdots L(\vec{v}_r)\right]$.

So far we have not used $L \in \Theta(V)$. Now we use it to set $\vec{x} \cdot \vec{y} = L(\vec{x}) \cdot L(\vec{y})$. Therefore the above polyad is $(\otimes^{p+r} L)(P\langle q \rangle R) =$

$$
\begin{aligned}
&= [L(\vec{x}_1) \cdot L(\vec{y}_1)] \cdots [L(\vec{x}_q) \cdot L(\vec{y}_q)]\,[L(\vec{u}_1) \cdots L(\vec{u}_p) L(\vec{v}_1) \cdots L(\vec{v}_r)] \\
&= [L(\vec{u}_1) \cdots L(\vec{u}_p) L(\vec{x}_1) \cdots L(\vec{x}_q)]\,\langle q \rangle\,[L(\vec{y}_1) \cdots L(\vec{y}_q) L(\vec{v}_1) \cdots L(\vec{v}_r)] \\
&= \left[(\otimes^{p+q} L)(\vec{u}_1 \cdots \vec{u}_p \vec{x}_1 \cdots \vec{x}_q)\right] \langle q \rangle \left[(\otimes^{q+r} L)(\vec{y}_1 \cdots \vec{y}_q \vec{v}_1 \cdots \vec{v}_r)\right] \\
&= \left[(\otimes^{p+q} L)(P)\right] \langle q \rangle \left[(\otimes^{q+r} L)(R)\right].
\end{aligned}
$$

QED

Corollary 8.2.19 The value of a rotated q-tensor at a rotated q-tuple of vectors is the value of the original tensor at the original q-tuple of vectors. The same is true for reflections. in general, if $L \in \Theta(V)$ and $Q \in \otimes^q V$ and $(\vec{v}_1 \cdots, \vec{v}_q) \in x^q V$ then

$$[(\otimes^q L)(Q)]\,[L(\vec{v}_1), \cdots, L(\vec{v}_q)] = Q(\vec{v}_1, \cdots, \vec{v}_q).$$

Proof:

$$
\begin{aligned}
&[(\otimes^q L)(Q)]\,[L(\vec{v}_1), \cdots, L(\vec{v}_q)] \\
&\qquad = [(\otimes^q L)(Q)]\,\langle q \rangle\,[L(\vec{v}_1), L(\vec{v}_2) \cdots L(\vec{v}_q)] \\
&\qquad = [(\otimes^q L)(Q)]\,\langle q \rangle\,[(\otimes^q L)(\vec{v}_1 \vec{v}_2 \cdots \vec{v}_q)] \\
&\qquad = Q\langle q \rangle (\vec{v}_1 \cdots \vec{v}_q) \text{ by remark 8.2.41} \\
&\qquad = Q(\vec{v}_1, \cdots, \vec{v}_q).
\end{aligned}
$$

8.3 Physical applications

Some physical properties of physical systems are tensors. For example, Ohm's law in a crystal says that if an electric field $\vec{E}$ is applied, the resulting current density $\vec{J}$ depends linearly on $\vec{E}$. This means that there is a second-order tensor $\overleftrightarrow{K}$ such that $\vec{J} = \vec{E} \cdot \overleftrightarrow{K}$. This $\overleftrightarrow{K}$ is the "conductivity tensor" of the crystal. In a liquid, $\vec{J} = \vec{E}K$ where K is a scalar, so $\overleftrightarrow{K} = K \overleftrightarrow{I}$, but in a crystal all that can be said in general is that if there is no magnetic field, $\overleftrightarrow{K}$ is symmetric ($\overleftrightarrow{K}^T = \overleftrightarrow{K}$). This is Onsager's theorem for dissipative thermodynamics.

A second example is an elastic crystal. As we will see later, its local state of deformation is described by a symmetric second order tensor $\overleftrightarrow{\sigma}$, the strain tensor. Its local state of stress is described by another symmetric second order tensor $\overleftrightarrow{S}$, the stress tensor. For small $\overleftrightarrow{\sigma}$ (i.e. $\| \overleftrightarrow{\sigma} \| << 10^{-2}$) $\overleftrightarrow{S}$ is a linear function of $\overleftrightarrow{\sigma}$, so there is a tensor of fourth order over ordinary 3-space, the elasticity tensor E, such that $\overleftrightarrow{S} = \overleftrightarrow{\sigma} \langle 2 \rangle E$. Thermodynamics assures that $(12)E = (34)E = (13)(24)E = E$, but E has no other symmetries in general so the dimenson of the space of all possible elasticity tensors is 21. (We will discuss all this in detail later.)

A third example is the second moment tensor $\overleftrightarrow{M}$ of a collection of mass points. If there is a mass m_v at position $\vec{r}_v$ for $v = 1, \cdots, N$, then

$$\overleftrightarrow{M} := \sum_{v=1}^{N} m_v \vec{r}_v \vec{r}_v. \tag{8.3.1}$$

This is the underlying tensor from which the inertia tensor and gravitational second degree harmonic tensor of a solid body are calculated.

If V is real 3-space and $L \in \mathcal{L}(V \to V)$, then "subjecting a physical system to the mapping L" means that we move all the mass points (atoms) which make up the system, so that a mass point which was originally at $\vec{r}_v$ is moved to the new location $L(\vec{r}_v) = \vec{r}_v \cdot \overleftrightarrow{L}$. If $L \in \Theta^+(V)$, L subjects the body to a rigid rotation. If $L \in \Theta(V)$ but $\det L = -1$, L subjects the body to a rigid rotation and also a reflection in some plane. If $L \in \Theta(V)$ but $\det L > 0$ then L is the product of a rigid rotation and a symmetric linear operator. The latter stretches the body by different amounts in different directions. If $\det L < 0$, a

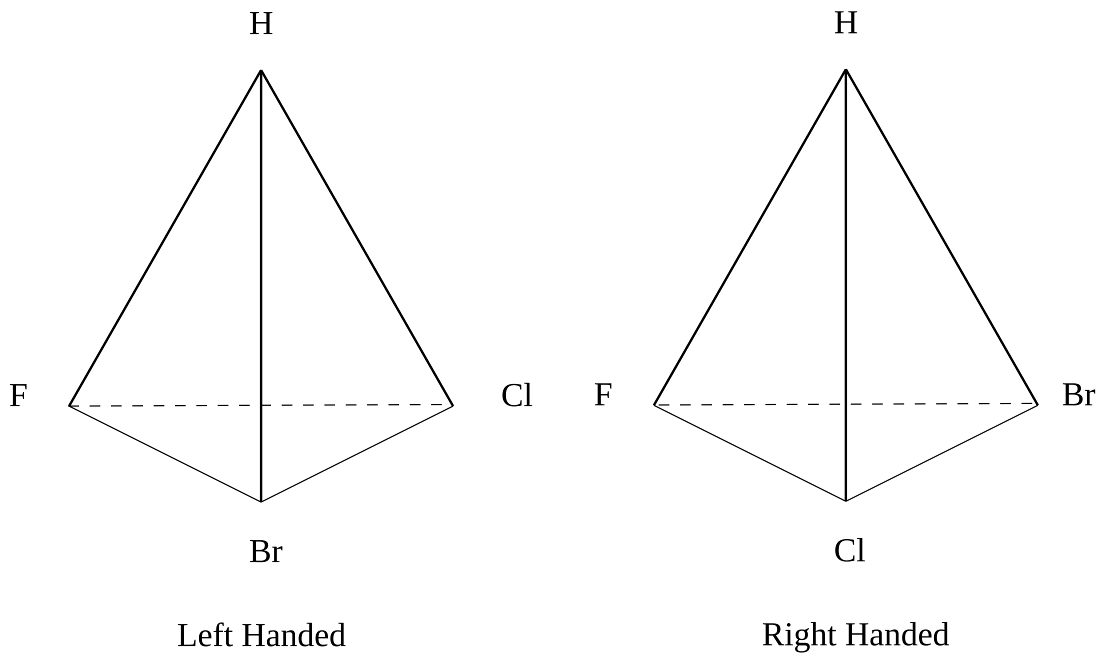


Figure 8.2:

reflection also occurs.

It is actually possible to imagine subjecting a real body to a reflection if we allow ourselves to disassemble and reassemble the body atom-by-atom. For example, iceland spar (calcite) crystals come in right-handed or left-handed version. Liquids can also be reflected. For most, reflection has no effect because the distribution of molecules in the liquid has no preferred directions and the molecules either have no right- or left-handedness, or the right- and left-handed versions are present in equal amounts. A liquid which is affected by reflection is bromochlorofluoromethane, C H F Cl Br. At one atmosphere of pressure, it is liquid between $-115°$C and $36°$C. the molecule is a tetrahedron with H, F, Cl, Br at the four vertices and C at the center. It can come in two forms, left handed or right handed. To reflect a pot of left-handed C H F Cl Br means to replace it by a pot of right-handed C H F Cl Br.

If we subject a system of mass points m_v at positions $\vec{r}_v$ to the mapping $L \in \mathcal{L}(V \to V)$, where V is real 3-space, then the second moment tensor $\overleftrightarrow{M}$ of that system (see (8.3.1)) changes to $\otimes^2 L(\overleftrightarrow{M})$. To see this, note that the second moment tensor of the deformed system is

$$\begin{aligned}\overleftrightarrow{M}_L &= \sum_{v=1}^{N} m_v L\left(\vec{r}_v\right) L\left(\vec{r}_v\right) = \sum_{v=1}^{N} m_v \otimes^2 L\left(\vec{r}_v \vec{r}_v\right) \\ &= \left(\otimes^2 L\right)\left(\sum_{v=1}^{N} m_v \vec{r}_v \vec{r}_v\right) = \left(\otimes^2 L\right)\left(\overleftrightarrow{M}\right).\end{aligned}$$

If we subject a homogeneously stressed crystal to the mapping L, the stress tensor $\overleftrightarrow{S}$ in that crystal will not change to $\otimes^2 L(\overleftrightarrow{S})$. Rather it will be determined by the elasticity tensor E of the crystal. These examples show that it is not always obvious how to calculate the effect of subjecting a physical system to a linear mapping L. Some tensor properties of that system simply have L applied to them, and others do not.

The situation is much simpler if $L \in \Theta(V)$. To see this, suppose $\vec{b}_1$, $\vec{b}_2$, $\vec{b}_3$ is any basis for real 3-space V. Suppose $Q \in \otimes^q V$ is a measurable tensor property of a particular physical system. Then

$$Q = Q^{i_1 \cdots i_q} \vec{b}_{i_1} \vec{b}_{i_2} \cdots \vec{b}_{i_q}. \tag{8.3.2}$$

"Measuring" Q means measuring the real numbers $Q^{i_1 \cdots i_q}$. Suppose we apply to the physical system a mapping $L \in \Theta(V)$. The new physical system will have instead of Q the tensor Q_L to describe that particular physical property. Suppose we apply L to the apparatus we used to measure $Q^{i_1 \cdots i_q}$, and we use this mapped apparatus to measure the contravariant components of Q_L relative to the basis $L(\vec{b}_1)$, $L(\vec{b}_2)$, $L(\vec{b}_3)$. We assume either that no gravity or electromagnetic field is present or, if they are, that they are also mapped by L. Then the experiment on Q_L is identical to that on Q, except that it has been rotated and possibly reflected relative to the universe. As far as we know, the univese does not care; the laws of nature are invariant under reflection and rigid rotation. The universe seems to be isotropic and without handedness. (As a matter of fact, a small preference for one orientation or handedness exists in the weak nuclear force governing β-decay; Yang, Lee and Wu won Nobel prizes for this discovery; it has no measurable effects at the macroscopic level of continuum mechanics in an old universe.) Therefore, the numbers we read on our dials (or digital voltmeters) will be the same in the original and mapped systems. That is, the contravariant components of Q_L relative to $L(\vec{b}_1)$,

$L(\vec{b}_2)$, $L(\vec{b}_3)$ will be the same as the contravariant components of Q relative to $\vec{b}_1$, $\vec{b}_2$, $\vec{b}_3$. Thus

$$\begin{aligned} Q_L &= Q^{i_1\cdots i_q} L(\vec{b}_{i_1}), L(\vec{b}_{i_2}) \cdots L(\vec{b}_{i_q}) \\ &= Q^{i_1\cdots i_q} (\otimes^q L)(\vec{b}_{i_1}\vec{b}_{i_2}\cdots \vec{b}_{i_q}) \\ &= (\otimes^q L)\left[Q^{i_1\cdots i_q}\vec{b}_{i_1}\vec{b}_{i_2}\cdots \vec{b}_{i_q}\right]. \end{aligned}$$

Therefore, if $L \in \Theta(V)$ and the physical system is mapped by L, the new tensor property Q_L is obtained by applying L to the original Q;

$$Q_L = (\otimes^q L)(Q). \tag{8.3.3}$$

A tensor property of a physical system rotates with that system, and is reflected if the system if reflected.

8.4 Invariance groups

Sometimes, subjecting a physical system to certain rotations or reflections will not affect any of its measurable physical properties. For example, suppose I work all morning on a cubic crystal of Na Cl. While I am at lunch, you rotate it in such a way as simply to permute its three crystal axes. When I return from lunch, I will be unable to detect that anything has happened to the crystal. The same is true if you manange to work so fast as to disassemble and reassemble the crystal so as to subject it to a reflection along one of its crystal axes while I am at lunch. If I was working not on Na Cl but on a pot of left-handed CH F CL Br before lunch, you can subject it to any rotation whatever. I will not be able to detect the rotation when I return. However, if you reflect it, I can detect that now I have a pot of right-handed CH F Cl Br. For example, it will rotate the plane of polarized light oppositely before and after lunch.

Let $\mathcal{S}$ be any physical system in real 3-space V. For any $L \in \Theta(V)$, let $L\mathcal{S}$ be the physical system obtained by subjecting $\mathcal{S}$ to L. Let $\mathcal{G}(\mathcal{S})$ be the set of all members of $\Theta(V)$ which, when applied to $\mathcal{S}$, change none of the measurable properties of $\mathcal{S}$. That is, $L \in \theta(V)$ is a member of $\mathcal{G}(\mathcal{S})$ if no measurement can distinguish between $\mathcal{S}$ and $L\mathcal{S}$.

These definitions have some simple but useful implications. Obviously

$$I \in \mathcal{G}(\mathcal{S}) \tag{8.4.1}$$

where I is the identity operator on V. And if L_1, $L_2 \in \mathcal{G}(\mathcal{S})$, then $L_2\mathcal{S}$ is indistinguishable from $\mathcal{S}$, so $L_1(L_2\mathcal{S})$ is indistinguishable from $\mathcal{S}$. Therefore

$$\text{if } L_1, L_2 \in \mathcal{G}(\mathcal{S}) \text{ then } L_1 \circ L_2 \in \mathcal{G}(\mathcal{S}). \tag{8.4.2}$$

Finally, we claim

$$\text{if } L \in \mathcal{G}(\mathcal{S}) \text{ then } L^{-1} \in \mathcal{G}(\mathcal{S}). \tag{8.4.3}$$

For $L\mathcal{S}$ and $\mathcal{S}$ are indistinguishable. Hence so are $L^{-1}(L\mathcal{S})$ and $L^{-1}\mathcal{S}$. But $L^{-1}(L\mathcal{S})$ is $\mathcal{S}$, so $\mathcal{S}$ and $L^{-1}\mathcal{S}$ are indistinguishable. Properties (8.4.1), (8.4.2), (8.4.3) show that $\mathcal{G}(\mathcal{S})$, is a subgroup of $\Theta(V)$. It is called the invariance group of the physical system $\mathcal{S}$.

Examples of invariance groups are these:

Example 8.4.5 If $\mathcal{S}$ is water, $\mathcal{G}(\mathcal{S}) = \Theta(V)$.

Example 8.4.6 If $\mathcal{S}$ is liquid, left-handed $\mathrm{C\,H\,F\,Cl\,Br}$, $\mathcal{G}(\mathcal{S}) = \Theta^+(V)$.

Example 8.4.7 If $\mathcal{S}$ is polycrystalline calcite, with the microcrystals randomly oriented, and with equal numbers of right and left-handed microcrystals, and if we make only measurements involving lengths $> >$ microcrystal diameter (e.g. study only seismic waves very long compared to the size of the microcrystals), $\mathcal{G}(\mathcal{S}) = \Theta(V)$.

Example 8.4.8 If in example 8.4.7, 30% of the microcrystals are left handed and 70% are right handed, $\mathcal{G}(\mathcal{S}) = \Theta^+(V)$.

Example 8.4.9 If $\mathcal{S}$ is an $\mathrm{Na\,Cl}$ crystal, $\mathcal{G}(\mathcal{S})$ is the "cubic group". It is the group of orthogonal operators which send into itself a cube centered on $\vec{0}$. If $\hat{x}_1, \hat{x}_2, \hat{x}_3$ are unit vectors parallel to the edges of the cube, $\mathcal{G}(\mathcal{S})$ consists of those $L \in \Theta(V)$ such that $L(\hat{x}_i) = \sigma_{(i)}\hat{x}_{\pi(i)}$ for $i = 1, 2, 3$, where $\pi \in S_3$ and $\sigma_1, \sigma_2, \sigma_3$ can be $+1$ or -1 independently.

If $Q \in \otimes^q V$ describes some measurable property of physical system $\mathcal{S}$, then clearly $(\otimes^q L)(Q) = Q$ for all $L \in \mathcal{G}(\mathcal{S})$. This fact limits the possible forms which Q can take, and is therefore of physical interest. It leads us to introduce

Definition 8.4.26 Let V be a Euclidean vector space. Let $\mathcal{G}$ be a subgroup of $\Theta(V)$. Then $\mathcal{I}^q(\mathcal{G})$ is the set of all $Q \in \otimes^q V$ such that

$$(\otimes^q L)(Q) = Q \text{ for all } L \in \mathcal{G}. \tag{8.4.4}$$

The set $\mathcal{I}^q(\mathcal{G})$ is called the space of q'th order tensors invariant under $\mathcal{G}$. The tensors in $\mathcal{I}^q(\Theta(V))$ are unchanged by any orthogonal operator. They are called "isotropic tensors". The tensors in $\mathcal{I}^q(\Theta^+(V))$ are unchanged by any rotation, but may be changed by reflections. They are called "skew isotropic tensors." If $\mathcal{G}$ is the cubic group, the tensors in $\mathcal{I}^q(\mathcal{G})$ are unchanged by any rotation or reflection which is physically undetectable in an Na Cl crystal.

The following are simple consequences of definition 8.4.26.

Corollary 8.4.20 $\mathcal{I}^q(\mathcal{G})$ is a subspace of $\otimes^q V$.

Proof:

> Clearly $\mathcal{I}^q(\mathcal{G}) \subseteq \otimes^q V$, so all we need prove is that any linear combination of members of $\mathcal{I}^q(\mathcal{G})$ is a member. If $Q_1, \cdots, Q_N \in \mathcal{I}^q(\mathcal{G})$ and $a_1, \cdots, a_N \in \mathcal{R}$ and $L \in \mathcal{G}$, then $(\otimes^q L)(Q_i) = Q_i$, so $(\otimes^q L)(a_i Q_i) = a_i(\otimes^q L)(Q_i) = a_i Q_i$ because $\otimes^q L$ is linear.

Corollary 8.4.21 If $\mathcal{G}_1 \subseteq \mathcal{G}_2$ then $\mathcal{I}^q(\mathcal{G}_2) \subseteq \mathcal{I}^q(\mathcal{G}_1)$.

Proof:

> If $Q \in \mathcal{I}^q(\mathcal{G}_2)$, Q is unchanged by any $L \in \mathcal{G}_2$, and then certainly by any $L \in \mathcal{G}_1$. Hence $Q \in \mathcal{I}^q(\mathcal{G}_1)$. (The bigger the group $\mathcal{G}$, the harder it is to be unchanged by all its members.)

Corollary 8.4.22 If $Q \in \mathcal{I}^q(\mathcal{G})$, and $\pi \in S_q$ then $\pi Q \in \mathcal{I}^q(\mathcal{G})$.

Proof:

For any $L \in \mathcal{L}(V \to V)$, equation 8.1.11 gives $(\otimes^q L) \circ \pi = \pi \circ (\otimes^q L)$. Then $(\otimes^q L)(\pi Q) = \pi\left[(\otimes^q L)(Q)\right]$. If $Q \in \mathcal{I}^q(\mathcal{G})$ then for any $L \in \mathcal{G}$, $(\otimes^q L)(Q) = Q$, so $(\otimes^q L)(\pi Q) = \pi Q$. Hence $\pi Q \in \mathcal{I}^q(\mathcal{G})$.

Corollary 8.4.23 If $P \in \mathcal{I}^{p+q}(\mathcal{G})$ and $R \in \mathcal{I}^{q+r}(\mathcal{G})$ then $P\langle q\rangle R \in \mathcal{I}^{p+r}(\mathcal{G})$.

Proof:

For any $L \in \mathcal{G}$, $L \in \Theta(V)$, so $(\otimes^{p+r} L)(P\langle q\rangle R) = \left[(\otimes^{p+q} L)(P)\right] \langle q\rangle \left[(\otimes^{q+r} L)(Q)\right] = P\langle q\rangle R$.

Corollary 8.4.24 If $Q \in \mathcal{I}^q(\mathcal{G})$ then the component array of Q is the same relative to any two ordered orthonormal bases $(\hat{x}_1, \cdots, \hat{x}_n)$ and $(\hat{x}'_1, \cdots, \hat{x}'_n)$ such that $\hat{x}'_i L(\hat{x}_i)$ for some $L \in \mathcal{G}$.

Proof:

For any $Q \in \otimes^q V$,

$$Q = Q_{i_1 \cdots i_q} \hat{x}_{i_1} \cdots \hat{x}_{i_q} = Q'_{i_1 \cdots i_q} \hat{x}'_{i_1} \cdots \hat{x}'_{i_q}. \tag{8.4.5}$$

For any $L \in \mathcal{L}(V \to V)$,

$$\left(\otimes^q L\right)(Q) = Q_{i_1 \cdots i_q} L\left(\hat{x}_{i_1}\right) \cdots L\left(\hat{x}_{i_q}\right). \tag{8.4.6}$$

If $Q \in \mathcal{I}^q(\mathcal{G})$ and $\hat{x}'_i = L(\hat{x}_i)$ for some $L \in \mathcal{G}$, then $(\otimes^q L)(Q) = Q$ and comparing (8.4.5) with (8.4.6) gives $Q_{i_1 \cdots i_q} = Q'_{i_1 \cdots i_q}$.

Corollary 8.4.25 Suppose $Q \in \otimes^q V$. Suppose V has one ordered orthonormal basis $(\hat{x}_1, \cdots, \hat{x}_n)$ such that whenever $L \in \mathcal{G}$ and $\hat{x}'_i = L(\hat{x}_i)$, then Q has the same component array relative to $(\hat{x}_1, \cdots, \hat{x}_n)$ and $(\hat{x}'_1, \cdots, \hat{x}'_n)$. Then $Q \in \mathcal{I}^q(\mathcal{G})$.

Proof:

Suppose $L \in \mathcal{G}$. Let $\hat{x}'_i = L(\hat{x}_i)$. By hypothesis $Q'_{i_1 \cdots i_q} = Q_{i_1 \cdots i_q}$. Comparing (8.4.5), and (8.4.6) gives $(\otimes^q L)(Q) = Q$. Since this is true for every $L \in \mathcal{G}$, $Q \in \mathcal{I}^q(\mathcal{G})$.

8.5 Isotropic and skew isotropic tensors

Definition 8.5.27 Let V be a Euclidean vector space. The tensors in $\mathcal{I}^q(\Theta(V))$ are called "isotropic". The tensors in $\mathcal{I}^q(\Theta^+(V))$ are called "skew isotropic".

Any tensor describing a measurable property of water, or of polycrystalline calcite with randomly oriented microcrystals and equal numbers of left- and right-handed microcrystals, is isotropic. Any tensor describing a property of left-handed liquid C H F Cl Br , or of polycrystalline calcite with randomly oriented microcrystals and more of one handedness than the other, is skew isotropic. Since $\Theta^+(V) \subseteq \Theta(V)$, corollary (8.4.21) implies

$$\mathcal{I}^q(\Theta(V)) \subseteq \mathcal{I}^q(\Theta^+(V)). \tag{8.5.1}$$

That is, every isotropic tensor is skew isotropic.

By corollaries 8.4.24 and 8.4.25, a tensor is isotropic iff it has the same array of components relative to every orthonormal basis for V.

Suppose (V, A) is an oriented Euclidean space. By corollary 8.4.25, a tensor is skew isotropic if it has the same array of components relative to every positively oriented orthonormal basis for V. Also, it is skew isotropic if it has the same array of components relative to every negatively oriented orthonormal basis for V. If a tensor is skew isotropic, by corollary 8.4.25 it has the same array of components relative to every positively oriented orthonormal basis, and the same array relative to every negatively oriented orthonormal basis. The two arrays need not be the same (indeed, if they are, the tensor is isotropic).

Example 8.5.10 The identity tensor $\overleftrightarrow{I}$ is isotropic. For, relative to any orthonormal basis $(\hat{x}_1, \cdots, \hat{x}_n)$ for V, $I_{ij} = \overleftrightarrow{I}(\hat{x}_i, \hat{x}_j) = \hat{x}_i \cdot \hat{x}_j = \delta_{ij}$.

Example 8.5.11 If A is a unimodular alternating tensor, A is skew isotropic but not isotropic. For as we saw on page 39, A has the same component array relative to every positively oriented orthonormal basis, but its component array changes sign for negatively oriented bases.

Example 8.5.12 By corollary 8.4.23, $\overleftrightarrow{I}\overleftrightarrow{I}$ is isotropic. So is $\overleftrightarrow{I}\overleftrightarrow{I}\overleftrightarrow{I}$, $\overleftrightarrow{I}\overleftrightarrow{I}\overleftrightarrow{I}\overleftrightarrow{I}$, etc. By corollary 8.4.22 , every permutation of these tensors is isotropic. By corollary 8.4.20, every linear combination of such permuted tensors is isotropic.

Example 8.5.13 All the tensors in example 8.5.12 are skew isotropic. By corollary 8.4.23, so are $\overleftrightarrow{I}A$, $\overleftrightarrow{I}\overleftrightarrow{I}A$, $\overleftrightarrow{I}\overleftrightarrow{I}\overleftrightarrow{I}A$, etc. By corollary 8.4.22, so are all permutations of these tensors. By corollary 8.4.20, so is every linear combination of such permuted tensors. We gain nothing new by using more than one factor A. The polar identity (10.2) shows that if $n = \dim V$ then AA is a linear combination of $n!$ permutations of $\overleftrightarrow{I}\overleftrightarrow{I}\cdots\overleftrightarrow{I}$ (with n factors $\overleftrightarrow{I}$).

On p. 64 of his book "The Classical Groups" (Princeton, 1946), Hermann Weyl shows that there are no isotropic tensors except those listed in example 8.5.12, and no skew isotropic tensors except those listed in example 8.5.13. Note the corollary that all nonzero isotropic tensors are of even order. We cannot consider Weyl's general proof, but we will discuss in detail the tensors of orders $q = 0, 1, 2, 3, 4$, which are of particular interest in continuum mechanics. We assume $\dim V \geq 2$.

$q = 0$ Tensors of order 0 are scalars. They are unaffected by any orthogonal transformation, so are isotropic and therefore also skew isotropic.

$q = 1$ If $\dim V \geq 2$, the only skew isotropic tensor of order 1 is $\vec{0}$. Therefore the only isotropic tensor of order 1 is $\vec{0}$.

Proof:

> Tensors of order 1 are vectors. We are claiming that if $\dim V \geq 2$ then the only vector which is unaffected by all rotations is the zero vector. A formal proof of this intuitively obvious fact is as follows. Suppose $\vec{v} \in \mathcal{I}^1(\Theta^+(V))$. Let $\hat{x}$ and $\hat{y}$ be any two unit vectors. Each is the first vector of a positively oriented ordered orthonormal basis (pooob), so by corollary 8.4.24, $\vec{v}(\hat{x}) = \vec{v}(\hat{y})$. That is $\vec{v} \cdot \hat{x} = \vec{v} \cdot \hat{y}$ for any two unit vectors

$\hat{x}$ and $\hat{y}$. If $\dim V \geq 2$, we can always find a $\hat{y}$ such that $\vec{v} \cdot \hat{y} = 0$. Then $\vec{v} \cdot \hat{x} = 0$ for all $\hat{x}$, so $\vec{v} = \vec{0}$.

$q = 2$

$$\mathcal{I}^2(\Theta(V)) = sp\{\overleftrightarrow{I}\} \tag{8.5.2}$$

$$\mathcal{I}^2(\Theta^+(V)) = sp\{\overleftrightarrow{I}\} \text{ if } \dim V \geq 3 \tag{8.5.3}$$

$$\mathcal{I}^2(\Theta^+(V)) = sp\{\overleftrightarrow{I}, A\} \text{ if } \dim V = 2 \text{ and } A \text{ is any nonzero member of } \Lambda^2 V. \tag{8.5.4}$$

Proof:

The containments $\supseteq$ have all been established in examples 8.5.12 and 8.5.13, so we need prove only $\subseteq$ in (8.5.2, 8.5.3, 8.5.4). In all these cases, suppose $T \in \mathcal{I}^2(\Theta^+(V))$. If $\hat{x}$ and $\hat{y}$ are any unit vectors, each is the first vector of an ordered orthonormal basis (oob), and by changing the sign of one later basis vector, if necessary, we can assume that these oob's have the same orientation. Therefore $T(\hat{x}, \hat{x}) = T(\hat{y}, \hat{y})$ because T has the same component array relative to all oobs with the same orientation. Thus there is a scalar a_T such that

$$T(\hat{x}, \hat{x}) = a_T \text{ for every unit vector } \hat{x}. \tag{8.5.5}$$

If $\{\hat{x}, \hat{y}\}$ is orthonormal, $(\hat{x}, \hat{y})$ is the first pair of vectors in an oob. Replacing $(\hat{x}, \hat{y})$ by $(\hat{y}, -\hat{x})$ in that oob gives a second oob with the same orientation. Hence, $T(\hat{x}, \hat{y}) = T(\hat{y}, -\hat{x}) = -T(\hat{y}, \hat{x})$.

$$T(\hat{x}, \hat{y}) = -T(\hat{y}, \hat{x}) \text{ if } \{\hat{x}, \hat{y}\} \text{ is orthonormal .} \tag{8.5.6}$$

When $\dim V = 2$, choose an oob $(\hat{x}_1, \hat{x}_2)$ and let $a = a_T$, $b = T(\hat{x}_1, \hat{x}_2) / A(\hat{x}_1, \hat{x}_2)$. Then relative to this oob, $T_{ij} = a\delta_{ij} + bA(\hat{x}_1, \hat{x}_2)\varepsilon_{ij} = (a \overleftrightarrow{I} + bA)_{ij}$, so $T = a \overleftrightarrow{I} + bA$, and (8.5.4) is proved. To treat (8.5.2, 8.5.3), let

$(\hat{x}, \hat{y})$ be orthonormal. Both $(\hat{x}, \hat{y})$ and $(\hat{y}, \hat{x})$ are the first pair of vectors in an oob, so if $T \in \mathcal{I}^2(\Theta(V))$,

$$T(\hat{x}, \hat{y}) = T(\hat{y}, \hat{x}) \text{ if } \{\hat{x}, \hat{y}\} \text{ is orthonormal.} \tag{8.5.7}$$

If $\dim V \geq 3$, both $(\hat{x}, \hat{y})$ and $(\hat{y}, \hat{x})$ are first pairs in oobs with the same orientation, namely $(\hat{x}, \hat{y}, \hat{x}_3, \cdots, \hat{x}_n)$ and $(\hat{y}, \hat{x}, -\hat{x}_3, \hat{x}_4, \cdots, \hat{x}_n)$. If $T \in \mathcal{I}(\Theta^+(V))$, we again have 8.5.7. In either case we have (8.5.6) and (8.5.7) so

$$T(\hat{x}, \hat{y}) = 0 \text{ if } \{\hat{x}, \hat{y}\} \text{ is orthonormal.} \tag{8.5.8}$$

Now let $(\hat{x}_1, \cdots, \hat{x}_n)$ be an oob. By (8.5.5) and (8.5.8), relative to this oob we have $T_{ij} = a_T \delta_{ij}$, so $T = a_T \overset{\leftrightarrow}{I}$. QED

$q = 3$

$$\begin{aligned}
\mathcal{I}^3(\Theta(V)) &= \{0\} && (8.5.9)\\
\mathcal{I}^3(\Theta + (V)) &= \{0\} \text{ if } \dim V \neq 3 && (8.5.10)\\
\mathcal{I}^3(\Theta^+(V)) &= sp\{A\} \text{ if } \dim V = 3 \text{ and } A \neq 0, A \in \Lambda^3 V. && (8.5.11)
\end{aligned}$$

Proof:

If $\dim V$ is even, let $(\hat{x}_1, \cdots, \hat{x}_n)$ be an oob for V. Then $(-\hat{x}_1, \cdots, -\hat{x}_n)$ is another oob with the same orientation. Then if $T \in \mathcal{I}^3(\Theta^+(V))$, $T(\hat{x}_i, \hat{x}_j, \hat{x}_k) = T(-\hat{x}_i, -\hat{x}_j, -\hat{x}_k) = -T(\hat{x}_i, \hat{x}_j, \hat{x}_k) = 0$ so $T_{ijk} = 0$.

If $\dim V$ is odd and $\dim V \geq 5$, let $(\hat{x}, \hat{y}, \hat{z})$ be any ordered orthonormal sequence. It is the first triple of an oob $(\hat{x}, \hat{y}, \hat{z}, \hat{x}_4, \cdots, \hat{x}_n)$ and $(-\hat{x}, \hat{y}, \hat{z}, -\hat{x}_4, \hat{x}_5, \cdots \hat{x}_n)$ is an oob with the same orientation, so $T(\hat{x}, \hat{y}, \hat{z}) = T(-\hat{x}, \hat{y}, \hat{z}) = -T(\hat{x}, \hat{y}, \hat{z}) = 0$. Similarly, $T(\hat{x}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{y}}) = -T(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{y}}) = 0$ and $T(\hat{\boldsymbol{x}}, \hat{\boldsymbol{x}}, \hat{\boldsymbol{x}}) == -T(\hat{\boldsymbol{x}}, \hat{\boldsymbol{x}}, \hat{\boldsymbol{x}}) = 0$. Thus if $\hat{\boldsymbol{x}}_1, \hat{\boldsymbol{x}}_2, \hat{\boldsymbol{x}}_3$ are orthonormal, $T(\hat{\boldsymbol{x}}_i, \hat{\boldsymbol{x}}_j, \hat{\boldsymbol{x}}_k) = 0$ for $\{i, j, k\} \subseteq \{1, 2, 3, \}$. This comment applies to any three vectors from an arbitrary oob for V, so $T_{ijk} = 0$

for $\{i,j,k\} \subseteq \{1,\cdots,n\}$ relative to any oob, and $T = 0$. Finally, suppose $\dim V = 3$. If $\{\hat{x},\hat{y}\}$ is orthonormal, $(\hat{x},\hat{y})$ is the first pair of an oob $(\hat{x},\hat{y},\hat{z})$, and the oob $(-\hat{x},\hat{y},-\hat{z})$ has the same orientation. Thus $T(\hat{x},\hat{y},\hat{y}) = T(-\hat{x},\hat{y},\hat{y}) = -T(\hat{x},\hat{y},\hat{y}) = 0$. Similarly $T(\hat{y},\hat{x},\hat{y}) = 0$ and $T(\hat{y},\hat{y},\hat{x}) = 0$. Also $T(\hat{x},\hat{x},\hat{x}) = T(-\hat{x},-\hat{x},-\hat{x}) = -T(\hat{x},\hat{x},\hat{x}) = 0$. Thus if $T \in \mathcal{I}^3(\Theta^+(V))$

$$T(\hat{x},\hat{x},\hat{x}) = 0 \text{ for every unit vector } \hat{x} \tag{8.5.12}$$

$$T(\hat{x},\hat{y},\hat{y}) = T(\hat{y},\hat{x},\hat{y}) = T(\hat{y},\hat{y},\hat{x}) \text{ if } \{\hat{x},\hat{y}\} \text{ is orthonormal.} \tag{8.5.13}$$

Finally, if $(\hat{x},\hat{y},\hat{z})$ is an oob, then the following are oobs with the same orientation: $(\hat{y},\hat{z},\hat{x})$, $(\hat{z},\hat{x},\hat{y})$, $(\hat{y},\hat{x},-\hat{z})$, $(\hat{x},-\hat{z},\hat{y})$, $(-\hat{z},\hat{y},\hat{x})$. Then if $T \in \mathcal{I}^3(\Theta^+(V))$

$$T(\hat{x},\hat{y},\hat{z}) = T(\hat{y},\hat{z},\hat{x}) = T(\hat{z},\hat{x},\hat{y}) = -T(\hat{y},\hat{x},\hat{z}) = -T(\hat{x},\hat{z},\hat{y}) = -T(\hat{z},\hat{y},\hat{x}) \tag{8.5.14}$$

if $\{\hat{x},\hat{y},\hat{z}\}$ is orthonormal. Let $(\hat{x}_1,\hat{x}_2,\hat{x}_3)$ be an oob for V. Let $c = T(\hat{x}_1,\hat{x}_2,\hat{x}_3)/A(\hat{x}_1,\hat{x}_2,\hat{x}_3)$. Then (8.5.12 , 8.5.13 , 8.5.14) imply $T_{ijk} = cA(\hat{x}_1,\hat{x}_2,\hat{x}_3)\varepsilon_{ijk} = cA_{ijk}$ so $T = cA$. QED.

Note that in all the foregoing arguments, only mutually $\perp$ vectors were used. These arguments remain valid if $\mathcal{G}$ is the cubic group and $T \in \mathcal{I}^q(\mathcal{G})$ with $q = 1,2,3$ and we work only with unit vectors and oobs in the crystal axis directions. Therefore the conclusions apply to tensors of orders $1,2,3$ known only to be invariant under the cubic group of Na Cl. For example, it follows that the conductivity tensor $\overleftrightarrow{K}$ of Na Cl must have the form $K\overleftrightarrow{I}$. In Na Cl, Ohm's law $\vec{J} = \vec{E}\cdot\overleftrightarrow{K}$ reduces to the isotropic form $\vec{J} = \vec{E}K$, even though an Na Cl crystal is not isotropic.

$q = 4$

$$\mathcal{I}^4(\Theta(V)) \;=\; sp\left\{\overleftrightarrow{I}\overleftrightarrow{I}, (23)\overleftrightarrow{I}\overleftrightarrow{I}, (24)\overleftrightarrow{I}\overleftrightarrow{I}\right\} \tag{8.5.15}$$

$$\mathcal{I}^4(\Theta^+(V)) \;=\; '' \text{if } \dim V \notin \{2,4\} \tag{8.5.16}$$

$$\begin{aligned}\mathcal{I}^4(\Theta^+(V)) \;&=\; sp\left\{\overset{\leftrightarrow}{I}\overset{\leftrightarrow}{I}, (23)\ \overset{\leftrightarrow}{I}\overset{\leftrightarrow}{I}, (24)\ \overset{\leftrightarrow}{I}\overset{\leftrightarrow}{I}, A\right\} \\ &\quad \text{if } \dim V = 4\end{aligned} \tag{8.5.17}$$

$$\begin{aligned}\mathcal{I}^4(\Theta^+(V)) \;&=\; sp\{\overset{\leftrightarrow}{I}\overset{\leftrightarrow}{I}, (23)\ \overset{\leftrightarrow}{I}\overset{\leftrightarrow}{I}, (24)\ \overset{\leftrightarrow}{I}\overset{\leftrightarrow}{I}, \\ &\quad \overset{\leftrightarrow}{I}A, (23)\ \overset{\leftrightarrow}{I}A, (24)\ \overset{\leftrightarrow}{I}A, \\ &\quad A\overset{\leftrightarrow}{I}, (23)A\overset{\leftrightarrow}{I}, (24)A\overset{\leftrightarrow}{I}\} \text{ if } \dim V = 2.\end{aligned} \tag{8.5.18}$$

Here A is any nonzero alternating tensor, $A \in \Lambda^n V$ with $n = \dim V$.

Proof:

The containments $\supseteq$ have all been established in examples 8.5.12 and 8.5.13, so we need prove only $\subseteq$. We consider only (8.5.15, 8.5.17), and leave (8.5.17, 8.5.18) to the reader with Sitzfleisch. Therefore we assume $T \in \mathcal{I}^4(\Theta^+(V))$, and if $\dim V \in \{2,4\}$ we assume $T \in \mathcal{I}^4(\Theta(V))$. If $\dim V \geq 5$, then any four orthonormal vectors $(\hat{w}, \hat{x}, \hat{y}, \hat{z})$ are the first four vectors of and oob for V, say $(\hat{w}, \hat{x}, \hat{y}, \hat{z}, \hat{x}_5, \cdots, \hat{x}_n)$. The vectors $(-\hat{w}, \hat{x}, \hat{y}, \hat{z})$ are the first four vectors of an oob with the same orientation, namely $(-\hat{w}, \hat{x}, \hat{y}, \hat{z}, -\hat{x}_5, \hat{x}_6, \cdots, \hat{x}_n)$. Since $T \in \mathcal{I}^4(\Theta^+(V))$, T has the same component array relative to both oob's, so $T(\hat{w}, \hat{x}, \hat{y}, \hat{z}) = T(-\hat{w}, \hat{x}, \hat{y}, \hat{z}) = -T(\hat{w}, \hat{x}, \hat{y}, \hat{z})$. Thus for $\dim V \geq 5$,

$$T(\hat{w}, \hat{x}, \hat{y}, \hat{z}) = 0 \text{ if } \{\hat{w}, \hat{x}, \hat{y}, \hat{z})\} \text{ is orthonormal.} \tag{8.5.19}$$

This result is also vacuously true for $\dim V = 3$, because then we cannot find four orthonormal vectors in V.

Next, suppose $\dim V \geq 3$ and $\{\hat{x}, \hat{y}, \hat{z}\}$ is orthonormal. Then $\{\hat{x}, \hat{y}, \hat{z}\}$ and $(-\hat{x}, \hat{y}, -\hat{z})$ are both the first triple in oobs with the same orientation, so $T(\hat{x}, \hat{y}, \hat{z}, \hat{z}) = T(-\hat{x}, \hat{y}, -\hat{z}, -\hat{z}) = -T(\hat{x}, \hat{y}, \hat{z}, \hat{z}) = 0$. Similarly, if $\{\hat{x}, \hat{y}, \hat{z}\}$ is orthonormal

$$\begin{aligned}&T\left(\hat{x}, \hat{y}, \hat{z}, \hat{z}\right) = T\left(\hat{x}, \hat{z}, \hat{y}, \hat{z}\right) = T\left(\hat{x}, \hat{z}, \hat{z}, \hat{y}\right) \\ &= T\left(\hat{z}, \hat{x}, \hat{y}, \hat{z}\right) = T\left(\hat{z}, \hat{x}, \hat{z}, \hat{y}\right) = T\left(\hat{z}, \hat{z}, \hat{x}, \hat{y}\right) = 0.\end{aligned} \tag{8.5.20}$$

Equations (8.5.20) are vacuously true if $\dim V = 2$ because then we cannot find three orthonormal vectors in V.

Next, suppose $\{\hat{x}, \hat{y}\}$ is orthonormal. Then $(\hat{x}, \hat{y})$ and $(-\hat{x}, \hat{y})$ are the first pairs in oobs, and if $\dim V \geq 3$ we may assume these oobs to have the same orientation, say $(\hat{x}, \hat{y}, \hat{x}_3, \cdots \hat{x}_n)$ and $(-\hat{x}, \hat{y}, -\hat{x}_3, \hat{x}_4, \cdots \hat{x}_n)$. Then $T(\hat{x}, \hat{y}, \hat{y}, \hat{y}) = T(-\hat{x}, \hat{y}, \hat{y}, \hat{y}) = -T(\hat{x}, \hat{y}, \hat{y}, \hat{y}) = 0$. Similarly

$$\begin{aligned} & T(\hat{x}, \hat{y}, \hat{y}, \hat{y}) = T(\hat{y}, \hat{x}, \hat{y}, \hat{y}) \\ & = T(\hat{y}, \hat{y}, \hat{x}, \hat{y}) = T(\hat{y}, \hat{y}, \hat{y}, \hat{x}) = 0 \text{ if } \{\hat{x}, \hat{y}\} \text{ orthonormal.} \end{aligned} \tag{8.5.21}$$

Next, suppose that $\{\hat{x}, \hat{y}\}$ and $\{\hat{x}', \hat{y}'\}$ are both orthonormal. Each pair $(\hat{x}, \hat{y})$ and $\{\hat{x}', \hat{y}'\}$ is the first pair in an oob, so if $T \in \mathcal{I}^4(\Theta(V))$ then $T(\hat{x}, \hat{x}, \hat{y}, \hat{y}) = T(\hat{x}', \hat{x}', \hat{y}', \hat{y}')$. If $\dim V \geq 3$, then $(\hat{x}, \hat{y})$ and $(\hat{x}', \hat{y}')$ are first pairs in oobs with the same orientation, say $(\hat{x}, \hat{y}, \hat{x}_3 \cdots, \hat{x}_n)$ and $(\hat{x}', \hat{y}', \pm\hat{x}_3, \hat{x}_4, \cdots, \hat{x}_n)$. Then in that case also, $T(\hat{x}, \hat{x}, \hat{y}, \hat{y}) = T(\hat{x}', \hat{x}', \hat{y}', \hat{y}')$. In either case, there is a scalar λ_T depending only on T such that

$$T(\hat{x}, \hat{x}, \hat{y}, \hat{y}) = \lambda_T \text{ if } \{\hat{x}, \hat{y}\} \text{ is orthonormal.} \tag{8.5.22}$$

Similarly, there are scalars u_T and v_T such that

$$T(\hat{x}, \hat{y}, \hat{x}, \hat{y}) = u_T \text{ if } \{\hat{x}, \hat{y}\} \text{ is orthonormal;} \tag{8.5.23}$$

$$T(\hat{x}, \hat{y}, \hat{y}, \hat{x}) = v_T \text{ if } \{\hat{x}, \hat{y}\} \text{ is orthonormal.} \tag{8.5.24}$$

Finally, if $\hat{x}$ and $\hat{y}$ are unit vectors, each is the first member of an oob, and the two oobs can have the same orientation, so $T(\hat{x}, \hat{x}, \hat{x}, \hat{x}) = T(\hat{y}, \hat{y}, \hat{y}, \hat{y})$. Therefore, there is a scalar α_T depending only on T such that

$$T(\hat{x}, \hat{x}, \hat{x}, \hat{x}) = \alpha_T \text{ for every unit vector } \hat{x}. \tag{8.5.25}$$

Then tensor $\overleftrightarrow{I}\overleftrightarrow{I}$ satisfies (8.5.19)-(8.5.25) with $\alpha = 1$, $\lambda = 1$, $u = 0$, $v = 0$. The tensor (23) $\overleftrightarrow{I}\overleftrightarrow{I}$ satisfies (8.5.19)-(8.5.25) with $\alpha = 1$, $\lambda = 0$, $u = 1$,

$v = 0$. The tensor (24) $\overleftrightarrow{I}\overleftrightarrow{I}$ satisfies (8.5.19)-(8.5.25) with $\alpha = 1$, $\lambda = 0$, $u = 0$, $v = 1$. Therefore, the tensor

$$S = T - \lambda_T \overleftrightarrow{I}\overleftrightarrow{I} - u_T(23)\overleftrightarrow{I}\overleftrightarrow{I} - v_T(24)\overleftrightarrow{I}\overleftrightarrow{I} \tag{8.5.26}$$

satisfies (8.5.19)-(8.5.25) with $\alpha_S = \alpha_T - \lambda_T - u_T - v_T$ and $\lambda_S = u_S = v_S = 0$. Now let $(\hat{x}, \hat{y})$ be orthonormal in V and let $(\hat{x}, \hat{y}, \hat{x}_3, \cdots, \hat{x}_n)$ be an oob for V. Then $(\hat{x}', \hat{y}', \hat{x}_3, \cdots, \hat{x}_n)$ is an oob with the same orientation if we take

$$\hat{x}' = \hat{x}\cos\theta + \hat{y}\sin\theta, \quad \hat{y}' = -\hat{x}\sin\theta + \hat{y}\cos\theta.$$

To see this, let $A \in \Lambda^n V$, $A \neq 0$. Then

$$\begin{aligned}
A(\hat{x}', \hat{y}', \hat{x}_3, \cdots \hat{x}_n) &= A(\hat{x}\cos\theta + \hat{y}\sin\theta, -\hat{x}\sin\theta + \hat{y}\cos\theta, \hat{x}_3, \cdots, \hat{x}_n) \\
&= -\cos\theta\sin\theta A(\hat{x}, \hat{x}, \hat{x}_3, \cdots, \hat{x}_n) + \cos^2\theta A(\hat{x}, \hat{y}, \hat{x}_3, \cdots, \hat{x}_n) \\
&\quad - \sin^2 A(\hat{y}, \hat{x}, \hat{x}_3, \cdots, \hat{x}_n) + \sin\theta\cos\theta A(\hat{y}, \hat{y}, \hat{x}_3, \cdots, \hat{x}_n) \\
&= \left(\cos^2\theta + \sin^2\theta\right) A(\hat{x}, \hat{y}, \hat{x}_3, \cdots, \hat{x}_n) = A(\hat{x}, \hat{y}, \hat{x}_3, \cdots, \hat{x}_n).
\end{aligned}$$

But if $(\hat{x}, \hat{y}, \hat{x}_3, \cdots, \hat{x}_n)$ and $(\hat{x}', \hat{y}', \hat{x}_3, \cdots, \hat{x}_n)$ have the same orientation and $S \in \mathcal{I}^4(\Theta^+(V))$, the $S(\hat{x}, \hat{x}, \hat{x}, \hat{x}) = S(\hat{x}', \hat{x}', \hat{x}', \hat{x}')$. The S of (8.5.26) is a linear combination of members of $\mathcal{I}^4(\Theta^+(V))$, so it is in $\mathcal{I}^4(\Theta^+(V))$. In addition, S satisfies (8.5.19)-(8.5.25) with $\lambda_S = u_S = v_S = 0$. Therefore

$$\begin{aligned}
S(\hat{x}, \hat{x}, \hat{x}, \hat{x}) &= S(\hat{x}', \hat{x}', \hat{x}', \hat{x}') \\
&= S(\hat{x}\cos\theta + \hat{y}\sin\theta, \\
&\quad \hat{x}\cos\theta + \hat{y}\sin\theta, \hat{x}\cos\theta + \hat{y}\sin\theta, \hat{x}\cos\theta + \hat{y}\sin\theta) \\
&= \cos^4\theta S(\hat{x}, \hat{x}, \hat{x}, \hat{x}) + \sin^4\theta S(\hat{y}, \hat{y}, \hat{y}, \hat{y})
\end{aligned} \tag{8.5.27}$$

since all other 14 terms in the multilinear expansion of S vanish. Thus $\alpha_S = (\cos^4\theta + \sin^4\theta)\alpha_S$. It is possible to choose θ so that $\cos^4\theta + \sin^4\theta \neq 1$. Therefore $\alpha_S = 0$. Therefore $S = 0$. Therefore

$$T = \lambda_T \overleftrightarrow{I}\overleftrightarrow{I} + u_T(23)\overleftrightarrow{I}\overleftrightarrow{I} + v_T(24)\overleftrightarrow{I}\overleftrightarrow{I} . \tag{8.5.28}$$

This completes the proofs of (8.5.15) and (8.5.17).

Notice that the proof of (8.5.26) works equally well if all the unit vectors are parallel to the crystal axes of an Na Cl crystal. Only (8.5.27) fails. But if $\hat{x}, \hat{y}, \hat{z}$ are crystal axes, then $S = \alpha_S(\hat{x}\hat{x}\hat{x}\hat{x} + \hat{y}\hat{y}\hat{y}\hat{y} + \hat{z}\hat{z}\hat{z}\hat{z})$. Therefore, we have proved that if T is invariant under the cubic group we can write

$$\begin{aligned} T &= \alpha_S\left(\hat{x}\hat{x}\hat{x}\hat{x} + \hat{y}\hat{y}\hat{y}\hat{y} + \hat{z}\hat{z}\hat{z}\hat{z}\right) \\ &+ \lambda_T \overleftrightarrow{I}\overleftrightarrow{I} + u_T(23)\overleftrightarrow{I}\overleftrightarrow{I} + v_T(24)\overleftrightarrow{I}\overleftrightarrow{I}, \end{aligned} \tag{8.5.29}$$

where $\alpha_S = \alpha_T - \lambda_T - u_T - v_T$.

Chapter 9

Differential Calculus of Tensors

9.1 Limits in Euclidean vector spaces

Definition 9.1.28 Let U be a Euclidean vector space. Let $\vec{u} \in U$. Let $t \in \mathcal{R}$, $t > 0$. We define three sets:

i) $B(\vec{u}, t) := \{\vec{u}' : \vec{u}' \in U \text{ and } \|\vec{u}' - \vec{u}\| < t\}$

ii) $\partial B(\vec{u}, t) := \{\vec{u}' : \vec{u}' \in U \text{ and } \|\vec{u}' - \vec{u}\| = t\}$

iii) $\bar{B}(\vec{u}, t) := \{\vec{u}' : \vec{u}' \in U \text{ and } \|\vec{u}' - \vec{u}\| \leq t\}$.

$B(\vec{u}, t)$, $\partial B(\vec{u}, t)$ and $\bar{B}(\vec{u}, t)$ are called, respectively, the open ball, the sphere, and the closed ball centered at $\vec{u}$ with radius t. See figure 9.1

Definition 9.1.29 A subset D of Euclidean space U is "open" if every point in D is the center of a ball lying entirely in D. That is, if $\vec{u} \in D$ then $\exists \varepsilon > 0 \ni B(\vec{u}, \varepsilon) \subseteq D$. See figure 9.2.

Definition 9.1.30 Suppose U and V are Euclidean vector spaces, $D \subseteq U$, $\vec{f} : D \to V$, $\vec{u}_\circ \in D$ and $\vec{v}_\circ \in V$. We say that $\lim_{\vec{u} \to \vec{u}_\circ} \vec{f}(\vec{u}) = \vec{v}_\circ$, or, equivalently $\vec{f}(\vec{u}) \to \vec{v}_\circ$, as $\vec{u} \to \vec{u}_\circ$ if for every $\varepsilon > 0$ there is a $\delta(\varepsilon) > 0$ such that whenever $\vec{u} \in D$ and $0 < \|\vec{u} - \vec{u}_\circ\| < \delta(\varepsilon)$ then $\|f(\vec{u}) - \vec{v}_0\| < \varepsilon$.

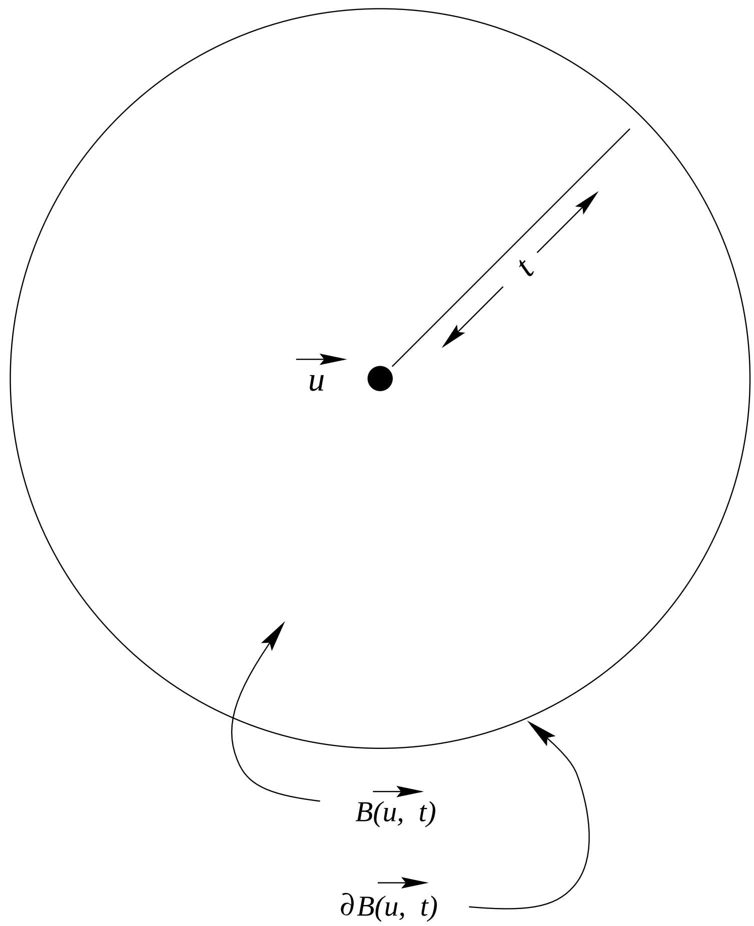


Figure 9.1:

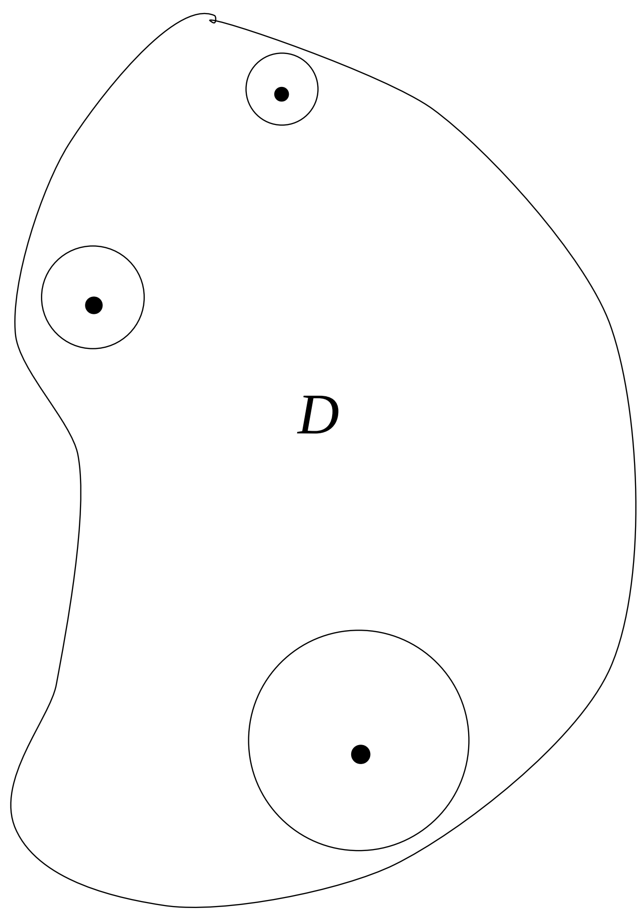


Figure 9.2:

Definition 9.1.31 If $\vec{f}(\vec{u}_\circ) = \vec{v}_\circ$ in definition 9.1.30 then $\vec{f}$ is "continuous at $\vec{u}_\circ$".

Definition 9.1.32 If x and y are real numbers, denote the algebraically smaller of them by $x \wedge y$ (read "x meet y") and the algebraically larger by $x \vee y$ (read "x join y"). Thus $1 \wedge 6 = 1$, $1 \vee 6 = 6$, $(-1) \wedge 0 = -1$, $(-1) \vee 0 = 0$.

Corollary 9.1.26 Suppose U and V are Euclidean vector spaces, $D \subseteq U$ and $f : D \to V$. Suppose $\vec{u}_0 \in D$, $\vec{v}_0 \in V$ and $\vec{f}(\vec{u}) \to \vec{v}_0$ as $\vec{u} \to \vec{u}_0$. Then if $\vec{u} \in D$ and $\|\vec{u} - \vec{u}_0\| < \delta(1)$, we have

$$\|f(\vec{u})\| < 1 + \|\vec{v}_0\|.$$

Proof:

If $\vec{u} \in D$ and $\|\vec{u} - \vec{u}_0\| < \delta(1)$, then $\|f(\vec{u}) - \vec{v}_0\| < 1$ so $\|\vec{f}(\vec{u})\| = \|\vec{f}(\vec{u}) - \vec{v}_0 + \vec{v}_0\| \leq \|f(\vec{u}) - \vec{v}_0\| + \|\vec{v}_0\| < 1 + \|\vec{v}_0\|$.

Corollary 9.1.27 Suppose U and V are Euclidean vector spaces, $D \subseteq U$, $\vec{f} : D \to V$, $\vec{f}' : D \to V$, $\vec{u}_0 \in D$, $\vec{v}_0$ and $\vec{v}_0' \in V$. Suppose that $\vec{f}(\vec{u}) \to \vec{v}_0$ and $\vec{f}'(\vec{u}) \to \vec{v}_0'$ as $\vec{u} \to \vec{u}_0$. Then $(\vec{f} - \vec{f}')(\vec{u}) \to \vec{v}_0 - \vec{v}_0'$ as $\vec{u} \to \vec{u}_0$.

Proof:

For any given $\varepsilon > 0$ we must find a $\delta''(\varepsilon) > 0$ such that $\vec{u} \in D$ and $\|\vec{u} - \vec{u}_0\| < \delta''(\varepsilon)$ implies $\|(f + f')(\vec{u}) - \vec{v}_0 - \vec{v}_0'\| < \varepsilon$. By hypothesis, there are $\delta(\varepsilon/2) > 0$ and $\delta'(\varepsilon/2) > 0$ such that if $\vec{u} \in D$ then

$$\|\vec{f}(\vec{u}) - \vec{v}_0\| < \varepsilon/2 \quad \text{if} \quad \|\vec{u} - \vec{u}_0\| < \delta(\varepsilon/2) \tag{9.1.1}$$

$$\|\vec{f}'(\vec{u}) - \vec{v}_0'\| < \varepsilon/2 \quad \text{if} \quad \|\vec{u} - \vec{u}_0\| < \delta'(\varepsilon/2). \tag{9.1.2}$$

Let $\delta''(\varepsilon) = \delta(\varepsilon/2) \wedge \delta'(\varepsilon/2)$. If $\|\vec{u} - \vec{u}_0\| < \delta''(\varepsilon)$, then both (9.1.1) and (9.1.2) are true, so $\|(\vec{f} + \vec{f}')(\vec{u}) - \vec{v}_0 - \vec{v}_0'\| = \|\vec{f}(\vec{u}) - \vec{v}_0 + \vec{f}'(\vec{u}) - \vec{v}_0'\| \leq \|\vec{f}(\vec{u}) - \vec{v}_0\| + \|\vec{f}'(\vec{u}) - \vec{v}_0'\| \leq \varepsilon/2 + \varepsilon/2 = \varepsilon$. QED.

Corollary 9.1.28 Suppose U, V, W, X are Euclidean vector spaces and $D \subseteq U$ and $R : D \to V \otimes W$ and $T : D \to W \otimes X$. Define $R \cdot T : D \to V \otimes X$ by requiring $(R \cdot T)(\vec{u}) = R(\vec{u}) \cdot T(\vec{u})$ for each $\vec{u} \in D$. Suppose $\vec{u}_0 \in D$ and as $\vec{u} \to \vec{u}_0$ we have $R(\vec{u}) \to R_0$ and $T(\vec{u}) \to T_0$. Then $(R \cdot T)(\vec{u}) \to R_0 \cdot T_0$.

Proof:

Choose $\varepsilon > 0$. We must find $\delta(\varepsilon) > 0$ such that if $\vec{u} \in D$ and $\|\vec{u} - \vec{u}_0\| < \delta(\varepsilon)$ then $\|(R \cdot T)(\vec{u}) - R_0 \cdot T_0\| < \varepsilon$. We note

$$\begin{aligned}(R \cdot T)(\vec{u}) - R_0 \cdot T_0 &= R(\vec{u}) \cdot T(\vec{u}) - R_0 \cdot T_0 \\ &= R(\vec{u}) \cdot T(\vec{u}) - R(\vec{u}) \cdot T_0 + R(\vec{u}) \cdot T_0 - R_0 \cdot T_0 \\ &= R(\vec{u}) \cdot [T(\vec{u}) - T_0] + [R(\vec{u}) - R_0] \cdot T_0.\end{aligned}$$

Thus

$$\|(R \cdot T)(\vec{u}) - R_0 \cdot T_0\| \leq \|R(\vec{u}) \cdot [T(\vec{u}) - T_0]\| + \|[R(\vec{u}) - R_0] \cdot T_0\|.$$

The generalized Schwarz inequality, remark 7.3.36, gives

$$\begin{aligned}\|R(\vec{u}) \cdot [T(\vec{u}) - T_0]\| &\leq \|R(\vec{u})\|\|T(\vec{u}) - T_0\| \\ \|(R(\vec{u}) - R_0) \cdot T_0\| &\leq \|R(\vec{u}) - R_0\|\|T_0\|,\end{aligned}$$

Thus if $\vec{u} \in D$

$$\|(R \cdot T)(\vec{u}) - R_0 \cdot T_0\| \leq \|R(\vec{u})\|\|T(\vec{u}) - T_0\| + \|R(\vec{u}) - R_0\|\|T_0\|. \tag{9.1.3}$$

Now if $\vec{u} \in D$ and

$$\left\{\begin{array}{l}\text{if } \|\vec{u} - \vec{u}_0\| < \delta_R(1) \text{ then } \|R(\vec{u})\| < 1 + \|R_0\| \text{ (see corollary 9.1.26)} \\ \text{if } \|\vec{u} - \vec{u}_0\| < \delta_T\left(\frac{\varepsilon}{2(1+\|R_0\|)}\right) \text{ then } \|T(\vec{u}) - T_0\| < \frac{\varepsilon}{2(1+\|R_0\|)} \\ \text{if } \|\vec{u} - \vec{u}_0\| < \delta_R\left(\frac{\varepsilon}{2(1+\|T_0\|)}\right) \text{ then } \|R(\vec{u}) - R_0\| < \frac{\varepsilon}{2(1+\|T_0\|)}.\end{array}\right. \tag{9.1.4}$$

Take $\delta(\varepsilon) = \delta_R(1) \wedge \delta_T\left(\frac{\varepsilon}{2(1+\|R_0\|)}\right) \wedge \delta_R\left(\frac{\varepsilon}{2(1+\|T_0\|)}\right)$. Then if $\|\vec{u} - \vec{u}_0\| < \delta(\varepsilon)$, we have all three of (9.1.4), so by (9.1.3),

$$\|(R \cdot T)(\vec{u}) - R_0 \cdot T_0\| < (1 + \|R_0\|) \frac{\varepsilon}{2(1 + \|R_0\|)} + \frac{\|T_0\|\varepsilon}{2(1 + \|T_0\|)} < \varepsilon.$$

QED.

Corollary 9.1.29 Suppose U and V are Euclidean vector spaces, and $D \subseteq U$, and $\vec{u}_0 \in D$ and $\vec{v}_0 \in V$. Suppose $\vec{f} : D \to V$ is defined by $\vec{f}(\vec{u}) = \vec{v}_0$ for all $\vec{u} \in D$. Then $\vec{f}(\vec{u}) \to \vec{v}_0$ as $\vec{u} \to \vec{u}_0$.

Proof:

Obvious from definition 9.1.30.

Corollary 9.1.30 Suppose U and V are Euclidean vector spaces, and $D \subseteq U$, and $\vec{u}_0 \in D$ and $\vec{v}_0 \in V$. Suppose $\vec{f} : D \to V$.

i) If $\vec{f}(\vec{u}) \to \vec{v}_0$ as $\vec{u} \to \vec{u}_0$, then $\vec{b}^i \cdot \vec{f}(\vec{u}) \to \vec{b}^i \cdot \vec{v}_0$ as $\vec{u} \to \vec{u}_0$, for any basis $(\vec{b}_1, \cdots, \vec{b}_n)$ of V, with dual basis $(\vec{b}^1, \cdots, \vec{b}^n)$.

ii) If there is one basis for V, $(\vec{b}_1, \cdots, \vec{b}_n)$, such that $\vec{b}^i \cdot \vec{f}(\vec{u}) \to \vec{b}^i \cdot \vec{v}_0$ as $\vec{u} \to \vec{u}_0$, then $\vec{f}(\vec{u}) \to \vec{v}_0$ as $\vec{u} \to \vec{u}_0$.

Proof:

Regard $\vec{b}^i$ as a function whose domain is D and whose value is $\vec{b}^i$ for each $\vec{u} \in D$. Then i) follows from corollaries 9.1.28 and 9.1.29. And ii) follows from corollaries 9.1.29 , 9.1.28, and 9.1.27.

Notice that the ranges of our functions can lie in tensor spaces since these are Euclidean vector spaces. The corollaries all apply verbatim. The dot products are, of course, the generalized dot products used to make tensor spaces Euclidean. The only corollary which perhaps deserves comment is 9.1.28. It now reads thus: Suppose U, $V_1, \cdots, V$ $W_1, \cdots, W_s$, $X_1, \cdots, X_t$ are Euclidean vector spaces, and $D \subseteq U$ and $R : D \to V_1 \otimes$

$\cdots \otimes V_r \otimes W_1 \otimes \cdots \otimes W_s$ and $T : D \to W_1 \otimes \cdots \otimes W_s \otimes X_1 \otimes \cdots \otimes X_t$. Define $R\langle s\rangle T$: $D \to V_1 \otimes \cdots \otimes V_r \otimes X_1 \otimes \cdots \otimes X_t$ by requiring $(R\langle s\rangle T)(\vec{u}) = R(\vec{u})\langle s\rangle T(\vec{u})$ for each $\vec{u} \in D$. Suppose $\vec{u}_0 \in D$ and as $\vec{u} \to \vec{u}_0$ we have $R(\vec{u}) \to R_0$ and $T(\vec{u}) \to T_0$. Then $(R\langle s\rangle T)(\vec{u}) \to R_0\langle s\rangle T_0$.

9.2 Gradients: Definition and simple properties

Definition 9.2.33 We define the "dot product" of two real numbers to be their ordinary arithmetic product. With this definition, $\mathcal{R}$ is a one-dimensional Euclidean vector space, and it has two orthonormal bases, $\{1\}$ and $\{-1\}$, which we write as $\hat{1}$ and $-\hat{1}$ when thinking of $\mathcal{R}$ as a Euclidean space. When we think of a real number as a vector, we write it with an arrow, like $\vec{2}$. Thus $\vec{2} = 2\hat{1}$, and $2 = \vec{2} \cdot \hat{1} = \hat{1} \cdot \vec{2}$.

Definition 9.2.34 Suppose U and V are Euclidean vector spaces, D is an open subset of U, $\vec{f} : D \to V$, and $\vec{u} \in D$. A linear mapping $L : U \to V$ is a "gradient mapping for $\vec{f}$ at $\vec{u}$" if there is a function $\vec{R} : U \to V$ (which may depend on $\vec{f}$ and $\vec{u}$ and L) such that

$$\begin{array}{ll} a) & \lim_{\vec{h}\to\vec{0}} \vec{R}(\vec{h}) = \vec{0}, \text{ and} \\ b) & \vec{f}(\vec{u}+\vec{h}) = \vec{f}(\vec{u}) + L(\vec{h}) + \|\vec{h}\|\vec{R}(\vec{h}) \text{ for any} \\ & \vec{h} \in U \text{ such that } \vec{u}+\vec{h} \in D. \end{array} \tag{9.2.1}$$

Remark 9.2.42 $\vec{f}$ has at most one gradient mapping at $\vec{u}$.

Proof:

If L_1 and L_2 are both gradient mappings for $\vec{f}$ at $\vec{u}$ then whenever $\vec{u} + \vec{h} \in D$ we have

$$\begin{aligned} \vec{f}(\vec{u}+\vec{h}) &= \vec{f}(\vec{u}) + L_1(\vec{h}) + \|\vec{h}\|\vec{R}_1(\vec{h}) \\ &= \vec{f}(\vec{u}) + L_2(\vec{h}) + \|\vec{h}\|\vec{R}_2(\vec{h}). \end{aligned}$$

Therefore if $\vec{u} + \vec{h} \in D$,

$$L_1(\vec{h}) + \|\vec{h}\|\vec{R}_1(\vec{h}) = L_2(\vec{h}) + \|\vec{h}\|\vec{R}_2(\vec{h}). \tag{9.2.2}$$

Now D is open, so there is an $\varepsilon > 0$ such that if $\|\vec{h}\| < \varepsilon$ then $\vec{u} + \vec{h} \in D$. Let $\vec{k}$ be any vector in U, and let t be any nonnegative real number. Then $\|t\vec{k}\| < \varepsilon$ as long as $0 \le t < \varepsilon/\|\vec{k}\|$. Therefore, if t is in this range, 9.2.2 is true for $\vec{h} = t\vec{k}$. That is

$$L_1(t\vec{k}) + \|t\vec{k}\|\vec{R}_1(t\vec{k}) = L_2(t\vec{k}) + \|tk\|R_2(t\vec{k}).$$

But $L_i(t\vec{k}) = tL_i(k)$ and $\|t\vec{k}\| = t\|\vec{k}\|$, so if $t > 0$

$$L_1(\vec{k}) + \|\vec{k}\|\vec{R}_1(t\vec{k}) = L_2(\vec{k}) + \|\vec{k}\|\vec{R}_2(t\vec{k}). \tag{9.2.3}$$

This is true for all t in $0 < t < \varepsilon/\|\vec{k}\|$. Let $t \to 0$, and both $\vec{R}_1(t\vec{k}) \to \vec{0}$ and $\vec{R}_2(t\vec{k}) \to \vec{0}$. Therefore $L_1(\vec{k}) = L_2(\vec{k})$. Since this is true for all $k \in U$, we have $L_1 = L_2$. QED

Definition 9.2.35 In definition 9.2.34 if $\vec{f} : D \to V$ has a gradient mapping L at $\vec{u}$, then $\vec{f}$ is "differentiable at $\vec{u}$", and $\vec{R}$ is the "remainder function for $\vec{f}$ at $\vec{u}$". The tensor $\overset{\leftrightarrow}{L} \in U \otimes V$ is called the "gradient tensor of f at $\vec{u}$" or the "gradient of $\vec{f}$ at $\vec{u}$". It is written $\vec{\nabla}\vec{f}(\vec{u})$. Since $L(\vec{h}) = \vec{h} \cdot \overset{\leftrightarrow}{L}$, we can rewrite 9.2.33 as

$$\begin{aligned} &a) \quad \lim_{\vec{h}\to\vec{0}} \vec{R}(\vec{h}) = \vec{0} \text{ and} \\ &b) \quad \vec{f}(\vec{u}+\vec{h}) = \vec{f}(\vec{u}) + \vec{h}\cdot\left[\vec{\nabla}\vec{f}(\vec{u})\right] + \|\vec{h}\|\vec{R}(\vec{h}) \text{ if } \vec{u}+\vec{h} \in D. \end{aligned} \tag{9.2.4}$$

Remark 9.2.43 Suppose U and V are Euclidean vector spaces, D is an open subset of U, $\vec{u} \in D$, and $\vec{f} : D \to V$. Suppose there is a tensor $\overset{\leftrightarrow}{L} \in U \otimes V$ and a function $\vec{R} : D \to V$ with these properties:

$$\begin{aligned} &a) \quad \lim_{\vec{h}\to\vec{0}} \vec{R}(\vec{h}) = \vec{0} \text{ and} \\ &b) \quad \vec{f}(\vec{u}+\vec{h}) = \vec{f}(\vec{u}) + \vec{h}\cdot L + \|\vec{h}\|\vec{R}(\vec{h}) \text{ if } \vec{u}+\vec{h} \in D. \end{aligned} \tag{9.2.5}$$

Then $\vec{f}$ is differentiable at $\vec{u}$, its gradient tensor at $\vec{u}$ is $\overset{\leftrightarrow}{L}$, and its remainder function at $\vec{u}$ is $\vec{R}$.

Proof:

$\vec{h} \mapsto \vec{h} \cdot \overleftrightarrow{L}$ is a linear mapping L, so (9.2.5) $\Leftrightarrow$ (9.2.1). Therefore $\vec{f}$ is differentiable. By Remark 9.2.42, L is the gradient mapping for $\vec{f}$ at $\vec{u}$, so $\overleftrightarrow{L}$ is the gradient tensor. Then the $\vec{R}$'s in (9.2.1b) and (9.2.5b) must be the same.

Example 9.2.14 Suppose $\vec{f} : D \to V$ is constant. Then $\vec{f}$ is differentiable at every $\vec{u} \in D$ and $\vec{\nabla}\vec{f}(\vec{u}) = \overleftrightarrow{0}$.

Proof:

The hypotheses of remark 9.2.43 are satisfied with $\overleftrightarrow{L} = \overleftrightarrow{0}$, $\overleftrightarrow{R}(\vec{h}) = \vec{0}$.

Example 9.2.15 Suppose $\vec{f} : D \to U$ is the identity mapping on D, i.e. $\vec{f}(\vec{u}) = \vec{u}$ for all $\vec{u} \in D$. Then $\vec{f}$ is differentiable everywhere in D, and for each $\vec{u} \in D$, $\vec{\nabla}\vec{f}(\vec{u}) = \overleftrightarrow{I}_U$, the identity tensor on U. (This result is often abbreviated $\vec{\nabla}\vec{u} = \overleftrightarrow{I}_U$.)

Proof:

The hypotheses of remark 9.2.43 are satisfied with $\overleftrightarrow{L} = \overleftrightarrow{I}_U$, $\vec{R}(\vec{h}) = \vec{0}$.

Example 9.2.16 Suppose U and V are Euclidean spaces, D is an open subset of U and $\overleftrightarrow{T}$ is a fixed tensor in $U \otimes V$. Suppose $\vec{f} : D \to V$ is defined by $\vec{f}(\vec{u}) = \vec{u} \cdot \overleftrightarrow{T}$ for every $\vec{u} \in D$. Then at every $\vec{u} \in D$, $\vec{f}$ is differentiable and $\vec{\nabla}\vec{f}(\vec{u}) = \overleftrightarrow{T}$. (Often abbreviated $\vec{\nabla}(\vec{u} \cdot \overleftrightarrow{T}) = \overleftrightarrow{T}$.)

Proof:

The hypotheses of remark 9.2.43 are satisfied with $\overleftrightarrow{L} = \overleftrightarrow{T}$, $\vec{R}(\vec{h}) = \vec{0}$.

Example 9.2.17 Suppose that for $i \in \{1, \cdots, n\}$ we have $a_i \in \mathcal{R}$ and $\vec{f}_i : D \to V$. Suppose all the $\vec{f}_i$ are differentiable at $\vec{u} \in D$. Then so is $a_i\vec{f}_i : D \to V$, and $\vec{\nabla}(a_i\vec{f}_i)(\vec{u}) = a_i(\vec{\nabla}\vec{f}_i)(\vec{u})$.

Proof:

Each $\vec{f}_i$ satisfies equation (9.2.4), so the hypotheses of remark 9.2.43 are satisfied for $a_i\vec{f}_i$ with $\overleftrightarrow{L} = a_i[\vec{\nabla}\vec{f}_i(\vec{u})]$ and $\vec{R}(\vec{h}) = a_i\vec{R}_i(\vec{h})$.

Example 9.2.18 Suppose that D is an open subset of Euclidean space U and $f : D \to \mathcal{R}$. In elementary calculus classes, f is defined to be differentiable at $\vec{u} \in D$ if there is a vector $\vec{L} \in U$ and a function $R : U \to \mathcal{R}$ such that

$$\begin{array}{ll} a) & \lim_{\vec{h}\to\vec{0}} R(\vec{h}) = 0 \text{ and} \\ b) & f(\vec{u}+\vec{h}) = f(\vec{u}) + \vec{h}\cdot\vec{L} + \|\vec{h}\|R(\vec{h}) \text{ if } \vec{u}+\vec{h} \in D. \end{array} \tag{9.2.6}$$

The vector $\vec{L}$ is called the gradient of f at $\vec{u}$, written $\vec{\nabla} f(\vec{u})$. If we think of $\mathcal{R}$ as a one-dimensional Euclidean space, then we can think of f as a vector-valued function $\vec{f} = f\hat{1}$ (the vector is one dimensional). The scalar valued function f is differentiable in the sense of equations (9.2.6) iff $\vec{f} = f\hat{1}$ is differentiable in the sense of equation (9.2.4), and then

$$\vec{\nabla}\vec{f}(\vec{u}) = \vec{\nabla} f(\vec{u})\hat{1}, \quad \vec{\nabla} f(\vec{u}) = \vec{\nabla}\vec{f}(\vec{u})\cdot\hat{1}. \tag{9.2.7}$$

Proof:

To obtain (9.2.4) from (9.2.6), multiply all terms in (9.2.6) on the right by $\hat{1}$. To obtain (9.2.6) from (9.2.4) dot $\hat{1}$ on the right in all terms in (9.2.4). Here clearly $\overleftrightarrow{L} = \vec{L}\hat{1} = \vec{\nabla} f(\vec{u})\hat{1}$.

Example 9.2.19 Suppose D is an open subset of $\mathcal{R}$, regarded as a one-dimensional Euclidean space. Suppose $\vec{f} : D \to V$ has an ordinary derivative at $u \in D$ (we will write this derivative as $\partial_u \vec{f}(u)$ rather than $d\vec{f}/du$, for reasons made clear later). Then $\vec{f}$ is differentiable at $\vec{u}$ and

$$\vec{\nabla}\vec{f}(u) = \hat{1}\partial_u \vec{f}(u). \tag{9.2.8}$$

Proof:

By hypothesis, $\lim_{h\to 0}\left[\frac{\vec{f}(u+h)-\vec{f}(u)}{h}\right] = \partial_u \vec{f}(u)$. Define $\vec{R}(h) = 0$ if $h = 0$ and for $h \neq 0$ then

$$\vec{R}(h) = \frac{h}{|h|}\left\{\frac{\vec{f}(u+h)-\vec{f}(u)}{h} - \partial_u \vec{f}(u)\right\}.$$

Then $\lim_{h\to 0} \vec{R}(h) = \vec{0}$, and

$$\vec{f}(u+h) = \vec{f}(u) + h\partial_u \vec{f}(u) + |h|\vec{R}(h). \tag{9.2.9}$$

But $h = \vec{h}\cdot\hat{1}$ so $h\partial_u\vec{f}(\vec{u}) = \vec{h}\cdot\hat{1}\partial_u\vec{f}(u)$. Also $|h| = \|\vec{h}\|$. Therefore (9.2.9) is (9.2.5) with $\overleftrightarrow{L} = \hat{1}\partial_u\vec{f}(u)$.

Example 9.2.20 Suppose $f : D \to \mathcal{R}$ and $\vec{g} : D \to V$ are differentiable at $\vec{u} \in D$. Then so is the product function $f\vec{g} : D \to V$, and

$$\vec{\nabla}(f\vec{g})(\vec{u}) = \left[\vec{\nabla}f(\vec{u})\right]\vec{g}(\vec{u}) + f(\vec{u})\left[\vec{\nabla}\vec{g}(\vec{u})\right]. \tag{9.2.10}$$

Proof:

By hypothesis, if $\vec{u}+\vec{h} \in D$ then

$$\begin{aligned} f(\vec{u}+\vec{h}) &= f(\vec{u}) + \vec{h}\cdot\vec{\nabla}f(\vec{u}) + \|\vec{h}\|R_f(\vec{h} \\ \vec{g}(\vec{u}+\vec{h}) &= \vec{g}(\vec{u}) + \vec{h}\cdot\vec{\nabla}\vec{g}(\vec{u}) + \|\vec{h}\|R_g(\vec{h}). \end{aligned}$$

Multiplying these two equations gives

$$f(\vec{u}+\vec{h})\vec{g}(\vec{u}+\vec{h}) = f(\vec{u})\vec{g}(\vec{u}) + \vec{h}\cdot\left[f(\vec{\nabla}\vec{g}) + (\vec{\nabla}f)\vec{g}\right] + \|\vec{h}\|\vec{R}_{fg}(\vec{h}) \tag{9.2.11}$$

where, by definition,

$$\begin{aligned} \vec{R}_{fg}(\vec{h}) &= R_f(\vec{h})\left[\vec{g}(\vec{u}) + \vec{h}\cdot\vec{\nabla}\vec{g}(\vec{u})\right] \\ &\quad + \vec{R}_g(\vec{h})\left[f(\vec{u}) + \vec{h}\cdot\vec{\nabla}f(\vec{u})\right] + \|\vec{h}\|R_f(\vec{h})\vec{R}_g(\vec{h}) \\ &\quad + \left[\vec{h}\cdot\vec{\nabla}f(\vec{u})\right]\left[\vec{h}\cdot\vec{\nabla}\vec{g}(\vec{u})\right]/\|\vec{h}\|. \end{aligned}$$

As $\vec{h} \to \vec{0}$, the three terms in $\vec{R}_{fg}$, which involve R_f or $\vec{R}_g$, obviously $\to 0$. So does the last term, because by Schwarz's inequality its length is $\leq \|\vec{h}\|\|\vec{\nabla}f(\vec{u})\|\|\vec{\nabla}\vec{g}(\vec{u})\|$. Therefore remark 9.2.43 can be applied to (9.2.11). QED.

Example 9.2.21 The Chain Rule. Suppose U, V, W are Euclidean vector spaces D_f is an open subset of U, D_g is an open subset of V, and $f : D_f \to D_g$ and $g : D_g \to W$. Suppose $\vec{u} \in D_f$ and $\vec{v} = \vec{f}(\vec{u})$. Suppose $\vec{f}$ is differentiable at $\vec{u}$ and $\vec{g}$ is differentiable at $\vec{v}$. Then their composition, $\vec{g} \circ \vec{f} : D_f \to W$, is differentiable at $\vec{u}$ and

$$\vec{\nabla}(\vec{g} \circ \vec{f})(\vec{u}) = \vec{\nabla}\vec{f}(\vec{u}) \cdot \vec{\nabla}\vec{g}(\vec{v}). \tag{9.2.12}$$

Proof:

Let $\overleftrightarrow{L}_f = \vec{\nabla}\vec{f}(\vec{u})$ and $\overleftrightarrow{L}_g = \vec{\nabla}\vec{g}(\vec{v})$. Let $\vec{R}_f$ and $\vec{R}_g$ be the remainder functions for $\vec{f}$ at $\vec{u}$ and $\vec{g}$ at $\vec{v}$. Then $\vec{R}_f(\vec{h}) \to \vec{0}$ as $\vec{h} \to \vec{0}$, and $\vec{R}_g(\vec{k}) \to \vec{0}$ as $\vec{k} \to \vec{0}$, and

$$\vec{f}(\vec{u} + \vec{h}) = \vec{f}(\vec{u}) + \vec{h} \cdot \overleftrightarrow{L}_f + \|\vec{h}\| \vec{R}_f(\vec{h}) \text{ if } \vec{u} + \vec{h} \in D_f; \tag{9.2.13}$$

$$\vec{g}(\vec{v} + \vec{k}) = \vec{g}(\vec{v}) + \vec{k} \cdot \overleftrightarrow{L}_g + \|\vec{k}\| \vec{R}_g(\vec{k}) \text{ if } \vec{v} + \vec{k} \in D_g. \tag{9.2.14}$$

We hope to find $\vec{R}_{g \circ f} : U \to W$ such that $\vec{R}_{g \circ f}(\vec{h}) \to \vec{0}$ as $\vec{h} \to \vec{0}$ and $(\vec{g} \circ \vec{f})(\vec{u} + \vec{h}) = (\vec{g} \circ \vec{f})(\vec{u}) + \vec{h} \cdot \left[\overleftrightarrow{L}_f \cdot \overleftrightarrow{L}_g\right] + \|\vec{h}\| \vec{R}_{g \circ f}(\vec{h})$ whenever $\vec{u} + \vec{h} \in D_f$. This equation is the same as

$$\vec{g}\left[\vec{f}(\vec{u}) + \vec{h})\right] = \vec{g}\left[\vec{f}(\vec{u})\right] + (\vec{h} \cdot \overleftrightarrow{L}_f) \cdot \overleftrightarrow{L}_g + \|\vec{h}\| \vec{R}_{g \circ f}(\vec{h}). \tag{9.2.15}$$

To obtain (9.2.15), choose any $\vec{h} \in U$ such that $\vec{u} + \vec{h} \in D_f$ and define

$$\vec{k} = \vec{f}(\vec{u} + \vec{h}) - \vec{f}(\vec{u}) = \vec{f}(\vec{u} + \vec{h}) - \vec{v}. \tag{9.2.16}$$

Then $\vec{k} \in V$ and $\vec{v} + \vec{k} = \vec{f}(\vec{u} + \vec{h}) \in D_g$, so (9.2.14) holds for the $\vec{k}$ of (9.2.16). Moreover, from (9.2.13)

$$\vec{k} = \vec{h} \cdot \overleftrightarrow{L}_f + \|\vec{h}\| \vec{R}_f(\vec{h}). \tag{9.2.17}$$

Substituting (9.2.17) in (9.2.14) gives

$$\vec{g}\left[\vec{f}(\vec{u} + \vec{h})\right] = \vec{g}\left[\vec{f}(\vec{u})\right] + \left[\vec{h} \cdot \overleftrightarrow{L}_f + \|\vec{h}\| \vec{R}_f(\vec{h})\right] \cdot \overleftrightarrow{L}_g + \|\vec{k}\| \vec{R}_g(\vec{k}).$$

This is exactly (9.2.15) if we define $\vec{R}_{g\circ f}(\vec{0}) = 0$ and, for $\vec{h} \neq \vec{0}$, define

$$\vec{R}_{g\circ f}(\vec{h}) = \vec{R}_f(\vec{h})\cdot \overleftrightarrow{L}_g + \frac{\|\vec{k}\|}{\|\vec{h}\|}\vec{R}_g(\vec{k}). \tag{9.2.18}$$

Applying to (9.2.17) the triangle and Schwarz inequalities gives

$$\|\vec{k}\| \leq \|\vec{h}\| \left[\| \overleftrightarrow{L}_f \| + \|\vec{R}_f(\vec{h})\|\right]. \tag{9.2.19}$$

Therefore, as $\vec{h} \to \vec{0}$, we have $\vec{k} \to \vec{0}$, and $\|\vec{k}\|/\|\vec{h}\|$ remains bounded. Hence $\vec{R}_{f\circ g}(\vec{h}) \to \vec{0}$ as $h \to 0$. QED.

Example 9.2.22 Suppose D_f is an open subset of Euclidean space U, and D_g is an open subset of $\mathcal{R}$, and $f : D_f \to D_g$ and $g : D_g \to \mathcal{R}$. Suppose f is differentiable at $\vec{u} \in D_f$ and that g has a derivative, $\partial_v g(v)$, at $v = f(\vec{u})$. Then $g \circ f : D_f \to \mathcal{R}$ is differentiable at $\vec{u} \to$ and

$$\vec{\nabla}(g \circ f)(\vec{u}) = [\partial_v g(v)]\, \vec{\nabla} f(\vec{u}), \quad v = f(\vec{u}). \tag{9.2.20}$$

Proof:

This fact is presumably well known to the reader. It is mentioned here only to show that it follows from the chain rule, example (9.2.21). If we regard $\mathcal{R}$ as a one-dimensional Euclidean space and consider $\vec{f} = f\hat{1}$ and $\vec{g} = g\hat{1}$, then $\vec{f}$ is differentiable at $\vec{u}$, $\vec{g}$ is differentiable at $\vec{v} = v\hat{1}$, so $\vec{g} \circ \vec{f}$ is differentiable at $\vec{u}$ and satisfies 9.2.12. But $\vec{\nabla}\vec{f} = \vec{\nabla} f)\hat{1}$, and $\vec{\nabla}(\vec{g} \circ \vec{f}) = \vec{\nabla}(g \circ f)\hat{1}$, and $\vec{\nabla}\vec{g} = \hat{1}\partial_v \vec{g} = \hat{1}\hat{1}\partial_v g$. Thus 9.2.12 is $\vec{\nabla}(g \circ f)\hat{1} = (\vec{\nabla}\vec{f}) \cdot \hat{1}\hat{1})\partial_v g = (\partial_v g)\vec{\nabla} f\hat{1}$. Dot $\hat{1}$ on the right to obtain 9.2.20.

Example 9.2.23 Suppose D_f is an open subset of $\mathcal{R}$, D_g is an open subset of V, and $\vec{f} : D_f \to D_g$, $\vec{g} : D_g \to W$. Suppose $\vec{f}$ has a derivative, $\partial_u \vec{f}(u)$, at $u \in D_f$, and suppose $\vec{g}$ is differentiable at $\vec{v} = \vec{f}(u)$. Then $\vec{g} \circ \vec{f}$ has a derivative at u and

$$\partial_u(\vec{g} \circ \vec{f})(u) = \partial_u \vec{f}(u) \cdot \vec{\nabla}\vec{g}(\vec{v}), \quad \vec{v} = \vec{f}(u). \tag{9.2.21}$$

Proof:

$\vec{f}$ is differentiable at $\vec{u} = u\hat{1}$, with $\vec{\nabla}\vec{f}(\vec{u}) = \hat{1}\partial_u\vec{f}(u)$. Therefore, by the chain rule, $\vec{g}\circ\vec{f}$ is differentiable. Therefore $\partial_u(\vec{g}\circ\vec{f})$ exists and $\vec{\nabla}(\vec{g}\circ\vec{f}) = \hat{1}\partial_u(\vec{g}\circ\vec{f})$. Substitute these expressions in (9.2.12) and dot $\hat{1}$ on the left to obtain (9.2.21).

Example 9.2.24 Suppose $\vec{f}$ and $\vec{g}$ are inverse to one another, and $\vec{f}$ is differentiable at $\vec{u}$ and $\vec{g}$ is differentiable at $\vec{v} = \vec{f}(\vec{u})$. Then $\vec{\nabla}\vec{f}(\vec{u})$ and $\vec{\nabla}\vec{g}(\vec{v})$ are inverse to one another.

Proof:

Let D_f be the domain of $\vec{f}$, an open subset of Euclidean space U. Let D_g be the domain of $\vec{g}$, an open subset of Euclidean space V. By hypothesis, $\vec{f}\circ\vec{g} = I_{D_g}$ and $\vec{g}\circ\vec{f} = I_{D_f}$. By the chain rule, $\vec{\nabla}\vec{f}(\vec{u})\cdot\vec{\nabla}\vec{g}(\vec{v}) = \vec{\nabla}I_{D_f}(\vec{u}) = \overleftrightarrow{I}_U$ and $\vec{\nabla}\vec{g}(\vec{v})\cdot\vec{\nabla}\vec{f}(\vec{u}) = \vec{\nabla}I_{D_g}(\vec{v}) = \overleftrightarrow{I}_V$. QED.

9.3 Components of gradients

Throughout this chapter, U and V are Euclidean vector spaces, $(\vec{b}_1, \cdots, \vec{b}_m)$ is a fixed basis for U with dual basis $(\vec{b}^1, \cdots, \vec{b}^m)$, and $(\vec{\beta}_1, \cdots, \vec{\beta}_n)$ is a fixed basis for V, with dual basis $(\vec{\beta}^1, \cdots, \vec{\beta}^n)$. D_f is an open subset of U, and $\vec{f} : D_f \to V$.

Definition 9.3.36 The i'th covariant component function of $\vec{f}$ relative to $\vec{\beta}_1, \cdots, \vec{\beta}_n$ is $f_i : D_f \to \mathcal{R}$ where, for every $\vec{u} \in D_f$,

$$f_i(\vec{u}) = \left[\vec{f}(\vec{u})\right]_i = \vec{f}(\vec{u})\cdot\vec{\beta}_i. \tag{9.3.1}$$

The ith contravariant component function of $\vec{f}$ relative to $\vec{\beta}_1, \cdots, \vec{\beta}_n$ is the i'th covariant component function of $\vec{f}$ relative to $\vec{\beta}^1, \cdots \vec{\beta}^n$. That is, it is $f^i : D_f \to \mathcal{R}$ where, for any $\vec{u} \in D_f$,

$$f^i(\vec{u}) = \left[\vec{f}(\vec{u})\right]^i = \vec{f}(\vec{u})\cdot\vec{\beta}^i.$$

Evidently

$$\vec{f} = f^i\vec{\beta}_i = f_i\vec{\beta}^i. \tag{9.3.2}$$

Theorem 9.3.18 Let $\vec{\beta}_1, \cdots, \vec{\beta}_n$ be any basis for V. Let D_f be an open subset of U. Then $\vec{f} : D_f \to V$ is differentiable at $\vec{u} \in D_f$ iff all n covariant component functions $f_i : D_f \to \mathcal{R}$ are differentiable at $\vec{u}$. If they are, then

$$\vec{\nabla} f_i(\vec{u}) = \left[\vec{\nabla}\vec{f}(\vec{u})\right]_i = \vec{\nabla}\vec{f}(\vec{u}) \cdot \vec{\beta}_i \tag{9.3.3}$$

$$\vec{\nabla}\vec{f}(\vec{u}) = \left[\vec{\nabla}\vec{f}(\vec{u})\right]_i \vec{\beta}^i = \vec{\nabla}\vec{f}_i(\vec{u})\vec{\beta}^i. \tag{9.3.4}$$

Proof:

$\Rightarrow$ Suppose $\vec{f}$ is differentiable at $\vec{u}$. Then it satisfies (9.2.4). Dotting $\vec{\beta}_i$ on the right in each term of (9.2.4) produces

$$\lim_{\vec{h}\to\vec{0}} R_i(\vec{h}) = 0$$

$$f_i(\vec{u} + \vec{h}) = f_i(\vec{u}) + \vec{h} \cdot \left[\vec{\nabla}\vec{f}(\vec{u}) \cdot \vec{\beta}_i\right] + \|\vec{h}\| R_i(\vec{h})$$

where $R_i = \vec{R} \cdot \vec{\beta}_i$. By remark 9.2.43, f_i is differentiable at $\vec{u}$, and its gradient is (9.3.3). Then (9.3.4) follows from (D-14) and (9.3.3).

$\Leftarrow$ Suppose all the f_i are differentiable at $\vec{u}$. We can regard $\vec{\beta}^i$ as a constant function on D_f, so it is differentiable at $\vec{u}$. Then by example 9.2.20 the product $f_{(i)}\vec{\beta}^{(i)}$ is differentiable at $\vec{u}$. By example 9.2.17 so is the sum $f_i\vec{\beta}^i = \vec{f}$. QED.

Definition 9.3.37 Let $\vec{b}_1, \cdots, \vec{b}_m$ be any basis for U. Let D_f be an open subset of U and suppose $\vec{f} : D_f \to V$. Then $\vec{f}(u^j\vec{b}_j)$ is a V-valued function of the real variables $u^1, \cdots, u^n$. If the partial derivative $\partial\vec{f}(u^j\vec{b}_j)/\partial u^i$ exists at $\vec{u} = u^j\vec{b}_j$, we abbreviate it as $\partial_{u^i}\vec{f}(\vec{u})$ or $\partial_i\vec{f}(\vec{u})$. Let $\vec{b}^1, \cdots, \vec{b}^m$ be the basis dual to $\vec{b}_1, \cdots, \vec{b}_m$. If the partial derivative $\partial\vec{f}(u_j\vec{b}^j)/\partial u_i$ exists at $\vec{u} = u_j\vec{b}^j$, abbreviate it as $\partial_{u_i}\vec{f}(\vec{u})$ or $\partial^i f(\vec{u})$. In summary

$$\partial_i\vec{f}(\vec{u}) := \partial_{u^i}\vec{f}(\vec{u}) := \frac{\partial\vec{f}(u^j\vec{b}_j)}{\partial u^i} \tag{9.3.5}$$

$$\partial^i\vec{f}(\vec{u}) := \partial_{u_i}\vec{f}(\vec{u}) := \frac{\partial\vec{f}(u_j\vec{b}^j)}{\partial u_i}. \tag{9.3.6}$$

Theorem 9.3.19 Suppose $\vec{b}_1, \cdots \vec{b}_m$ is any basis for U, and D_f is an open subset of U, and $\vec{f} : D_f \to V$. If $\vec{f}$ is differentiable at $\vec{u} \in D_f$ then the partial derivatives $\partial_i \vec{f}(\vec{u})$ and $\partial^i \vec{f}(\vec{u})$ all exist, and

$$\partial_i \vec{f}(\vec{u}) = \vec{b}_i \cdot \vec{\nabla} \vec{f}(\vec{u}) \tag{9.3.7}$$

$$\partial^i \vec{f}(\vec{u}) = \vec{b}^i \cdot \vec{\nabla} \vec{f}(\vec{u}) \tag{9.3.8}$$

$$\vec{\nabla} \vec{f}(\vec{u}) = \vec{b}^i \partial_i \vec{f}(\vec{u}) = \vec{b}_i \partial^i \vec{f}(\vec{u}). \tag{9.3.9}$$

Proof:

By example 9.2.23, if $\vec{f}$ is differentiable at $\vec{u} = u^j \vec{b}_j$ then the partial derivative of $\vec{f}(u^j \vec{b}_j)$ with respect to u^i exists (hold all u^j fixed except for u^i) and is equal to $\left[\partial_i (u^j \vec{b}_j)\right] \cdot \vec{\nabla} \vec{f}(\vec{u})$. But obviously, by the definition of ∂_i, $\partial_i u^j = \delta_i{}^j$ so

$$\partial_i (u^j \vec{b}_j) = \vec{b}_i.$$

Thus we have (9.3.7). 9.3.8 is obtained similarly. To get (9.3.9) use the fact that $\overleftrightarrow{I}_U = \vec{b}^i \vec{b}_i = \vec{b}_i \vec{b}^i$. Then $\vec{\nabla} \vec{f}(\vec{u}) = \overleftrightarrow{I}_U \cdot \vec{\nabla} \vec{f}(\vec{u}) = \vec{b}^i \vec{b}_i \cdot \vec{\nabla} \vec{f}(\vec{u}) = \vec{b}^i \partial_i \vec{f}(\vec{u})$ and $\vec{\nabla} \vec{f}(\vec{u}) = \overleftrightarrow{L}_U \cdot \vec{\nabla} \vec{f}(\vec{u}) = \vec{b}_i \vec{b}^i \cdot \vec{\nabla} \vec{f}(\vec{u}) = \vec{b}_i \partial^i \vec{f}(\vec{u})$.

Note: If $\dim U = 1$, the converse of theorem 9.3.19 is true. See example 9.2.19. However, if $\dim U \geq 2$, that converse is false. A famous example is this. Let $\dim U = 2$, $V = \mathcal{R}$. Let $\hat{x}$, $\hat{y}$ be an orthonormal basis for U and define $f : U \to \mathcal{R}$ by $f(\vec{0}) = 0$, $f(x\hat{x} + y\hat{y}) = xy/\sqrt{x^2 + y^2}$ if $x^2 + y^2 \neq 0$. Then $\partial_x f(\vec{0}) = 0$, $\partial_y f(\vec{0}) = 0$, so if f is differentiable at $\vec{0}$ then $\vec{\nabla} f(\vec{0}) = 0\hat{x} + 0\hat{y} = \vec{0}$. Then by example 9.2.23, $\partial_t f(t\hat{x} + t\hat{y}) = [\partial_t (t\hat{x} + t\hat{y})] \cdot \vec{\nabla} f(\vec{0}) = 0$ at $t = 0$. But $f(t\hat{x} + t\hat{y}) = t$, so $\partial_t f(t\hat{x} + t\hat{y}) = 1$ at $t = 0$. This contradiction shows that f is not differentiable at $\vec{0}$.

Theorem 9.3.20 Suppose $(\vec{b}_1, \cdots, \vec{b}_m)$ and $(\vec{\beta}_1, \cdots, \vec{\beta}_n)$ are bases for U and V respectively, with dual bases $(\vec{b}^1, \cdots, \vec{b}^m)$ and $(\vec{\beta}^1, \cdots, \vec{\beta}^n)$. Suppose D_f is an open subset of U and $\vec{f} : D_f \to V$ and $\vec{u} \in D_f$. Suppose f is differentiable at $\vec{u}$. Then the partial derivatives $\partial_i \vec{f}(\vec{u})$ and $\partial_i f_j(\vec{u})$ exist, and so do the gradients $\vec{\nabla} f_j(\vec{u})$. Moreover

$$\left[\partial_i \vec{f}(\vec{u})\right]_j = \partial_i f_j(\vec{u}) \tag{9.3.10}$$

$$\left[\vec{\nabla} f_j(\vec{u})\right]_i = \partial_i f_j(\vec{u}) \tag{9.3.11}$$

$$\left[\vec{\nabla} \vec{f}(\vec{u})\right]_{ij} = \partial_i f_j(\vec{u}). \tag{9.3.12}$$

Proof:

By theorem 9.3.18, f_j is differentiable at $\vec{u}$ so $\vec{\nabla} f_j(\vec{u})$ exists. By theorem 9.3.19, $\partial_i \vec{f}(\vec{u})$ and $\partial_i f_j(\vec{u})$ exist because $\vec{f}$ and f_j are differentiable at $\vec{u}$. By 9.3.4,

$$\vec{\nabla} \vec{f}(\vec{u}) = \left[\vec{\nabla} f_j(\vec{u})\right] \vec{\beta}^j. \tag{9.3.13}$$

By (9.3.7), $\vec{b}_i \cdot \vec{\nabla} \vec{f}(\vec{u}) = \partial_i \vec{f}(\vec{u})$ and $\vec{b}_i \cdot \vec{\nabla} f_j(\vec{u}) = \partial_i f_j(\vec{u})$. Therefore dotting $\vec{b}_i$ on the left throughout (9.3.13) gives

$$\partial_i \vec{f}(\vec{u}) = \partial_i f_j(\vec{u}) \vec{\beta}^j. \tag{9.3.14}$$

This is equivalent to (9.3.10). Equation (9.3.9) applied to f_j gives

$$\vec{\nabla} f_j(\vec{u}) = \vec{b}^i \partial_i f_j(\vec{u}). \tag{9.3.15}$$

This is equivalent to (9.3.11). Substituting it in (9.3.13) gives

$$\vec{\nabla} \vec{f}(\vec{u}) = \vec{b}^i \partial_i f_j(\vec{u}) \vec{\beta}^j. \tag{9.3.16}$$

This is equivalent to (9.3.12).

9.4 Gradients of dot products

In this chapter we suppose that U, V, W, X are Euclidean spaces, that D is an open subset of U, and that $\overleftrightarrow{f}\colon D \to V \otimes W$ and $\overleftrightarrow{g}\colon D \to W \otimes X$. We suppose that $\overleftrightarrow{f}$ and $\overleftrightarrow{g}$ are differentiable at $\vec{u} \in D$. We want to show that $\overleftrightarrow{f} \cdot \overleftrightarrow{g}\colon D \to V \otimes X$ is differentiable at $\vec{u}$, and we want to calculate $\vec{\nabla}(\overleftrightarrow{f} \cdot \overleftrightarrow{g})(\vec{u})$.

Method 1: No Bases.

By hypothesis,

$$\begin{aligned}\overleftrightarrow{f}(\vec{u}+\vec{h}) &= \overleftrightarrow{f}(\vec{u}) + \vec{h}\cdot\vec{\nabla}\overleftrightarrow{f}(\vec{u}) + \|\vec{h}\|\overleftrightarrow{R}_f(\vec{h})\\ \overleftrightarrow{g}(\vec{u}+\vec{h}) &= \overleftrightarrow{g}(\vec{u}) + \vec{h}\cdot\vec{\nabla}\overleftrightarrow{g}(\vec{u}) + \|\vec{h}\|\overleftrightarrow{R}_g(\vec{h}).\end{aligned}$$

Dotting the first equation into the second gives

$$\begin{aligned}(\overleftrightarrow{f}\cdot\overleftrightarrow{g})(\vec{u}+\vec{h}) &= (\overleftrightarrow{f}\cdot\overleftrightarrow{g})(\vec{u}) + \left[\vec{h}\cdot\vec{\nabla}\overleftrightarrow{f}(\vec{u})\right]\cdot\overleftrightarrow{g}(\vec{u})\\ &\quad + \overleftrightarrow{f}(\vec{u})\cdot\left[\vec{h}\cdot\vec{\nabla}\overleftrightarrow{g}(\vec{u})\right] + \|\vec{h}\|\overleftrightarrow{R}_{f\cdot g}(\vec{h}) \qquad (9.4.1)\end{aligned}$$

where

$$\begin{aligned}\overleftrightarrow{R}_{f\cdot g}(\vec{h}) &= \overleftrightarrow{f}(\vec{u})\cdot\overleftrightarrow{R}_g(\vec{h}) + \left[\vec{h}\cdot\vec{\nabla}\overleftrightarrow{f}(\vec{u})\right]\cdot\left[\vec{h}\cdot\vec{\nabla}\overleftrightarrow{g}(\vec{u})\right]/\|\vec{h}\|\\ &\quad + \left[\vec{h}\cdot\vec{\nabla}\overleftrightarrow{f}(\vec{u})\right]\cdot\overleftrightarrow{R}_g(\vec{h}) + \overleftrightarrow{R}_f(\vec{h})\cdot\overleftrightarrow{g}(\vec{u})\\ &\quad + \overleftrightarrow{R}_f(\vec{h})\cdot\left[\vec{h}\cdot\vec{\nabla}\overleftrightarrow{g}(\vec{u})\right] + \|\vec{h}\|\overleftrightarrow{R}_f(\vec{h})\cdot\overleftrightarrow{R}_g(\vec{h}).\end{aligned}$$

An application of Schwarz's inequality proves that $\overleftrightarrow{R}_{f\cdot g}(\vec{h}) \to \overleftrightarrow{0}$ as $\vec{h}\to\vec{0}$. In (9.4.1) the expression

$$\left[\vec{h}\cdot\vec{\nabla}\overleftrightarrow{f}(\vec{u})\right]\cdot\overleftrightarrow{g}(\vec{u}) + \overleftrightarrow{f}(\vec{u})\cdot\left[\vec{h}\cdot\vec{\nabla}\overleftrightarrow{g}(\vec{u})\right] \qquad (9.4.2)$$

depends linearly on $\vec{h}$, so (9.4.1) shows that $\overleftrightarrow{f}\cdot\overleftrightarrow{g}$ is differentiable at $\vec{u}$ and that $\vec{\nabla}(\overleftrightarrow{f}\cdot\overleftrightarrow{g})(\vec{u})$ is that tensor in $U\otimes V\otimes X$ such that $\vec{h}\cdot\vec{\nabla}(\overleftrightarrow{f}\cdot\overleftrightarrow{g})(\vec{u})$ is equal to (9.4.2) for all $\vec{h}\in U$. But what is $\vec{\nabla}(\overleftrightarrow{f}\cdot\overleftrightarrow{g})(\vec{u})$? We need

Lemma 9.4.23 If $P\in V\otimes W$, $\vec{h}\in U$, and $Q\in U\otimes W\otimes X$ then

$$P\cdot\left[\vec{h}\cdot Q\right] = \vec{h}\cdot\left\{(12)\left[P\cdot(12)Q\right]\right\}.$$

Proof:

Since both sides of this equation are linear in P and Q, it suffices to prove the equation when P and Q are polyads, say $P=\vec{v}\vec{w}$, $Q=\vec{u}\vec{w}'\vec{x}$. Then

$\vec{P} \cdot \left[\vec{h} \cdot Q\right] = (\vec{h} \cdot \vec{u})(\vec{w} \cdot \vec{w}')\vec{v}\vec{x}$, and

$$\begin{aligned}
\vec{h} \cdot \{(12)\,[P \cdot (12)Q]\} &= \vec{h} \cdot \{(12)\,[P \cdot \vec{w}'\vec{u}\vec{x}]\} \\
&= \vec{h}\,\{(12)\,[\vec{v}(\vec{w} \cdot \vec{w}')\vec{u}\vec{x}]\} \\
&= (\vec{w} \cdot \vec{w}')\vec{h} \cdot (12)\,[\vec{v}\vec{u}\vec{x}] \\
&= (\vec{w} \cdot \vec{w}')\vec{h} \cdot [\vec{u}\vec{v}\vec{x}] = (\vec{w} \cdot \vec{w}')(\vec{h} \cdot \vec{u})\vec{v}\vec{x}.
\end{aligned}$$

From lemma (9.4.23) it follows that we can write (9.4.2) as

$$\vec{h} \cdot \left\{\vec{\nabla} \overset{\leftrightarrow}{f} (\vec{u}) \cdot \overset{\leftrightarrow}{g} (\vec{u}) + (12)\left[\overset{\leftrightarrow}{f} (\vec{u}) \cdot 12\vec{\nabla} \overset{\leftrightarrow}{g} (\vec{u})\right]\right\}.$$

Therefore

$$\vec{\nabla}(\overset{\leftrightarrow}{f} \cdot \overset{\leftrightarrow}{g})(\vec{u}) = \vec{\nabla} \overset{\leftrightarrow}{f} (\vec{u}) \cdot \overset{\leftrightarrow}{g} (\vec{u}) + (12)\left\{\overset{\leftrightarrow}{f} (\vec{u}) \cdot \left[(12)\vec{\nabla} \overset{\leftrightarrow}{g} (\vec{u})\right]\right\}. \tag{9.4.3}$$

Method 2:

Introduce bases in U, V, W, X and take components. This is the procedure likely to be used in practical calculations, and the results are generally easier to use than (9.4.3).

Since $\overset{\leftrightarrow}{f}$ and $\overset{\leftrightarrow}{g}$ are differentiable at $\vec{u}$, so are their component functions f_{jk} and $g^k{}_l$. By example (9.2.20), so are the products of $f_{j(k)}g^{(k)}{}_l$. By example (9.2.17) so are the sums $f_{jk}g^k{}_l$, and from examples 9.2.20 and 9.2.17

$$\vec{\nabla}(f_{jk}g^k{}_l) = (\vec{\nabla} f_{jk})g^k{}_l + f_{jk}(\vec{\nabla} g^k{}_l). \tag{9.4.4}$$

Then (9.4.4) and (9.3.11) imply

$$\partial_i(f_{jk}g^k{}_l) = (\partial_i f_{jk})g^k{}_l + f_{jk}(\partial_i g^k{}_l). \tag{9.4.5}$$

Now $f_{jk}g^k{}_l = (\overset{\leftrightarrow}{f} \cdot \overset{\leftrightarrow}{g})_{jl}$ and so (9.4.5) and (9.3.10) imply

$$\left[\partial_i(\overset{\leftrightarrow}{f} \cdot \overset{\leftrightarrow}{g})\right]_{jl} = \left[\partial_i \overset{\leftrightarrow}{f}\right]_{jk} g^k{}_l + f_{jk}\left[\partial_i \overset{\leftrightarrow}{g}\right]^k{}_l;$$

therefore

$$\left[\partial_i(\overset{\leftrightarrow}{f} \cdot \overset{\leftrightarrow}{g})\right]_{jl} = \left[\left(\partial_i \overset{\leftrightarrow}{f}\right) \cdot \overset{\leftrightarrow}{g}\right]_{jl} + \left[\overset{\leftrightarrow}{f} \cdot \left(\partial_i \overset{\leftrightarrow}{g}\right)\right]_{jl},$$

and

$$\partial_i(\overleftrightarrow{f} \cdot \overleftrightarrow{g})(\vec{u}) = \left[\partial_i \overleftrightarrow{f}(\vec{u})\right] \cdot \overleftrightarrow{g}(\vec{u}) + \overleftrightarrow{f}(\vec{u}) \cdot \left[\partial_i \overleftrightarrow{g}(\vec{u})\right]. \tag{9.4.6}$$

If we multiply (9.4.6) on the right by $\vec{b}^i$, where $\vec{b}^1, \cdots \vec{b}^n$ is the dual to the basis we have introduced in U, we obtain

$$\vec{\nabla}(\overleftrightarrow{f} \cdot \overleftrightarrow{g}) = (\vec{\nabla} \overleftrightarrow{f}) \cdot \overleftrightarrow{g} + \vec{b}^i \left[\overleftrightarrow{f} \cdot (\partial_i \overleftrightarrow{g})\right]. \tag{9.4.7}$$

The last terms in (9.4.3) and (9.4.7) can be shown to be equal by lemma 9.4.23. In fact, neither (9.4.7) nor (9.4.3) is very useful. One usually works with (9.4.6) or (9.2.10).

The hoped for formula, $\vec{\nabla}(\overleftrightarrow{f} \cdot \overleftrightarrow{g}) = \vec{\nabla}(\overleftrightarrow{f}) \cdot \overleftrightarrow{g} + (\overleftrightarrow{f}) \cdot (\vec{\nabla} \overleftrightarrow{g})$, is generally false. It is true if one of $\vec{\nabla}$ or $\overleftrightarrow{f}$, is one-dimensional. If U is one dimensional, (9.4.6) reduces to

$$\partial_u(\overleftrightarrow{f} \cdot \overleftrightarrow{g})(u) = \partial_u \overleftrightarrow{f}(u) \cdot \overleftrightarrow{g}(u) + \overleftrightarrow{f}(u) \cdot \partial_u \overleftrightarrow{g}(u). \tag{9.4.8}$$

If $\overleftrightarrow{f}$ is one dimensional, (9.2.10) is

$$\vec{\nabla}(f \overleftrightarrow{g})(u) = \vec{\nabla} f(\vec{u}) \overleftrightarrow{g}(u) + f(u) \vec{\nabla} \overleftrightarrow{g}(\vec{u}). \tag{9.4.9}$$

9.5 Curvilinear coordinates

The set $\mathcal{R}^n$ of real n-tuples $\vec{u} = (u^1, \cdots, u^n)$ is a vector space if, for $\vec{u}$ and $\vec{v} = (v^1, \cdots, v^n) \in \mathcal{R}$ we define $\vec{u} + \vec{v} = (u^1 + v^1, \cdots, u^n + v^n)$ and, for $\vec{u} \in \mathcal{R}^n$ and $a \in \mathcal{R}$ we define $a\vec{u} = (au^1, \cdots, au^n)$. $\mathcal{R}^n$ becomes a Euclidean vector space if we define $\vec{u} \cdot \vec{v} = u^i v^i$. An orthonormal basis for $\mathcal{R}^n$ is $\hat{e}_1 = (1, 0, \cdots, 0)$, $\hat{e}_2 = (0, 1, 0, \cdots, 0)$, $\cdots \hat{e}_n = (0, \cdots, 0, 1)$.

Definition 9.5.38 Suppopse D_V is an open subset of Euclidean space V. A "coordinate system" on D_V is an open subset D_U of $\mathcal{R}^n$, together with an everywhere differentiable bijection $\vec{\mathcal{V}} : D_U \to D_V$ whose inverse mapping $\vec{\mathcal{C}} : D_V \to D_U$ is also everywhere differentiable. The real number $\mu_i = c^i(\vec{v}) = \hat{e}_i \cdot \vec{\mathcal{C}}(\vec{v})$ is the "value of the i'th coordinate at the point $\vec{v}$, and $\vec{v}$ is uniquely determined by the values of its n coordinates $u^1, \cdots, u^n$.

We have

$$\begin{aligned}\vec{v}(u^1,\cdots,u^n) &= \vec{\mathcal{V}}(u^i\hat{e}_i) = \vec{\mathcal{V}}(\vec{u})\\ u^i &= c^i(\vec{v}) = \hat{e}_i \cdot \vec{\mathcal{C}}(\vec{v}).\end{aligned}$$

Definition 9.5.39 Suppose W is a Euclidean space and $\vec{g} : D_U \to W$ is differentiable. Then we will write $\partial_i \vec{g}$ for the partial derivative with respect to u^i. That is

$$\partial_i \vec{g}(\vec{u}) = \frac{\partial \vec{g}(u^j\hat{e}_j)}{\partial u^i}. \tag{9.5.1}$$

If $\vec{f} : D_V \to W$ is differentiable, then so is $\vec{f} \circ \vec{\mathcal{V}} : D_U \to W$ (by the chain rule). We will abbreviate $\partial_i(\vec{f} \circ \mathcal{V})$ as $\partial_i \vec{f}$. Thus

$$\partial_i \vec{f}(\vec{v}) = \frac{\partial \vec{f}[\vec{\mathcal{V}}(u^1,\cdots,u^n)]}{\partial u^i} \tag{9.5.2}$$

if $\vec{v} = \vec{\mathcal{V}}(u^1,\cdots,u^n)$.

Note that by theorem 9.3.18, the coordinate functions $c^i : D_V \to \mathcal{R}$ are differentiable at every $\vec{v} \in D_V$. And by theorem 9.3.19 the partial derivatives $\partial_i \vec{\mathcal{V}}$ exist at every $\vec{u} \in D_U$. The key to understanding vector and tensor fields in curvilinear coordinates is

Theorem 9.5.21 Suppose $\vec{\mathcal{V}} : D_U \to D_V$ is a curvilinear coordinate system on open subset D_V of Euclidean vector space V. Suppose $n = \dim U$ is the number of coordinates. Then $\dim V = n$. For any $\vec{v} \in D_V$, let $\vec{u} = \mathcal{C}(\vec{v})$. Then the two n-tuples of vectors in V, $(\partial_1\vec{\mathcal{V}})(\vec{u}),\cdots,\partial_n\vec{\mathcal{V}}(\vec{u})$ and $(\vec{\nabla}c^1(\vec{v}),\cdots,\vec{\nabla}c^n(\vec{v}))$, are dual bases for V.

Proof:

By the definition of a coordinate system, $\vec{\mathcal{C}} \circ \vec{\mathcal{V}} = I_U|D_U$. By the chain rule, $\vec{\nabla}(\vec{\mathcal{C}} \circ \vec{\mathcal{V}})(\vec{u}) = \vec{\nabla}\vec{\mathcal{V}}(\vec{u}) \cdot \vec{\nabla}\vec{\mathcal{C}}(\vec{v})$. By example 9.2.15, $\vec{\nabla}(I_U|D_U) = \overleftrightarrow{I}_U$. Hence,

$$\vec{\nabla}\vec{\mathcal{V}}(\vec{u}) \cdot \vec{\nabla}\vec{\mathcal{C}}(\vec{v}) = \overleftrightarrow{I}_U\ . \tag{9.5.3}$$

Similarly, $\vec{\mathcal{V}} \circ \vec{\mathcal{C}} = I_V|D_V$, so

$$\vec{\nabla}\vec{\mathcal{C}}(\vec{v}) \cdot \vec{\nabla}\vec{\mathcal{V}}(\vec{u}) = \overleftrightarrow{I}_V\ . \tag{9.5.4}$$

Now $\vec{\nabla}\vec{\mathcal{V}}(\vec{u}) \in U \otimes V$ so $\vec{\nabla}\vec{\mathcal{V}}(\vec{u}) = \overleftrightarrow{L}$ for some $L \in \mathcal{L}(U \to V)$, and $\vec{\nabla}\vec{\mathcal{C}}(\vec{v}) \in V \otimes U$, so $\vec{\nabla}\vec{\mathcal{C}}(\vec{v}) = \overleftrightarrow{M}$ for some $M \in \mathcal{L}(V \to U)$. According to (9.5.3) and (7.4.20), $\overleftrightarrow{L} \cdot \overleftrightarrow{M} = \overleftrightarrow{I}_U$ so

$$M \circ L = I_U. \tag{9.5.5}$$

According to (9.5.4) and (7.4.20), $\overleftrightarrow{M} \cdot \overleftrightarrow{L} = \overleftrightarrow{I}_V$ so

$$L \circ M = I_V. \tag{9.5.6}$$

According to (D-1), (9.5.5) and (9.5.6) together imply that L and M are bijections and $M = L^{-1}$. Therefore U and V are isomorphic. Hence they have the same dimension.

Lemma 9.5.24 If U and V are finite dimensional vector spaces and $L : U \to V$ is a linear bijection, and $\vec{u}_1, \cdots, \vec{u}_n$ is a basis for U, then $L(\vec{u}_1), \cdots, L(\vec{u}_n)$ is a basis for V.

Proof of lemma:

Suppose $a_i L(\vec{u}_i) = \vec{0}$ for $a_1, \cdots, a_n \in \mathcal{R}$. Then $L(a_i\vec{u}_i) = \vec{0}$. Since L is a bijection, $a_i\vec{u}_i = \vec{0}$. Since $\vec{u}_1, \cdots, \vec{u}_n$ are linearly independent, $a_i = \cdots = a_n = 0$. Hence $L(\vec{u}_1), \cdots, L(\vec{u}_n)$ are linearly independent. Next, let $\vec{v} \in V$. There is a $\vec{u} \in U$ such that $\vec{v} = L(\vec{u})$. There are $a_1, \cdots, a_n \in \mathcal{R}$ such that $\vec{u} = a_i\vec{u}_i$. Then $\vec{v} = L(\vec{u}) = L(a_i\vec{u}_i) = a_i L(\vec{u}_i)$. Hence $L(\vec{u}_1), \cdots, L(\vec{u}_n)$ span V.

Now by (9.3.8), $\hat{e}_i \cdot \vec{\nabla}\vec{\mathcal{V}}(\vec{u}) = \partial_i\vec{\mathcal{V}}(\vec{u})$. By (9.3.3) , $\vec{\nabla}\vec{\mathcal{C}}(\vec{v}) \cdot \hat{e}_j = \vec{\nabla}c^j(\vec{v})$. And $\hat{e}_i \cdot \overleftrightarrow{I}_U \cdot \hat{e}_j = \hat{e}_i \cdot \hat{e}_j = \delta_i{}^j$. Therefore if we dot $\hat{e}_i$ on the left and $\hat{e}_j$ on the right on each side of (9.5.3) we obtain

$$\partial_i\vec{\mathcal{V}}(\vec{u}) \cdot \vec{\nabla}c^j(\vec{v}) = \delta_i{}^j. \tag{9.5.7}$$

Thus, the two n-tuples of vectors $(\partial_1\vec{\mathcal{V}}(\vec{u}), \cdots, \partial_n\vec{\mathcal{V}}(\vec{u}))$ and $(\vec{\nabla}c^1(\vec{v}), \cdots, \vec{\nabla}c^n(\vec{v}))$ are dual to one-another. Hence they are linearly independent. Since $\dim V = n$, they span V. Hence they are bases for V, and each is the basis dual to the other.[1]

[1] A more intuitive way to put the proof of (9.5.7) is as follows. $u^j = c^j(\vec{\mathcal{V}}(u^1, \cdots, u^n))$ so $\delta_i{}^j = \partial_i u^j = \partial_i[c^j(\vec{\mathcal{V}}(u^1, \cdots, u^n))] = \partial_i\vec{\mathcal{V}} \cdot \vec{\nabla}c^j$.

Corollary 9.5.31 If $\vec{f} : D_V \to W$ is differentiable (W being a Euclidean space) then

$$\vec{\nabla}\vec{f}(\vec{v}) = \vec{\nabla}c^i(\vec{v})\partial_i\vec{f}(\vec{v}). \tag{9.5.8}$$

Proof:

(9.5.8) is (9.3.9) with $\vec{b}^i = \vec{\nabla}c^i(\vec{v})$.

A function $\vec{f} : D_V \to W$ is often called a vector field on D_V. If W is a tensor product of other vector spaces, $\vec{f}$ is called a tensor field. It is often convenient to express $\vec{f}$ in terms of a basis for W which varies from point to point in D_V. Suppose that for $i \in \{1, \cdots, N\}$, $\beta_i : D_V \to W$ is differentiable everywhere in D_V. Suppose also that at each $\vec{v} \in D_V$, $\vec{\beta}_1(\vec{v}), \cdots, \vec{\beta}_N(\vec{v})$ is a basis for W, with dual basis $\vec{\beta}^1(\vec{v}), \cdots, \vec{\beta}^N(\vec{v})$. Then $\vec{\beta}^i : D_V \to W$ is also differentiable in D_V. At each $\vec{v} \in D_V$ we can write

$$\vec{f}(\vec{v}) = f^j(\vec{v})\vec{\beta}_j(\vec{v}) = f_j(\vec{v})\vec{\beta}^j(\vec{v}).$$

Then, by (9.5.8),

$$\begin{aligned} \vec{\nabla}\vec{f}(\vec{v}) &= (\vec{\nabla}c^i)\partial_i\left[f^j\vec{\beta}_j\right] \\ &= (\vec{\nabla}c^i)(\partial_i f^j)\vec{\beta}_j + f^j(\vec{\nabla}c^i)(\partial_i\vec{\beta}_j). \end{aligned}$$

Now $\partial_i\vec{\beta}j(\vec{v})$ is a vector in W, so there are coefficients $\Gamma_i{}^j{}_k(\vec{v}) \in \mathcal{R}$ such that

$$\partial_i\vec{\beta}_k(\vec{v}) = \Gamma_i{}^j{}_k(\vec{v})\vec{\beta})_j(\vec{v}). \tag{9.5.9}$$

Then

$$\vec{\nabla}\vec{f}(\vec{v}) = \left[\partial_i f^j + \Gamma_i{}^j{}_k f^k\right](\vec{\nabla}c^i)\vec{\beta}_j(\vec{v}). \tag{9.5.10}$$

Equation (9.5.9) is called the "connection formula" for the basis $\vec{\beta}^1(\vec{v}), \cdots, \vec{\beta}^N(\vec{v})$ in the coordinate system $\vec{\mathcal{V}} : D_U \to V$. The scalars $\Gamma_i{}^j{}_k(\vec{v})$ are the "Christoffel symbols" at $\vec{v}$. The expression $\partial_i f^j + \Gamma_i{}^j{}_k f^k$ is often abbreviated $D_i f^j$ and, in the older literature, is called the covariant derivative of the vector field $\vec{f} : D_V \to W$.

If $W = \otimes^q V$, then $(\vec{\beta}^1, \cdots, \vec{\beta}^N)$ is often taken to be a q'th order polyad basis constructed from $(\vec{\nabla} c^1, \cdots, \vec{\nabla} c^n)$ and $(\partial_1 \vec{\mathcal{V}}, \cdots \partial_n \vec{\mathcal{V}})$. For example, if $W = V \otimes V$ and $\vec{f} : D_V \to V \otimes V$, we might write

$$\overleftrightarrow{f} = f^{ij} \partial_i \vec{\mathcal{V}} \partial_j \vec{\mathcal{V}}. \tag{9.5.11}$$

The Christoffel symbols $\Gamma_i{}^j{}_k$ for $\partial_i \vec{\mathcal{V}}, \cdots, \partial_n \vec{\mathcal{V}}$ are defined by

$$\partial_i (\partial_k \vec{\mathcal{V}}) = \Gamma_i{}^j{}_k (\partial_j \vec{\mathcal{V}}). \tag{9.5.12}$$

If we use $(\partial_k \vec{\mathcal{V}})(\partial_l \vec{\mathcal{V}})$ as the basis $\vec{\beta}_1, \cdots, \vec{\beta}_N$ for $V \otimes V$, we have from (9.5.12)

$$\partial_i \left[(\partial_k \vec{\mathcal{V}})(\partial_l \vec{\mathcal{V}}) \right] = \Gamma_i{}^j{}_k (\partial_j \vec{\mathcal{V}})(\partial_l \vec{\mathcal{V}}) + \Gamma_i{}^j{}_l (\partial_k \vec{\mathcal{V}})(\partial_j \vec{\mathcal{V}}). \tag{9.5.13}$$

This is the connection formula (9.5.9) for $(\partial_k \vec{\mathcal{V}})(\partial_l \vec{\mathcal{V}})$. It leads to

$$\vec{\nabla} \overleftrightarrow{f} = (D_i f^{jk})(\vec{\nabla} c^i)(\partial_j \vec{\mathcal{V}})(\partial_k \vec{\mathcal{V}}) \tag{9.5.14}$$

where

$$D_i f^{jk} := \partial_i f^{jk} + \Gamma_i{}^j{}_l f^{lk} + \Gamma_i{}^k{}_l f^{jl}. \tag{9.5.15}$$

Note that (9.5.12) implies

$$\Gamma_i{}^j{}_k = \Gamma_k{}^j{}_i. \tag{9.5.16}$$

In general this will not be true for $\Gamma_i{}^j{}_k$ in (9.5.9).

Instead of (9.5.11), we might prefer to write

$$\overleftrightarrow{f} = f_{ij} (\vec{\nabla} c^i)(\vec{\nabla} c^j). \tag{9.5.17}$$

Then we need the connection formula for $(\vec{\nabla} c^k)(\vec{\nabla} c^l)$ analogous to (9.5.13). Therefore we need the analogue of (9.5.12) for $\vec{\nabla} c^k$. The Christoffel symbols introduced by that analogue will not be new. From (9.5.7), $\vec{\nabla} c^j \cdot \partial_k \vec{\mathcal{V}} = \delta^j{}_k$. Then $\partial_i (\vec{\nabla} c^j) \cdot \partial_k \vec{\mathcal{V}} + \vec{\nabla} c^j \cdot \partial_i (\partial_k \vec{\mathcal{V}}) = 0$. From (9.5.12),

$\partial_i(\vec{\nabla}c^j) \cdot \partial_k\vec{\mathcal{V}} = -\Gamma_i{}^l{}_k\vec{\nabla}c^j \cdot \partial_l\vec{\mathcal{V}} = -\Gamma_i{}^l{}_k\delta^j{}_l = -\Gamma_i{}^j{}_k$. Thus $-\Gamma_i{}^j{}_k\vec{\nabla}c^k = \partial_i(\vec{\nabla}c^j) \cdot (\partial_k\vec{\mathcal{V}})(\vec{\nabla}c^k) = \partial_i(\vec{\nabla}c^j) \cdot \overleftrightarrow{I}_V = \partial_i(\vec{\nabla}c^j)$. Hence,

$$\partial_i(\vec{\nabla}c^j) = -\Gamma_i{}^j{}_k(\vec{\nabla}c^k). \tag{9.5.18}$$

Therefore,

$$\vec{\nabla}\overleftrightarrow{f}(\vec{u}) = (D_i f_{jk})(\vec{\nabla}c^i)(\vec{\nabla}c^j)(\vec{\nabla}c^k) \tag{9.5.19}$$

where

$$D_i f_{jk} = \partial_i f_{jk} - \Gamma_i{}^l{}_j f_{lk} - \Gamma_i{}^l{}_k f_{jl}. \tag{9.5.20}$$

9.6 Multiple gradients and Taylor's formula

Let U and V be Euclidean spaces with fixed bases, $(\vec{b}_1, \cdots, \vec{b}_m)$ and $(\vec{\beta}_1, \cdots, \vec{\beta}_n)$ respectively. Take all components with respect to these bases. Let D be an open subset of U and suppose $\vec{f} : D \to V$ is differentiable at each $\vec{u} \in D$. Then $\vec{\nabla}\vec{f}(\vec{u}) \in U \otimes V$ for each $\vec{u} \in U$, so we can define a function

$$\vec{\nabla}\vec{f} : D \to U \otimes V. \tag{9.6.1}$$

This function may itself be differentiable everywhere in V, in which case we have another function.

$$\vec{\nabla}\vec{\nabla}\vec{f} : D \to U \otimes U \otimes V. \tag{9.6.2}$$

The process can continue. If $\vec{\nabla}^p\vec{f} : D \to (\otimes^p U) \otimes V$ exists, we say $\vec{f}$ is p times differentiable in D.

Now suppose $\vec{f} : D \to V$ is P times differentiable in D. For any $\vec{u} \in D$ write $\vec{u} = u^i\vec{b}_i$. Let $\partial_i = \partial/\partial^i$. Then

$$\begin{aligned}
\vec{\nabla}\vec{f} &= \vec{b}^i\partial_i\vec{f} \\
\vec{\nabla}\vec{\nabla}\vec{f} &= \vec{\nabla}\left[\vec{b}^i\partial_i\vec{f}\right] = \vec{b}^j\partial_j\left[\vec{b}^i\partial_i\vec{f}\right] \\
&= \vec{b}^j\vec{b}^i\partial_j\partial_i\vec{f} \\
&\vdots \\
\vec{\nabla}^p\vec{f} &= \vec{b}^{i_1}\cdots\vec{b}^{i_p}\partial_{i_1}\cdots\partial_{i_p}\vec{f}.
\end{aligned}$$

Thus, for any $\vec{h} \in U$, if we write

$$\vec{h}^p = \vec{h}\vec{h}\cdots\vec{h}$$

(p factors) we have

$$\vec{h}^p\langle p\rangle\vec{\nabla}^p\vec{f} = h^{i_1}\cdots h^{i_p}\partial_{i_1}\cdots\partial_{i_p}\vec{f}.$$

Fix $\vec{h} \in U$ so that $\vec{u}+\vec{h} \in D$. Let

$$\vec{g}(t) = \vec{f}(\vec{u}+t\hat{h}) \tag{9.6.3}$$

where

$$\hat{h} = \vec{h}/\|\vec{h}\|.$$

By the ordinary Taylor formula in one independent variable

$$\vec{g}(t) = \sum_{p=0}^{P}\frac{t^p}{p!}\partial_g^p(0) + |t|^P\vec{R}_P(T)$$

where $\vec{R}_P(t) \to 0$ as $t \to 0$. But

$$\begin{aligned}
\partial_t\vec{g}(t) &= \hat{h}\cdot\hat{\nabla}\vec{f}(\vec{u}+t\hat{h}) = \hat{h}^i\partial_i\vec{f}(\vec{u}+t\hat{h})\\
\partial_t^2\vec{g}(t) &= (\hat{h}\cdot\vec{\nabla})\left[\hat{h}^i\partial_i\vec{f}(\vec{u}+t\hat{h})\right]\\
&= \hat{h}^j\partial_j\left[\hat{h}^i\partial_i\vec{f}(\vec{u}+t\hat{h})\right]\\
&= \hat{h}^j\hat{h}^i\partial_j\partial_i\vec{f}(\vec{u}+t\hat{h})\\
&= (\hat{h})^2\langle 2\rangle\vec{\nabla}^2\vec{f}(\vec{u}+t\hat{h}).
\end{aligned}$$

Similarly $\partial_t^p\vec{g}(t) = (\hat{h})^p\langle p\rangle\vec{\nabla}^p\vec{f}(\vec{u}+t\hat{h})$ so

$$\partial_t^p\vec{g}(0) = (\hat{h})^p\langle p\rangle\vec{\nabla}^p\vec{f}(\vec{u}),$$

and

$$t^p\partial_t^p\vec{g}(0) = (t\hat{h})^p\langle p\rangle\vec{\nabla}^p\vec{f}(\vec{u}).$$

Thus,

$$\vec{g}(t) = \sum_{p=0}^{P}\frac{1}{p!}(t\hat{h})^p\langle p\rangle\vec{\nabla}^p\vec{f}(\vec{u}) + |t|^p\vec{R}_p(t).$$

Setting $t = \|\vec{h}\|$ in this formula gives

$$\begin{aligned}\vec{f}(\vec{u}+\vec{h}) &= \sum_{p=0}^{P} \frac{1}{p!}(\vec{h})^p\langle p\rangle(\vec{\nabla})^p\vec{f}(\vec{u}) + \|\vec{h}\|^p\vec{R}_p(t)\\ &= \sum_{p=0}^{P} \frac{1}{p!}h^{i_1}\cdots h^{i_p}\partial_{i_1}\cdots\partial_{i_p}\vec{f}(\vec{u}) + \|\vec{h}\|^p\vec{R}_p(\vec{h})\end{aligned} \tag{9.6.4}$$

with $\vec{R}_p(\vec{h}) \to 0$ as $\vec{h} \to \vec{0}$.

9.7 Differential identities

Definition 9.7.40 Suppose U and V are Euclidean spaces, D is an open subset of U, and $\overset{\leftrightarrow}{T}: D \to U \otimes V$ is differentiable at $\vec{u} \in D$. Then $\vec{\nabla}\overset{\leftrightarrow}{T}(\vec{u}) \in U \otimes U \otimes V$. We define the divergence of $\overset{\leftrightarrow}{T}$ at $\vec{u}$ to be

$$\vec{\nabla}\cdot\overset{\leftrightarrow}{T}(\vec{u}) = \overset{\leftrightarrow}{I}_U\langle 2\rangle\vec{\nabla}\overset{\leftrightarrow}{T}(\vec{u}). \tag{9.7.1}$$

Remark 9.7.44 Suppose $\vec{b}_1, \cdots, \vec{b}_m$ is a fixed basis for U and $\vec{\beta}_1, \cdots, \vec{\beta}_n$ is a fixed basis for V. Write $\vec{u} = u^i\vec{b}_i$, $\partial_i = \frac{\partial}{\partial u^i}$, $\overset{\leftrightarrow}{T} = T^{ik}\vec{b}_i\vec{\beta}_k$. Then

$$\vec{\nabla}\cdot\overset{\leftrightarrow}{T}(\vec{u}) = \partial_i T^{ik}\vec{\beta}_k \text{ or, equivalently,} \tag{9.7.2}$$

$$\left[\vec{\nabla}\cdot\overset{\leftrightarrow}{T}(\vec{u})\right]^k = \partial_i T^{ik}. \tag{9.7.3}$$

Proof:

In theorem 9.3.20, replace V by $U \otimes V$ and $\vec{f}: D \to V$ by $\overset{\leftrightarrow}{T}: D \to U \otimes V$. For $U \otimes V$ use the basis $\vec{b}_j\vec{\beta}_k$. Then by (9.3.16) $\vec{\nabla}\overset{\leftrightarrow}{T}(\vec{u})\partial_i T^{jk}\vec{b}^i\vec{b}_j\vec{\beta}_k$. Then $\vec{\nabla}\cdot\overset{\leftrightarrow}{T}(\vec{u}) = \overset{\leftrightarrow}{I}_U\langle 2\rangle\vec{\nabla}\overset{\leftrightarrow}{T}(\vec{u}) = \overset{\leftrightarrow}{I}_U\langle 2\rangle\left[\partial_i T^{jk}\vec{b}^i\vec{b}_j\vec{\beta}_n\right] = \partial_i T^{jk}\left[\overset{\leftrightarrow}{I}_U\langle 2\rangle\vec{b}^i\vec{b}_j\right]\vec{\beta}_k = \partial_i T^{jk}(\vec{b}^i\cdot\vec{b}_j)\vec{\beta}_k, = \partial_i T^{jk}\delta^i{}_j\vec{\beta}_k = \partial_i T^{ik}\vec{\beta}_k$.

Corollary 9.7.32 Suppose $f: D \to U$ and $\vec{g}: D \to V$. Then

$$\vec{\nabla}\cdot(\vec{f}\vec{g}) = (\vec{\nabla}\cdot\vec{f})\vec{g} + \vec{f}\cdot(\vec{\nabla}\vec{g}). \tag{9.7.4}$$

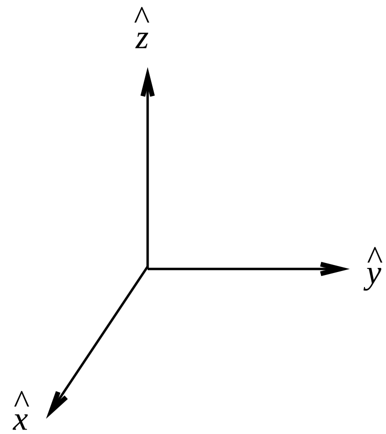


Figure 9.3:

Proof:

Use components. Then, because of (9.7.2), (9.7.4) is equivalent to $\partial_i(f^i g^k) = (\partial_i f^i)g^k + f^i(\partial_i g^k)$. In this form it is obvious.

Definition 9.7.41 Suppose (U, A) is an oriented *three-dimensional* Euclidean space and V is another Euclidean space. Suppose D is an open subset of U and $\overleftrightarrow{T}: D \to U \otimes V$ is differentiable at $\vec{u} \in D$. Then the A-curl of $\overleftrightarrow{T}$ at $\vec{u}$, written $\vec{\nabla}^A \times \overleftrightarrow{T}(\vec{u})$, or simply $\vec{\nabla} \times \overleftrightarrow{T}(\vec{u})$, is defined to be

$$\vec{\nabla} \times \overleftrightarrow{T}(\vec{u}) = A\langle 2\rangle \vec{\nabla} \overleftrightarrow{T}(\vec{u}). \tag{9.7.5}$$

Note 9.7.5 There are two curls for $\overleftrightarrow{T}: D \to U \otimes V$, one for each of the two unimodular alternating tensors over U. We will always choose the right handed A defined on page 61 so Figure 9.3 is positively oriented. The other curl is the negative of the one we use, because the two A's differ only in sign.

Remark 9.7.45 Suppose $(\vec{b}_1, \cdots, \vec{b}_m)$ is a basis for U and $(\vec{\beta}_1, \cdots \vec{\beta}_n)$ is a basis for V and $\overleftrightarrow{T} = T_{kl}\vec{b}^k\vec{\beta}^l$. Then

$$\vec{\nabla} \times \overleftrightarrow{T} = \left[A^{ijk}\partial_j T_{kl}\right]\vec{b}_i\vec{\beta}^l, \tag{9.7.6}$$

or equivalently,

$$\left(\vec{\nabla} \times \overleftrightarrow{T}\right)^i{}_l = A^{ijk}\partial_j T_{kl}. \tag{9.7.7}$$

Proof:

$(\vec{\nabla} \overleftrightarrow{T})_{jkl} = \partial_j T_{kl}$ and $(A\langle 2\rangle \vec{\nabla} \overleftrightarrow{T})^i{}_l = A^{ijk}(\vec{\nabla} \overleftrightarrow{T})_{jkl}$.

Corollary 9.7.33 If $\vec{f} : D \to V$ is twice continuously differentiable on D (i.e. $\vec{\nabla}\vec{\nabla}\vec{f} : D \to U \otimes U \otimes V$ exists and is continuous at every $\vec{u} \in D$) then

$$\vec{\nabla} \times \vec{\nabla} \vec{f} = \overleftrightarrow{0} . \tag{9.7.8}$$

Proof:

$(\vec{\nabla} \times \vec{\nabla} \vec{f})^i{}_l = A^{ijk}\partial_j(\vec{\nabla}\vec{f})_{kl} = A^{ijk}\partial_j\partial_k f_l$. Because of the continuity assumption, $\partial_j\partial_k = \partial_k\partial_j$ so $A^{ijk}\partial_j\partial_k f_l = A^{ijk}\partial_k\partial_j f_l = A^{ikj}\partial_j\partial_k f_l = -A^{ijk}\partial_j\partial_k f_l = 0$.

Corollary 9.7.34 If $\overleftrightarrow{T}: D \to U \otimes V$ is twice continuously differentiable on D then

$$\vec{\nabla} \cdot (\vec{\nabla} \times \overleftrightarrow{T}) = 0. \tag{9.7.9}$$

Proof:

$\left[\vec{\nabla} \cdot \vec{\nabla} \times \overleftrightarrow{T}\right]_l = \partial_i(\vec{\nabla} \times \overleftrightarrow{T})^i{}_l = \partial_i A^{ijk}\partial_j T_{kl} = A^{ijk}\partial_i\partial_j T_{kl}$.

Again this $= 0$ because $\partial_i\partial_j = \partial_j\partial_i$ and $A^{ijk} = -A^{ikj}$.

Corollary 9.7.35 If $\overleftrightarrow{T}: D \to U \otimes V$ is twice continuously differentiable then $\vec{\nabla} \times (\vec{\nabla} \times \overleftrightarrow{T}) = \vec{\nabla}(\vec{\nabla} \cdot \overleftrightarrow{T}) - \vec{\nabla} \cdot (\vec{\nabla} \overleftrightarrow{T})$. (N.B. $\vec{\nabla} \cdot (\vec{\nabla} \overleftrightarrow{T})$ is often written $\nabla^2 \overleftrightarrow{T}$.)

Proof:

Let $(\vec{b}_1, \vec{b}_2, \vec{b}_3)$ be orthonormal and positively oriented, so (9.7.7) reads $(\vec{\nabla}\times \overset{\leftrightarrow}{T})_{il} = \varepsilon_{ijk}\partial_j T_{kl}$. Then

$$\begin{aligned}
\left[\vec{\nabla} \times (\vec{\nabla}\times \overset{\leftrightarrow}{T})\right]_{il} &= \varepsilon_{ijk}\partial_j(\vec{\nabla}\times \overset{\leftrightarrow}{T})_{kl} \\
&= \varepsilon_{ijk}\partial_j\left[\varepsilon_{kmn}\partial_m T_{nl}\right] \\
&= \varepsilon_{ijk}\varepsilon_{kmn}\partial_j\partial_m T_{nl} = \left(\delta_{im}\delta_{jn} - \delta_{in}\delta_{jm}\right)\partial_j\partial_m T_{nl} \\
&= \partial_j\partial_i T_{jl} - \partial_j\partial_j T_{il} = \partial_i(\partial_j T_{jl}) - \partial_j(\partial_j T_{il}) \\
&= \partial_i(\vec{\nabla}\cdot \overset{\leftrightarrow}{T})_l - \partial_j(\vec{\nabla}\, \overset{\leftrightarrow}{T})_{jil} \\
&= \left[\vec{\nabla}(\vec{\nabla}\cdot \overset{\leftrightarrow}{T})\right]_{il} - \left[\vec{\nabla}\cdot(\vec{\nabla}\, \overset{\leftrightarrow}{T})\right]_{il}.
\end{aligned}$$

Thus, $\left[\vec{\nabla} \times (\vec{\nabla}\times \overset{\leftrightarrow}{T})\right]_{il} = \left[\vec{\nabla}(\vec{\nabla}\cdot \overset{\leftrightarrow}{T}) - \vec{\nabla}\cdot(\vec{\nabla}\, \overset{\leftrightarrow}{T})\right]_{il}$. QED.

Chapter 10

Integral Calculus of Tensors

10.1 Definition of the Integral

10.1.1 Mass distributions or measures

Let D be any subset of Euclidean space U. A function m which assigns real numbers to certain subsets of D is called a mass distribution, or measure, on D if it has these properties: (if S is a subset of D to which m does assign a mass, that mass is written $m(S)$, and S is called "measurable-m")

i) If S is measurable-m, then $m(S) \geq 0$,

ii) D and the empty set are measurable-m.

iii) If $S_1, S_2, \cdots$ is any finite or infinite sequence of sets which are measurable-m , then $S_1 \cap S_2$, $S_1 / \, S_2$ and $S_1 \cup S_2 \cup \cdots$ are measurable-m ; and $m(S_1 \cup S_2 \cup \cdots) = \sum_\nu m(S_\nu)$ if no S_i and S_j have common points (i.e., if $S_1, S_2, \cdots$ are mutually disjoint.)

The general theory of mass distributions is discussed in Halmos, "Measure Theory," Van Nostrand, 1950. An example of a "mass distribution" is $(dim\, U)$- dimensional volume. Some subsets of D are so irregular that a volume cannot be defined for them. All bounded open sets D do have volumes. As with volume, any mass distribution can be extended to permit infinite masses. The reader will be presumed to be familiar with the following mass distributions (but not perhaps with the general theory).

Example 10.1.25 D is an open subset of U. $\rho(\vec{u})$ is the density of mass per unit of $(\dim U)$-dimensional volume at $\vec{u}$, and the mass in any measurable set S is $m(S) = \int_S dV(\vec{u})\rho(\vec{u})$ where $dV(\vec{u})$ is the infinitesimal element of $(\dim U)$-dimensional volume in U.

Example 10.1.26 D is a $(\dim U) - 1$ dimensional "surface" in U. $\sigma(\vec{u})$ is the density of mass per unit of $(\dim U) - 1$ dimensional "area" at $\vec{u}$, $dA(\vec{u})$ is the infinitesimal element of such area at $\vec{u}$, and the mass in any measurable set S is $m(S) = \int_S dA(\vec{u})\sigma(\vec{u})$.

Example 10.1.27 D is a curve in U, and $\tau(\vec{u})$ is the density of mass per unit of length at $\vec{u}$, while $dl(\vec{u})$ is the infinitesimal element of such length at $\vec{u}$. Then the mass in any measurable subset S of D is $m(S) = \int_S dl(\vec{u})\tau(\vec{u})$.

Example 10.1.28 D is a finite set of points, $\vec{u}_1, \cdots, \vec{u}_N$, and these points have masses $m_1, \cdots, m_N$. The mass of any subset S of D is the sum of the masses M_ν for which $\vec{u}_\nu \in S$.

10.1.2 Integrals

Let $F : D \to \mathcal{R}$ be a real-valued function on D. The reader is presumed to know how to calculate the integral of f on D with respect to a mass distribution m. We will denote this integral by $\int_D dm(\vec{u})f(\vec{u})$. In the foregoing examples, this integral is as follows:

Example 10.1.25

$$\int_D dm(\vec{u})f(\vec{u}) = \int_D dV(\vec{u})\rho(\vec{u})f(\vec{u}).$$

Example 10.1.26

$$\int_D dm(\vec{u})f(\vec{u}) = \int_S dA(\vec{u})\sigma(\vec{u})f(\vec{u}).$$

Example 10.1.27

$$\int_D dm(\vec{u})f(\vec{u}) = \int_C dl(\vec{u})\tau(\vec{u})f(\vec{u}).$$

Example 10.1.28

$$\int_D dm(\vec{u})f(\vec{u}) = \sum_{v=1}^{N} m_v f(\vec{u}_v).$$

It is possible to invent functions $f : D \to \mathcal{R}$ so discontinuous that $\int_D dm(\vec{u})f(\vec{u})$ cannot be defined in examples 10.1.25 , 10.1.26, 10.1.27. (Clearly $\int_D dm(\vec{u})f(\vec{u})$ is always defined in example 10.1.28.) If the integral can be defined, f is said to be integrable with respect to the mass distribution m, or integrable-dm. We will not need the general theory of mass distributions. All we need are the following properties of integrals:

Remark 10.1.46 Suppose $D_1, \cdots, D_N$ are subsets of U , with mass distributions $m_1, \cdots, m_N$. On $D = D_1 \cup \cdots \cup D_N$, define a mass distribution m by $m(S) = \sum_{v=1}^N m_v(S \cap D_v)$. Suppose $f : D \to \mathcal{R}$ and for each v, $f|D_v$ is integrable-dm_v. Then f is integrable-dm and

$$\int_D dm(\vec{u})f(\vec{u}) = \sum_{v=1}^{N} \int_{D_v} dm_v(\vec{u})(f|D_v)(\vec{u}). \tag{10.1.1}$$

Remark 10.1.47 Suppose m is a mass distribution on D and for each $v \in \{1, \cdots, N\}$, $c_v \in \mathcal{R}$ and $f_v : D \to \mathcal{R}$ is measurable-m. Then $c_v f_v : D \to \mathcal{R}$ is integrable-dm and

$$\int_D dm(\vec{u})(c_v f_v)(\vec{u}) = c_v \int_D dm(\vec{u})f_v(\vec{u}). \tag{10.1.2}$$

Remark 10.1.48 Suppose m is a mass distribution on D, and $c \in \mathcal{R}$, and S is an m-measurable subset of D. Suppose $f : D \to \mathcal{R}$ is defined by $f(\vec{u}) = c$ if $\vec{u} \in S$, $f(\vec{u}) = 0$ if $\vec{u} \notin S$. Then f is integrable-dm and

$$\int_D dm(\vec{u})f(\vec{u}) = cm(S). \tag{10.1.3}$$

Remark 10.1.49 Suppose m is a mass distribution on D and $f : D \to \mathcal{R}$ and $g : D \to \mathcal{R}$ are both integrable-dm and $f(\vec{u}) \leq g(\vec{u})$ for every $\vec{u} \in U$. Then

$$\int_D dm(\vec{u})f(\vec{u})) \leq \int_D dm(\vec{u})g(\vec{u}). \tag{10.1.4}$$

Remark 10.1.50 Suppose m is a mass distribution on D and $f : D \to \mathcal{R}$ is integrable-dm. Then so is $|f| : D \to \mathcal{R}$, and (by 10.1.4)

$$\left|\int_D dm(\vec{u}) f(\vec{u})\right| \leq \int_D dm(\vec{u}) \left|f(\vec{u})\right|. \tag{10.1.5}$$

Amended version of remark 10.1.50:

Remark 10.1.51 Suppose $f : D \to \mathcal{R}$ is measurable-m. Then so is $|f|$, and f is integrable-dm iff $|f|$ is integrable-dm. If f is integrable-dm then

$$\left|\int_D dm(\vec{u}) f(\vec{u})\right| \leq \int_D dm(\vec{u}) \left|f(\vec{u})\right|.$$

Remark 10.1.52 Suppose $f : D \to \mathcal{R}$ and $g : D \to \mathcal{R}$ are measurable-m, and g is integrable-dm, and $|f| \leq g$. Then f is integrable-dm.

Our goal is simply to extend the idea of integrating a function f with respect to a mass distribution so that we can integrate vector-valued and tensor-valued function. Since $V_1 \otimes \cdots \otimes V_q$ is a Euclidean vector space when $V_1, \cdots, V_q$ are, it suffices to consider vector-valued functions.

Definition 10.1.42 Let U and V be Euclidean vector spaces. Let D be a subset of U, and let m be a mass distribution on D. Let $\vec{f} : D \to V$. We say that $\vec{f}$ is integrable-dm if for every fixed $\vec{v} \in V$ the function $(\vec{v} \cdot \vec{f}) : D \to \mathcal{R}$ is integrable-dm. (Obviously we define $\vec{v} \cdot \vec{f}$ by requiring $(\vec{v} \cdot \vec{f})(\vec{u}) = \vec{v} \cdot \vec{f}(\vec{u})$ for all $\vec{u} \in D$.) We define $I_{m,\vec{f}} : V \to \mathcal{R}$ by requiring for each $\vec{v} \in V$

$$I_{m,\vec{f}}(\vec{v}) = \int_D dm(\vec{u})(\vec{v} \cdot \vec{f})(\vec{u}). \tag{10.1.6}$$

Corollary 10.1.36 If $\vec{f} : D \to V$ is integrable-dm, then $I_{m,\vec{f}} \in \mathcal{L}(V \to \mathcal{R})$.

Proof:

Suppose $\vec{v}_1, \cdots, \vec{v}_N \in V$ and $c_1, \cdots, c_N \in \mathcal{R}$. Then $c_v \vec{v}_v \in V$, so $(c_v \vec{v}_v) \cdot \vec{f}$ is integrable-dm. Moreover, $(c_v \vec{v}_v) \cdot \vec{f} = c_v(\vec{v}_v \cdot \vec{f})$, so by remark (10.1.47)

$$\begin{aligned} I_{m,\vec{f}}(c_v \vec{v}_v) &= \int_D dm(\vec{u}) \left[(c_v \vec{v}_v) \cdot \vec{f}\right](\vec{u}) \\ &= \int_D dm(\vec{u}) \left[c_v(\vec{v}_v \cdot \vec{f})\right](\vec{u}) \\ &= c_v \int_D dm(\vec{u})(\vec{v}_v \cdot \vec{f})(\vec{u}) = c_v I_{m,\vec{f}}(\vec{v}_v). \end{aligned}$$

Corollary 10.1.37 If $\vec{f} : D \to V$ is integrable-dm, there is a unique vector $\vec{\mathcal{V}}(m, \vec{f}\,) \in V$ such that for all $\vec{v} \in V$, $I_{m,\vec{f}}(\vec{v}) = \vec{v} \cdot \vec{\mathcal{V}}(m, \vec{f}\,)$.

Proof:

Immediate from theorem 10 and corollary 10.1.36.

Definition 10.1.43 If $\vec{f} : D \to V$ is integrable-dm, we denote the vector $\vec{\mathcal{V}}(m, \vec{f}\,)$ by $\int_D dm(\vec{u})\vec{f}(\vec{u})$ and call it the integral of $\vec{f}$ with respect to m. From this definition and (10.1.6), it follows that for every $\vec{v} \in V$

$$\left[\int_D dm(\vec{u})\vec{f}(\vec{u})\right] \cdot \vec{v} = \int_D dm(\vec{u})(\vec{v} \cdot \vec{f})(\vec{u}) \tag{10.1.7}$$

and that the vector $\int_D dm(\vec{u})\vec{f}(\vec{u})$ is uniquely determined by the demand that (10.1.7) hold for all $\vec{v} \in V$.

Note 10.1.6 If $V = V_1 \otimes \cdots \otimes V_q$, then the dot product on V is the generalized dot product $\langle q \rangle$. Then $T : D \to V_1 \otimes \cdots \otimes V_q$ is integrable-dm if for every $Q \in V_1 \otimes \cdots \otimes V_q$, the function $Q\langle q \rangle T : D \to \mathcal{R}$ is integrable-dm. If it is, then $\int_D dm(\vec{u})T(\vec{u})$ is the unique tensor in $V_1 \otimes \cdots \otimes V_q$ such that for every $Q \in V_1 \otimes \cdots \otimes V_q$

$$\left[\int_D dm(\vec{u})T(\vec{u})\right] \langle q \rangle Q = \int_D dm(\vec{u})(Q\langle q \rangle T)(\vec{u}). \tag{10.1.8}$$

Here $(Q\langle q \rangle T)(\vec{u}) := Q\langle q \rangle [T(\vec{u})]$. If $Q = \vec{v}_1\vec{v}_2 \cdots \vec{v}_q$ then $(Q\langle q \rangle T)(\vec{u}) = [T(\vec{u})]\langle q \rangle Q = [T(\vec{u})](\vec{v}_1, \cdots, \vec{v}_q)$ so (10.1.8) implies the special case

$$\left[\int_D dm(\vec{u})T(\vec{u})\right](\vec{v}_1, \cdots, \vec{v}_q) = \int_D dm(\vec{u})\{\,[T(\vec{u})]\,(\vec{v}_1, \cdots, \vec{v}_q)\} \tag{10.1.9}$$

for any $(\vec{v}_1, \cdots, \vec{v}_q) \in V_1 \times \cdots \times V_q$.

10.2 Integrals in terms of components

Remark 10.2.53 Suppose U and V are Euclidean spaces, $D \subseteq U$, and m is a mass distribution on D. Suppose $\vec{f} : D \to V$ is integrable-dm. Suppose $(\vec{\beta}_1, \cdots, \vec{\beta}_n)$ is a basis for V, with dual basis $(\vec{\beta}^1, \cdots, \vec{\beta}^n)$. Define $f^i := \vec{\beta}^i \cdot \vec{f}$. Then $f^i : D \to \mathcal{R}$ is integrable-dm and

$$\left[\int_D dm(\vec{u})\vec{f}(\vec{u})\right]^i = \int_D dm(\vec{u})\vec{f}^{\,i}(\vec{u}). \tag{10.2.1}$$

Proof:

This is a special case of definitions (10.1.42, 10.1.43) with $\vec{v} = \vec{\beta}^i$ in (10.1.43).

Lemma 10.2.25 Suppose $V_1, \cdots, V_q$ are Euclidean spaces and, for $p \in \{1, \cdots, q\}$, $(\vec{\beta}_1^{(p)}, \cdots, \vec{\beta}_{n_p}^{(p)})$ is a basis for V_p, with dual basis $(\vec{\beta}^1_{(p)}, \cdots, \vec{\beta}^{n_p}_{(p)})$. Then the polyad bases $\{\vec{\beta}_{i_1}^{(1)} \cdots \vec{\beta}_{i_q}^{(q)}\}$ and $\{\vec{\beta}^{i_1}_{(1)} \cdots \vec{\beta}^{i_q}_{(q)}\}$ are dual to one another.

Proof:

$$\left(\vec{\beta}_{i_1}^{(1)} \cdots \vec{\beta}_{i_q}^{(q)}\right) \langle q \rangle \left(\vec{\beta}^{j_1}_{(1)} \cdots \vec{\beta}^{j_q}_{(q)}\right) = \delta_{i_1}{}^{j_1} \cdots \delta_{i_q}{}^{j_q}.$$

Corollary 10.2.38 Suppose U, $V_1, \cdots, V_q$ are Euclidean vector spaces, D, m is a mass distribution on D, and $T : D \to V_1 \otimes \cdots \otimes V_q$ is integrable-dm. For $p \in \{1, \cdots, q\}$, suppose $(\vec{\beta}_1^{(p)}, \cdots, \vec{\beta}_{n_p}^{(p)})$ is a basis for V_p, with dual basis $(\vec{\beta}^1_{(p)}, \cdots, \vec{\beta}^{n_p}_{(p)})$. Suppose $T^{i_1 \cdots i_q} = (\vec{\beta}^{i_1}_{(1)} \cdots \vec{\beta}^{i_q}_{(q)})\langle q \rangle T$. Then $T^{i_1 \cdots i_q} : D \to \mathcal{R}$ is integrable-dm and

$$\int_D dm(\vec{u})T^{i_1 \cdots i_q}(\vec{u}) = \left[\int_D dm(\vec{u})T(\vec{u})\right]^{i_1 \cdots i_q}. \tag{10.2.2}$$

Proof:

Immediate from lemma 10.2.25 and remark 10.2.53 by setting, in remark 10.2.53 , $\{\vec{\beta}_1, \cdots, \vec{\beta}_n\} = \{\vec{\beta}_{i_1}^{(1)} \cdots \vec{\beta}_{i_q}^{(1)}\}$ and $\{\vec{\beta}^1, \cdots, \vec{\beta}^n\} = \{\vec{\beta}^{i_1}_{(1)} \cdots \vec{\beta}^{i_q}_{(q)}\}$.

Theorem 10.2.22 Suppose U and V are Euclidean vector spaces, D is a subset of U, m is a mass distribution on D, and $\vec{f} : D \to V$. Suppose $(\vec{\beta}_1, \cdots, \vec{\beta}_n)$ is a basis for V and $f^i = \vec{\beta}^i \cdot \vec{f}$. Suppose that for each $i \in \{1, \cdots, n\}$, $f^i : D \to \mathcal{R}$ is integrable-dm. Then $\vec{f}$ is integrable-dm.

Proof:

Let $\vec{v} \in V$, and $v_i = \vec{\beta}_i \cdot \vec{v}$. Then $\vec{v} \cdot \vec{f} = v_i f^i$. Each f^i is integrable-dm, so $v_i f^i$ is integrable-dm by remark 10.1.47. Thus, $\vec{v} \cdot \vec{f}$ is integrable-dm. Since this is true for any $\vec{v} \in V$, definition 10.1.42 is satisfied, and $\vec{f}$ is integrable-dm.

Corollary 10.2.39 Suppose U , $V_1, \cdots, V_q$ are Euclidean spaces, $D \subseteq U$, m is a mass distribution on D, and $T : D \to V_1 \otimes \cdots \otimes V_q$. Suppose $(\vec{\beta}_1^{(p)}, \cdots, \vec{\beta}_{n_p}^{q})$ is a basis for V_p with dual basis $(\beta_{(p)}^1, \cdots, \vec{\beta}_{(p)}^{n_p})$. Suppose $T^{i_1 \cdots i_q} := (\beta_{(1)}^{i_1} \cdots \vec{\beta}_{(q)}^{i_q})\langle q \rangle T$ and that each $T^{i_1 \cdots i_q} : D \to \mathcal{R}$ is integrable-dm. Then T is integrable-dm.

Proof:

Immediate from theorem 10.2.22 and lemma 10.2.25.

Now it becomes obvious that remarks 10.1.46, 10.1.47, 10.1.48 are true if $\vec{f} : D \to V$ instead of $f : D \to \mathcal{R}$, and if in remark 10.1.48, $\vec{c} \in V$. We simply apply those remarks to the component functions of $\vec{f}$ relative to a basis for V. Remark 10.1.49 has no analogue for vector-valued functions, because $\vec{f} < \vec{g}$ is not defined for vectors. However, (10.1.5) does have a vector analogue. Suppose $\vec{f} : D \to V$, and suppose $\vec{f}$ and also $\|\vec{f}\| : D \to \mathcal{R}$ are integrable-dm. Then

$$\| \int_D dm(\vec{u}) \vec{f}(\vec{u}) \| \leq \int_D dm(\vec{u}) \|\vec{f}(\vec{u})\|. \tag{10.2.3}$$

Proof:

For any fixed $\vec{v} \in V$, $\vec{v} \cdot \vec{f}$ is integrable-dm and so is $\|\vec{v}\| \, \|\vec{f}\|$. Also, $|\vec{v} \cdot \vec{f}| \leq \|\vec{v}\| \, \|\vec{f}\|$. Therefore, using (10.1.4),

$$\begin{aligned} \left| \vec{v} \cdot \int_D dm(\vec{u}) \vec{f}(\vec{u}) \right| &= \left| \int_D dm(\vec{u}) (\vec{v} \cdot \vec{f})(\vec{u}) \right| \leq \int_D dm(\vec{u}) \|\vec{v}\| \|\vec{f}(\vec{u})\| \\ &= \|\vec{v}\| \int_D dm(\vec{u}) \|\vec{f}(\vec{u})\|. \end{aligned}$$

If we set $\vec{v} = \int_D dm(\vec{u}) \vec{f}(\vec{u})$ in this inequality, we obtain

$$\|\vec{v}\|^2 \leq \|\vec{v}\| \int_D dm(\vec{u}) \left\| \vec{f}(\vec{u}) \right\|.$$

Cancelling one factor $\|\vec{v}\|$ gives (10.2.3).

We also have from remark 10.2.53, and theorem 10.2.22,

Remark 10.2.54 Suppose U, V, W, X, Y are Euclidean spaces, $D \subseteq U$, and m is a mass distribution on D. Suppose $P \in V \otimes V \otimes W$, $Q \in X \otimes Y$, and $T : D \to W \otimes X$ is integrable-dm. Then so are $P \cdot T : D \to V \otimes X$ and $T \cdot Q : D \to W \otimes Y$. Moreover

$$\int_D dm(\vec{u})(P \cdot T)(\vec{u}) = P \cdot \int_D dm(\vec{u}) T(\vec{u}) \tag{10.2.4}$$

$$\left[\int_D dm(\vec{u})(T \cdot Q)(\vec{u})\right] = \left[\int_D dm(\vec{u}) T(\vec{u})\right] \cdot Q. \tag{10.2.5}$$

Proof:

> We consider $P \cdot T$. The proof for $T \cdot Q$ is similar. Choose bases for V, W, X, Y and take components with respect to these bases. Then $(P \cdot T)(\vec{u}) = P \cdot [T(\vec{u})]$ so
>
> $$(P \cdot T)^{ik}(\vec{u}) = [(P \cdot T)(\vec{u})]^{ik} = [P \cdot T(\vec{u})]^{ik} = P^i{}_j T^{jk}(\vec{u}).$$
>
> By hypothesis, T is integrable-dm. Hence so is T^{jk}. Hence so is $P^i{}_j T^{jk}$. Hence so is $(P \cdot T)^{ik}$. Hence so is $P \cdot T$. And
>
> $$\begin{aligned} \left[\int_D dm(\vec{u})(P \cdot T)(\vec{u})\right]^{ik} &= \int_D dm(\vec{u})(P \cdot T^{ik})(\vec{u}) \\ = \;& \int_D dm(\vec{u}) P^i_j T^{jk}(\vec{u}) = P^i{}_j \int_D dm(\vec{u}) T^{jk}(\vec{u}) \\ = \;& P^i{}_j \left[\int_D dm(\vec{u}) T(\vec{u})\right]^{jk} = \left[P \cdot \int_D dm(\vec{u}) T(\vec{u})\right]^{ik}. \end{aligned}$$

Corollary 10.2.40 In remark 10.2.54, take $W = W_1 \otimes \cdots \otimes W_p$ and $X = X_1 \otimes \cdots \otimes X_q$. Then

$$\int_D dm(\vec{u})(P\langle p\rangle T)(\vec{u}) = P\langle p\rangle \int_D dm(\vec{u}) T(\vec{u}) \tag{10.2.6}$$

$$\int_D dm(\vec{u})(T\langle q\rangle Q)(\vec{u}) = \left[\int_D dm(\vec{u}) T(\vec{u})\right] \langle q\rangle Q. \tag{10.2.7}$$

Chapter 11

Integral Identities

11.1 Linear mapping of integrals

Remark 11.1.55 Suppose U, V, W are Euclidean spaces, $D \subseteq U$, m is a mass distribution on D, $L \in \mathcal{L}(V \to W)$, and $\vec{f} : D \to V$ is integrable-dm. Then $L \circ \vec{f} : D \to W$ is integrable-dm and

$$\int_D dm(\vec{u})(\text{Ł} \circ \vec{f})(\vec{u}) = L\left[\int_D dm(\vec{u})\vec{f}(\vec{u})\right]. \tag{11.1.1}$$

In other words, L can be applied to $\int_D dm(\vec{u})\vec{f}(\vec{u})$ by applying it to the integrand.

Proof:

Let $\overleftrightarrow{L} \in V \otimes W$ be the tensor corresponding to L. Then $(L \circ \vec{f})(\vec{u}) = L[\vec{f}(\vec{u})] = \vec{f}(\vec{u}) \cdot \overleftrightarrow{L} = (\vec{f} \cdot \overleftrightarrow{L})(\vec{u})$ so $L \circ \vec{f} = \vec{f} \cdot \overleftrightarrow{L}$, and remark 10.2.54 says this is integrable and that

$$\begin{aligned}
\int_D dm(\vec{u})(L \circ \vec{f})(\vec{u}) &= \int_D dm(\vec{u})(\vec{f} \cdot \overleftrightarrow{L})(\vec{u}) \\
&= \left[\int_D dm(\vec{u})\vec{f}(\vec{u})\right] \cdot \overleftrightarrow{L} \\
&= L\left[\int_D dm(\vec{u})\vec{f}(\vec{u})\right].
\end{aligned}$$

QED.

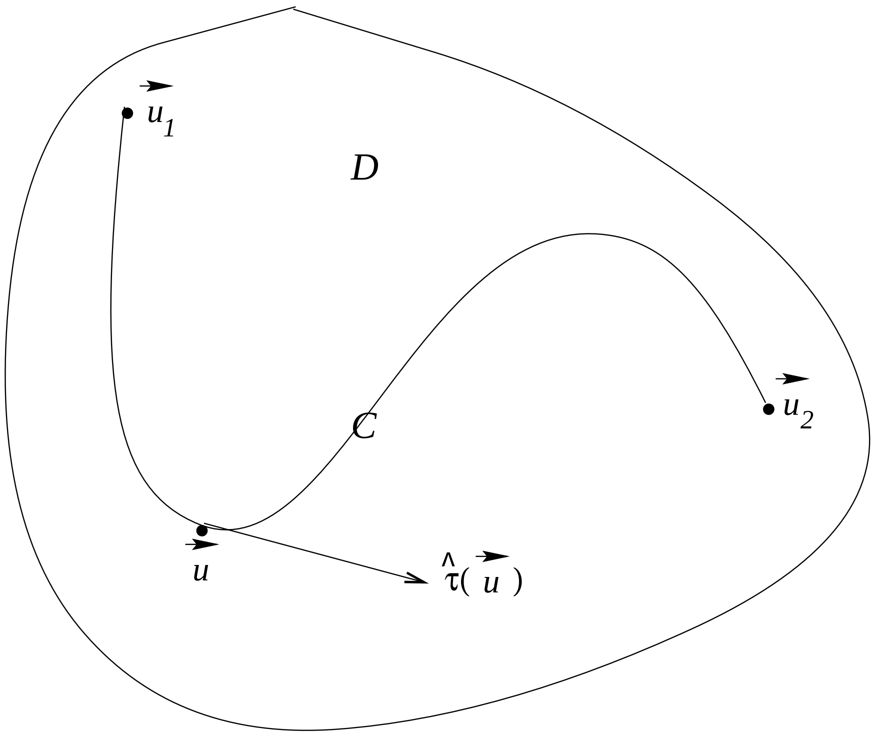


Figure 11.1:

Corollary 11.1.41 Suppose U, V, W, D, L and m as in remark 11.1.55. Suppose $T : D \to \otimes^q V$ is integrable-dm. Then $(\otimes^q L) \circ T : D \to \otimes^q W$ is integral-dm and

$$\int_D dm(\vec{u}) \left[(\otimes^q L) \circ T \right] (\vec{u}) = (\otimes^q L) \left[\int_D dm(\vec{u}) T(\vec{u}) \right]. \tag{11.1.2}$$

Proof:

A special case of remark 11.1.55, with V, W, L replaced by $\otimes^q V$, $\otimes^q W$ and $\otimes^q L$.

11.2 Line integral of a gradient

Theorem 11.2.23 Suppose U and V are Euclidean spaces, D is an open subset of U, and $\vec{f} : D \to V$ is continuously differentiable at each $\vec{u} \in D$. Suppose C is a smooth curve in D whose two endpoints are $\vec{u}_1$ and $\vec{u}_2$. Suppose $dl(\vec{u})$ is the infinitesimal element of arc length on C, and $\hat{\tau}(\vec{u})$ is the unit vector tangent to C at $\vec{u}$ and pointing along C in the direction from $\vec{u}_1$ to $\vec{u}_2$. Then

$$\int_C dl(\vec{u})(\hat{\tau}\cdot\vec{\nabla}\vec{f})(\vec{u}) = \vec{f}(\vec{u}_2) - \vec{f}(\vec{u}_1). \tag{11.2.1}$$

Proof:

We assume that the reader knows this theorem when $V = \mathcal{R}$, and $\vec{f}$ is a real-valued function. Let $(\vec{\beta}_1, \cdots, \vec{\beta}_n)$ be a basis for V, with dual basis $(\vec{\beta}^1, \cdots, \vec{\beta}^n)$, and define $f^i := \vec{\beta}^i \cdot \vec{f}$. Since $\vec{f}$ is differentiable at each $\vec{u}$, so is f^i. Thus $\vec{\nabla} f^i(\vec{u})$ exists. Moreover, $\vec{\nabla}\vec{\beta}_i = \overleftrightarrow{0}$, so by equation (9.4.9), $\vec{\nabla} f^i(\vec{u}) = \vec{\nabla}(\vec{f}\cdot\vec{\beta}_i)(\vec{u}) = \vec{\nabla}\vec{f}(\vec{u})\cdot\vec{\beta}_i$. Both $\vec{\nabla}\vec{f}$ and $\vec{\beta}_i$ are continuous, and then so is $\vec{\nabla}\vec{f}\cdot\vec{\beta}_i$. Thus $f^i : D \to \mathcal{R}$ is continously differentiable. By the known scalar version of (11.2.1),

$$\int_C dl(\vec{u})\hat{\tau}\cdot\vec{\nabla} f^i(\vec{u}) = f^i(\vec{u}_2) - f^i(\vec{u}_1).$$

If we multiply this equation by $\vec{\beta}_i$ and sum over i we obtain

$$\begin{aligned}
\vec{f}(\vec{u}_2) - \vec{f}(\vec{u}_1) &= \left[\int_C dl(\vec{u})\hat{\tau}\cdot\vec{\nabla} f^i(\vec{u})\right]\vec{\beta}_i \\
&= \int_C dl(\vec{u})\left[(\hat{\tau}\cdot\vec{\nabla} f^i)\vec{\beta}_i\right](\vec{u}) \text{ by 10.2.7 with } q=0 \\
&= \int_C dl(\vec{u})(\hat{\tau}\cdot\vec{\nabla}\vec{f})(\vec{u}) \text{ because} \\
(\hat{\tau}\cdot\vec{\nabla} f^i)\vec{\beta}_i &= \hat{\tau}\cdot\left[(\vec{\nabla} f^i)\vec{\beta}_i\right] = \hat{\tau}\cdot\vec{\nabla}(f^i\vec{\beta}_i) = \hat{\tau}\cdot\vec{\nabla}\vec{f}.
\end{aligned}$$

QED.

Corollary 11.2.42 Suppose D is open and arcwise connected (any two points in D can be joined by a smooth curve lying wholly in D). Suppose $\vec{f} : D \to V$ is continuously differentiable and $\vec{\nabla}\vec{f}(\vec{u}) = \overleftrightarrow{0}$ for all $\vec{u}\in D$. Then $\vec{f}$ is a constant function.

Proof:

For any $\vec{u}_1, \vec{u}_2 \in D$, $\vec{f}(\vec{u}_1) = \vec{f}(\vec{u}_2)$ by theorem 11.2.23.

11.3 Gauss's theorem

Theorem 11.3.24 Suppose U and V are Euclidean spaces, D is an open subset of U, ∂D is the boundary of D, and $\hat{n}(\vec{u})$ is the unit outward normal at $\vec{u} \in \partial D$. Let $\bar{D} = D \cup \partial D =$ closure of D. Suppose $\overleftrightarrow{f} \colon \bar{D} \to U \otimes V$ is continuous and $\overleftrightarrow{f} \,|D$ is continuously differentiable at each $\vec{u} \in D$. Let $dV(\vec{u})$ and $dA(\vec{u})$ be the infinitesimal volume element in D and the infinitesimal surface area element on ∂D respectively ($dV(\vec{u})$ has dimension $\dim U$, and $dA(\vec{u})$ has dimension one less). Then

$$\int_{\partial D} dA(\vec{u})(\hat{n} \cdot \overleftrightarrow{f})(\vec{u}) = \int_D dV(\vec{u})(\vec{\partial} \cdot \overleftrightarrow{f})(\vec{u}). \tag{11.3.1}$$

Proof:

We assume that the reader knows the corresponding theorem for vector-valued functions, $\vec{f} : \bar{D} \to U$. Let $(\vec{\beta}_1, \cdots, \vec{\beta}_n)$ be a basis for V, with dual basis $(\vec{\beta}^1, \cdots, \vec{\beta}^n)$. Define $\vec{f}^{\,i} : \bar{D} \to U$ by $\vec{f}^{\,i} := \overleftrightarrow{f} \cdot \vec{\beta}^i$, so $\vec{f}^i(\vec{u}) = \overleftrightarrow{f}(\vec{u}) \cdot \vec{\beta}^i$. Note that $\vec{f}^{\,i}(\vec{u})$ is not a component but a vector in U. By corollary 9.1.28, $\vec{f}^{\,i} : \bar{D} \to U$ is continuous, and by section 9.4, $\vec{f}^{\,i}$ is differentiable. By corollary 9.1.28, $\vec{\partial} \vec{f}^{\,i}$ is continuous. Thus, $\vec{f}^{\,i} : \bar{D} \to U$ is continuous and $\vec{f}^{\,i}|D$ is continuously differentiable. Therefore, $\vec{f}^{\,i}$ obeys Gauss's theorem:

$$\int_{\partial D} dA(\vec{u})(\hat{n} \cdot \vec{f}^{\,i})(\vec{u}) = \int_D dV(\vec{u})(\vec{\nabla} \cdot \vec{f}^{\,i})(\vec{u}).$$

If we multiply on the right by $\vec{\beta}_i$ and sum over i, (10.2.7) with $q = 0$ gives

$$\int_{\partial D} dA(\vec{u}) \left[(\hat{n} \cdot \vec{f}^{\,i})\vec{\beta}_i\right](\vec{u}) = \int_D dV(\vec{u}) \left[(\vec{\nabla} \cdot \vec{f}^{\,i})\vec{\beta}_i\right](\vec{u}). \tag{11.3.2}$$

But $\vec{f}^{\,i}\vec{\beta}_i = \overleftrightarrow{f}(\cdot \vec{\beta}^i \vec{\beta}_i) = \overleftrightarrow{f} \cdot (\vec{\beta}^i \vec{\beta}_i) = (\overleftrightarrow{f} \cdot \overleftrightarrow{I}_V) = \overleftrightarrow{f}$ so $(\hat{n} \cdot \vec{f}^{\,i})\vec{\beta}_i = \hat{n} \cdot (\vec{f}^{\,i}\vec{\beta}_i) = \hat{n} \cdot \vec{f}$, and by section 9.4

$$\begin{aligned}
(\vec{\nabla} \cdot \vec{f}^{\,i})\vec{\beta}_i &= \left[\overleftrightarrow{I}_U \langle 2 \rangle \vec{\nabla} \vec{f}^{\,i}\right] \vec{\beta}_i = \overleftrightarrow{I}_U \langle 2 \rangle \left[(\vec{\nabla} \vec{f}^{\,i} \vec{\beta}_i)\right] \\
&= \overleftrightarrow{I}_U \langle 2 \rangle \left[\vec{\nabla}(\vec{f}^i \vec{\beta}_i)\right] = \overleftrightarrow{I}_U \langle 2 \rangle \vec{\nabla} \overleftrightarrow{f} = \vec{\nabla} \cdot \overleftrightarrow{f}\ .
\end{aligned}$$

Substituting these expressions in (11.3.2) gives 11.3.1.

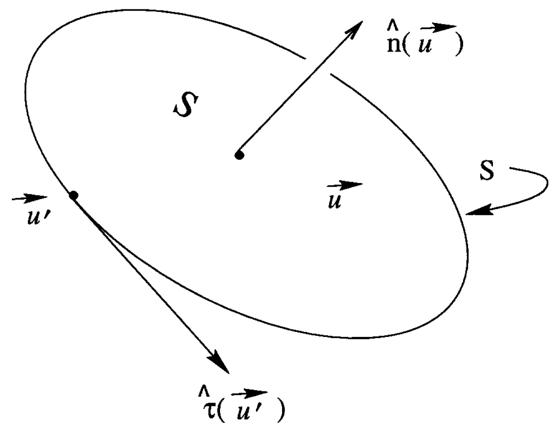


Figure 11.2:

11.4 Stokes's theorem

Theorem 11.4.25 Suppose (U, A) is an oriented Euclidean 3-space, and A is used to define cross products and curls. Suppose D is an open subset of U, and S is a smooth two-sided surface in D, with boundary curve ∂S. Suppose $\hat{n} : S \to U$ is one of the two continuous functions such that $\hat{n}(\vec{u})$ is a unit vector in U perpendicular to S at $\vec{u} \in S$. If $\vec{u} \in \partial S$, let $\hat{\tau}(\vec{u})$ be the unit vector tangent to ∂S at $\vec{u}$ and pointing so that $\hat{n}(\vec{u}) \times \hat{\tau}(\vec{u})$ points into S. Suppose that $\overleftrightarrow{f}: D \to U \otimes V$ is continuously differentiable everywhere in D. Suppose that $dA(\vec{u})$ is the element of area on S and $dl(\vec{u})$ is the element of arc length on ∂S. Then

$$\int_{\partial S} dl(\vec{u})(\hat{\tau} \cdot \overleftrightarrow{f})(\vec{u}) = \int_S dA(\vec{u}) \left[\hat{n} \cdot (\vec{\nabla} \times \overleftrightarrow{f})\right](\vec{u}). \tag{11.4.1}$$

Proof:

Again, we assume the reader knows this theorem when $V = \mathcal{R}$, i.e. for vector fields $\vec{f} : D \to U$. Let $(\vec{\beta}_1, \cdots, \vec{\beta}_n)$ be a basis for V, with dual basis $(\vec{\beta}^1, \cdots, \vec{\beta}^n)$, and define $\vec{f}^i = \overleftrightarrow{f} \cdot \vec{\beta}^i$. Then we can use Stoke's theorem for $\vec{f}^i$, i.e., replace $\overleftrightarrow{f}$ by $\vec{f}^i$ in (11.4.1). If we multiply both sides of the result on the right by $\vec{\beta}_i$, sum over i, and use (10.2.7) with $q = 0$, we get

$$\int_{\partial S} dl(\vec{u}) \left[(\hat{\tau} \cdot \vec{f}^{\,i})\vec{\beta}_i\right](\vec{u}) = \int_S dA(\vec{u}) \left\{\left[\hat{n} \cdot (\vec{\nabla} \times \vec{f}^{\,i})\right] \vec{\beta}_i\right\}(\vec{u}).$$

As in the proof of Gauss's theorem, $\vec{f}^{\,i}\vec{\beta}_i = \overleftrightarrow{f}$, so $(\hat{\tau} \cdot \vec{f}^{\,i})\vec{\beta}_i = \hat{\tau} \cdot (\vec{f}^{\,i}\vec{\beta}_i) = \hat{\tau} \cdot \overleftrightarrow{f}$. And $[\hat{n} \cdot (\vec{\nabla} \times \vec{f}^i)]\vec{\beta}_i = \hat{n} \cdot [(\vec{\nabla} \times \vec{f}^{\,i})\vec{\beta}_i]$. Finally

$$\begin{aligned}
\left(\vec{\nabla} \times \vec{f}^{\,i}\right) \vec{\beta}_i &= \left(A\langle 2\rangle \vec{\nabla}\vec{f}^{\,i}\right) \vec{\beta}_i \\
&= A\langle 2\rangle \left[(\vec{\nabla}\vec{f}^{\,i})\vec{\beta}_i\right] \\
&= A\langle 2\rangle \left[\vec{\nabla}(\vec{f}^{\,i}\vec{\beta}_i)\right] = A\langle 2\rangle \left[\vec{\nabla}\, \overleftrightarrow{f}\right] \\
&= \vec{\nabla} \times \overleftrightarrow{f}\,.
\end{aligned}$$

This finishes the proof.

11.5 Vanishing integral theorem

Lemma 11.5.26 Suppose D is an open subset of Euclidean space U, $f : D \to \mathcal{R}$ is continuous, and $\int_{D'} dV(\vec{u}) f(\vec{u}) = 0$ for every open subset D' of D. Here $dV(\vec{u})$ is the volume element in U. Then $f(\vec{u}) = 0$ for all $\vec{u} \in D$.

Proof:

Suppose $\vec{u}_0 \in D$ and $f(\vec{u}_0) \neq 0$. If f satisfies the hypotheses so does $-f$, so if $f(\vec{u})_0) < 0$ we may replace f by $-f$. Therefore, we may assume $f(\vec{u}_0) > 0$. Then there is a $\delta > 0$ such that if $|\vec{u} - \vec{u}_0| < \delta$ then $\vec{u} \in D$ and $|f(\vec{u}) - f(\vec{u}_0)| < f(\vec{u}_0)/2$. But then $-f(\vec{u}_0)/2 < f(\vec{u}) - f(\vec{u}_0) < f(\vec{u}_0)/2$, so $f(\vec{u}) > f(\vec{u}_0)/2$ if $\vec{u} \in B(\vec{u}_0, \delta)$, the open ball of radius δ centered at $\vec{u}_0$. Then

$$\int_{B(\vec{u}_0,\delta)} dV(\vec{u}) f(\vec{u}) > \int_{B(\vec{u}_0,\delta)} dV(\vec{u}) \frac{f(\vec{u}_0)}{2} = |B(\vec{u_0}, \delta)| \frac{f(\vec{u}_0)}{2},$$

where $|B(\vec{u}_0, \delta)|$ = volume of $B(\vec{u}_0, \delta)$. But $|B(\vec{u}_0, \delta)| f(\vec{u}_0)/2 > 0$, and $B(\vec{u}_0, \delta)$ is an open subset of D, so we must have $\int_{B(\vec{u}_0,\delta)} dV(\vec{u}) f(\vec{u}) = 0$. This contradiction shows $f(\vec{u}_0) = 0$.

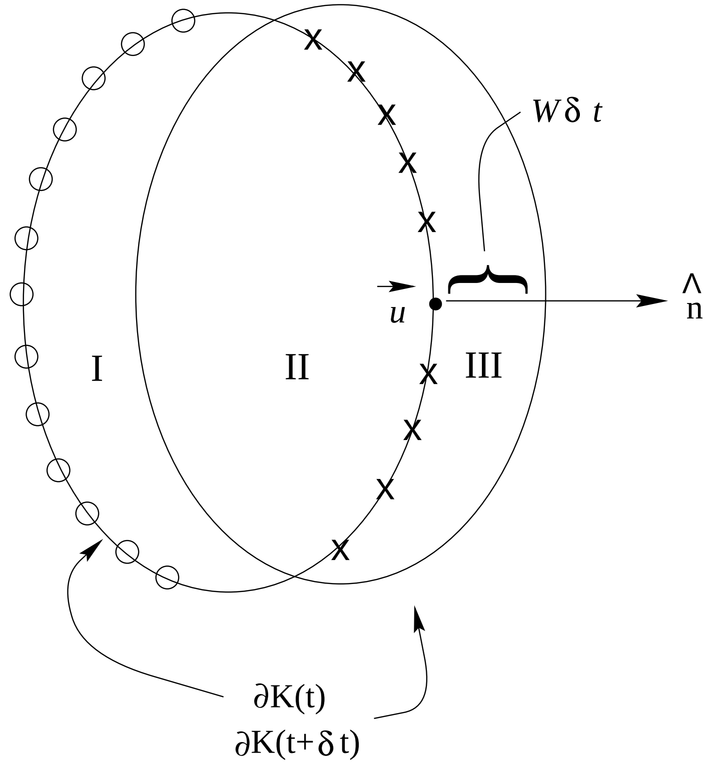


Figure 11.3:

Theorem 11.5.26 Suppose U and V are Euclidean spaces, D is an open subset of U, and $\vec{f}: D \to V$ is continuous. Suppose that for every open subset D' of D, we have $\int_{D'} dV(\vec{u})\vec{f}(\vec{u}) = \vec{0}_V$. Then $\vec{f}(\vec{u}) = \vec{0}_V$ for all $\vec{u} \in D$.

Proof:

Let $(\vec{\beta}_1, \cdots, \vec{\beta}_n)$ be a basis for V, with dual basis $(\vec{\beta}^1, \cdots, \vec{\beta}^n)$. Let $f^i = \vec{\beta}^i \cdot \vec{f}$. Then $f^i : D \to \mathcal{R}$ is continuous, and for every open subset D' of D, $\int_{D'} dV(\vec{u}) f^i(\vec{u}) = 0$. Hence, $f^i = 0$ by lemma 11.5.26. Hence $\vec{f} = f^i \vec{\beta}_i = \vec{0}_V$.

11.6 Time derivative of an integral over a moving volume

Suppose that for each instant t we are given an open set $K(t)$ in Euclidean vector space U, with boundary surface $\partial K(t)$ and outward unit normal $\hat{n}$. Suppose that $K(t)$ changes with time sufficiently smoothly that its surface $\partial K(t)$ has a well defined speed normal to itself. At $\vec{u} \in \partial K(t)$, this speed is $W(\vec{u}, t)$, and it is > 0 if $\partial K(t)$ is moving outward from $K(t)$, < 0 if inward.

In figure 11.3, $W > 0$ on the part of $\partial K(t)$ marked with x's and $W < 0$ on the part of $\partial K(t)$ marked with o's.

Suppose that all the sets $\overline{K(t)}$ are subsets ($\overline{K(t)} = K(t) \cup \partial K(t) =$ closure of $K(t)$) of an open set D, and $\vec{f} : D \times \mathcal{R} \to V$, where V is another Euclidean space. Thus for each $\vec{u} \in D$ and $t \in \mathcal{R}$, $\vec{f}(\vec{u}, t)$ is a vector in V. We suppose this vector depends continuously on $\vec{u}$ and is continuously differentiable in t (i.e. $\vec{f}(\cdot, t)$ is continuous for each $t \in \mathcal{R}$, and $\vec{f}(\vec{u}, \cdot)$ is continuously differentiable for each $\vec{u} \in D$). Then

$$\begin{aligned} \frac{d}{dt} \int_{K(t)} dV(\vec{u}) \vec{f}(\vec{u}, t) &= \int_{K(t)} dV(\vec{u}) \partial_t \vec{f}(\vec{u}, t) \\ &+ \int_{\partial K(t)} dA(\vec{u}) (W\vec{f})(\vec{u}, t). \end{aligned} \tag{11.6.1}$$

Here $\partial_t \vec{f}(\vec{u}, t)$ stands for $\partial[\vec{f}(\vec{u}, \cdot)](t)$, or $\partial \vec{f}(\vec{u}, t)/\partial t$.

Proof:

Let $\partial K(t)^+$ stand for the part of $\partial K(t)$ marked with x's in figure 11.3, while $\partial K(t)^-$ stands for the part marked with o's. Let II be the set $K(t) \cap K(t+\delta t)$, while I is $K(t) \backslash K(t+\delta t)$, and III is $K(t+\delta t) \backslash K(t)$. See figure 11.3. Let

$$\vec{F}(t) = \int_{K(t)} dV(\vec{u}) \vec{f}(\vec{u}, t).$$

Then

$$\vec{F}(t) = \int_{\mathrm{I}} dV(\vec{u}) \vec{f}(\vec{u}, t) + \int_{\mathrm{II}} dV(\vec{u}) \vec{f}(\vec{u}, t).$$

$$\vec{F}(t+\delta t) = \int_{\mathrm{II}} dV(\vec{u}) \vec{f}(\vec{u}, t+\delta t) + \int_{\mathrm{III}} dV(\vec{u}) \vec{f}(\vec{u}, t+\delta t).$$

Correct to first order in δt,

$$\begin{aligned}\vec{f}(\vec{u}, t+\delta t) &= \vec{f}(\vec{u}, t) + (\delta t)\ \partial_t \vec{f}(\vec{u}, t) \text{ in } D \\ dV(\vec{u}) &= (\delta t) W(\vec{u}, t) dA(\vec{u}) \text{ in III}, \text{ where } W > 0 \\ dV(\vec{u}) &= -(\delta t) W(\vec{u}, t) dA(\vec{u}) \text{ in I}, \text{ where } W < 0.\end{aligned}$$

Thus, correct to first order in δt,

$$\begin{aligned}\int_{\mathrm{I}} dV(\vec{u}) \vec{f}(\vec{u}, t) &= -(\delta t) \int_{\partial K(t)^-} dA(\vec{u}) W \vec{f}(\vec{u}, t) \\ \int_{\mathrm{III}} dV(\vec{u}) \vec{f}(\vec{u}, t) &= (\delta t) \int_{\partial K(t)^+} dA(\vec{u}) W \vec{f}(\vec{u}, t).\end{aligned}$$

Therefore,

$$\int_{\mathrm{III}} dV(\vec{u}) \vec{f}(\vec{u}, t) - \int_{\mathrm{I}} dV(\vec{u}) \vec{f}(\vec{u}, t) = (\delta t) \int_{\partial K(t)} dA(\vec{u}) W \vec{f}(\vec{u}, t). \tag{11.6.2}$$

Also,

$$\begin{aligned}&\int_{\mathrm{II}} dV(\vec{u}) \left[\vec{f}(\vec{u}, t+\delta t) - \vec{f}(\vec{u}, t) \right] = (\delta t) \int_{\mathrm{II}} dV(\vec{u}) \partial_t \vec{f}(\vec{u}, t) \\ &= (\delta t) \int_{K(t)} dV(\vec{u}) \partial_t \vec{f}(\vec{u}, t) - (\delta t) \int_{\mathrm{I}} dV(\vec{u}) \partial_t \vec{f}(\vec{u}, t).\end{aligned}$$

But $\int_{\mathrm{I}} dV(\vec{u}) \partial_t \vec{f} = -(\delta t) \int_{\partial K(t)^-} dA(\vec{u}) W \partial_t f$, correct to first order in δt. Therefore, correct to first order in δt,

$$\int_{\mathrm{II}} dV(\vec{u}) \left[\vec{f}(\vec{u}, t+\delta t) - \vec{f}(\vec{u}, t) \right] = (\delta t) \int_{K(t)} dV(\vec{u}) \partial_t \vec{f}(\vec{u}, t). \tag{11.6.3}$$

Combining (11.6.2, 11.6.3), we have, correct to first order in δt,

$$\vec{F}(t+\delta t) - \vec{F}(t) = \delta t \left[\int_{\partial K(t)} dA(\vec{u}) W \vec{f}(\vec{u}, t) + \int_{K(t)} dV(\vec{u}) \partial_t \vec{f}(\vec{u}, t) \right].$$

Dividing by δt and letting $\delta t \to 0$ gives (11.6.1).

11.7 Change of variables of integration

Suppose (U, A_U) and (V, A_V) are two oriented n-dimensional Euclidean spaces. Suppose $L \in \mathcal{L}(U \to V)$. For any $\vec{u}_1, \cdots, \vec{u}_n \in U$, define $\mathcal{A}(\vec{u}_1, \cdots, \vec{u}_n) = A_V[L(\vec{u}_1), \cdots, L(\vec{u}_n)]$.

Clearly $\mathcal{A} \in \Lambda^n U$, so there is a scalar k_L such that $\mathcal{A} = k_L A_U$. We denote this scalar by $\det L$ and call it the determinant of L. Then

$$A_V [L(\vec{u}_1), \cdots, L(\vec{u}_n)] = (\det L) A_U(\vec{u}_1, \cdots, \vec{u}_n) \tag{11.7.1}$$

for all $(\vec{u}_1, \cdots, \vec{u}_n) \in X^n U$. Note that $\det L$ changes sign if we replace either A_U or A_V by its negative. Thus $\det L$ depends not only on L but on how U and V are oriented. For a given L, $\det L$ can have two values, one the negative of the other, and which value it has depends on which of the two unimodular alternating tensors $\pm A_U$ and $\pm A_V$ are used to orient U and V.

If $(\hat{x}_1, \cdots, \hat{x}_n)$ is a pooob in U and $(\hat{y}_1, \cdots, \hat{y}_n)$ is a pooob in V and $L_{ij} = \hat{x}_i \cdot \overleftrightarrow{L} \cdot \hat{y}_j$, then[1] $L(\hat{x}_i) = L_{ij}\hat{y}_j$ and so

$$\begin{aligned} \det L &= (\det L) A_U(\hat{x}_1, \cdots, \hat{x}_n) = A_V [L(\hat{x}_1), \cdots, L(\hat{x}_n)] \\ &= A_V [L_{ij_1}\hat{y}_{j_1}, \cdots, L_{nj_n}\hat{y}_{j_n}] = L_{1j_1} \cdots L_{nj_n} A_V(\hat{y}_{j_1}, \cdots \hat{y}_{j_n}). \\ &\ \text{Therefore } \det L = L_{1j_1} \cdot L_{nj_n} \varepsilon_{i_1 \cdots i_n} = \det(L_{ij}). \end{aligned} \tag{11.7.2}$$

Now suppose H and K are open subsets of U and V respectively and $\vec{\mathbf{r}} : H \to K$ is a continuously differentiable bijection of H onto K. For $\vec{x} \in H$ write $\vec{x} = x_i \hat{x}_i$ and define $r_j = \hat{y}_j \cdot \vec{\mathbf{r}}$. At any $\vec{x} \in H$ we have $\partial_i r_j = (\vec{\nabla}\,\vec{\mathbf{r}})_{ij} = \hat{x}_i \cdot [\vec{\nabla}\,\vec{\mathbf{r}}] \cdot \hat{y}_j$ so the determinent of the matrix $\partial_i r_j(\hat{x})$ is $\det \vec{\nabla}\,\vec{\mathbf{r}}\,(\vec{x})$, as we see from (11.7.2).

Finally, suppose $f : K \to \mathcal{R}$ is integrable with respect to volume in V. Then $f \circ \mathbf{r} : H \to \mathcal{R}$ is integrable with respect to volume in U and

$$\int_K f(r_j \hat{y}_j) dr_1 \cdots dr_n = \int_H f\left[\vec{\mathbf{r}}\,(x_k \hat{x}_k)\right] \left| \det \frac{\partial\,\mathbf{r}_j}{\partial x_i} \right| dx_1 \cdots dx_n. \tag{11.7.3}$$

The reader is presumed to know this formula for changing variables in a multiple integral. The determinant $\det(\partial r_j / \partial x_i)$ is the Jacobian of the coordinate transformation. We note that $\partial r_j / \partial x_i = \partial_i r_j$, so by (11.7.2), at $\vec{x} \in H$

$$\det \left(\frac{\partial r_j}{\partial x_i} \right) = \det \vec{\nabla}\,\vec{\mathbf{r}}\,(\vec{x}). \tag{11.7.4}$$

[1] Proof: $L_{ij}\hat{y}_j = x_i \cdot \overleftrightarrow{L} \cdot \hat{y}_j \hat{y}_j = \hat{x}_i \cdot \overleftrightarrow{L}$

The local volume elements in U and V are $dV_U(\vec{x}) = dx_1 \cdots dx_n$ and $dV_V(\vec{r}) = dr_1 \cdots dr_n$, so (11.7.3) can be written

$$\int_K dV_V(\vec{r}) f(\vec{r}) = \int_H dV_U(\vec{x}) \left|\det \vec{\nabla}\, \vec{\mathbf{r}}\,(\vec{x})\right| (\vec{f} \circ \vec{\mathbf{r}}\,)(\vec{x}). \tag{11.7.5}$$

Now suppose W is another Euclidean space, with basis $(\vec{\beta}_1, \cdots, \vec{\beta}_p)$ and dual basis $(\vec{\beta}^1, \cdots, \vec{\beta}^p)$. Suppose $\vec{f} : K \to W$ is integrable. Then so is $f_i = \vec{\beta}_i \cdot \vec{f}$, and (11.7.5) holds for each f_i. If we multiply the resulting equations by $\vec{\beta}^i$ and sum over i we obtain

$$\int_K dV_V(\vec{r}) \vec{f}(\vec{r}) = \int_H dV_U(\vec{x}) \left|\det \vec{\nabla}\, \vec{\mathbf{r}}\,(\vec{x})\right| (\vec{f} \circ \vec{\mathbf{r}}\,)(\vec{x}). \tag{11.7.6}$$

Part II

Elementary Continuum Mechanics

Chapter 12

Eulerian and Lagrangian Descriptions of a Continuum

12.1 Discrete systems

For some purposes, some physical systems (for example, the solar system) can be regarded as consisting of a finite number N of mass points. A complete *description* of the motion of such a system consists simply of N vector-valued functions of time; the value of the ν'th function at time t is the position of the ν'th particle at time t. If the values of the positions $\vec{r}_1(t), \cdots, \vec{r}_N(t)$ of all particles are known for all times, nothing more can be said about the system.

To *predict* the motion of such a "discrete" system, one must usually add other properties to the model. In the case of the solar system, one ascribes to each particle a mass, m_ν being the mass of the ν'th particle. Then, if all the velocities and positions are known at some time t_0, all the functions $\vec{r}_\nu(t)$ can be calculated from Newton's laws of motion and gravitation.

If the system of particles is a classical (pre-quantum) atom, one needs not only their masses but their charges in order to calculate their motion.

And, of course, for some purposes the discrete model of the solar system is too crude. It cannot model or predict the angular acceleration of the earth's moon due to the tidal bulges on the earth. A point mass has no tidal bulges. Nevertheless, the discrete model of the solar system is very useful. No one would suggest abandoning it because the solar system has properties not described by such a model.

12.2 Continua

A parcel of air, water or rock consists of such a large number of particles that a discrete model for it would be hopelessly complicated. A different sort of model has been developed over the last three centuries to describe such physical systems. This model, called a "continuum", exploits the fact that in air, water and rock nearby particles behave similarly.

The continuum model for a lump of material regards it as consisting of infinitely many points, in fact too many to count by labeling them even with all the integers. (There are too many real numbers between 0 and 1 to label them as $x_1, x_2, \cdots$. Any infinite sequence of real numbers between 0 and 1 will omit most of the numbers in that interval. This theorem was proved by Georg Cantor in about 1880.) Nearby points are given nearby labels. One way to label the points in the continuum is to give the Cartesian coordinates of their positions at some particular instant t_0, relative to some Cartesian coordinate system. Then each particle is a point to which is attached a number triple (x_1, x_2, x_3) giving the Cartesian coordinates of that particle at the "labelling time" t_0. This number triple is a label which moves with the particle and remains attached to it. Nearby particles have nearly equal labels, and particles with nearly equal labels are close to one another. The collection of all particles required to describe the motion will usually fill up an open set in ordinary real 3-space.

Of course there are many ways to label the particles in a continuum. The label attached to a particle could be the number triple (r, θ, λ) giving the values of its radius, colatitude and longitude in a system of polar spherical coordinates at the labelling time t_0.

Another label would be simply the position vector $\vec{x}$ at time t_0 of the particle relative to some fixed origin and axes which are unaccelerated and non-rotating (so that Newton's laws can be used). In fact, labelling by number triples can also be thought of as labelling by vectors, since a number triple is a vector in the Euclidean vector space R^3. (R^n is the set of real n-tuples, with addition and multiplication by scalars defined coordinate-wise and with the dot product of $(u_1, \cdots, u_n)$ and $(v_1, \cdots, v_n)$ defined to be $u_i v_i$.)

In order to leave ourselves freedom to choose different ways of labelling, we will not specify one particular scheme. We will simply assume that there is a three-dimensional oriented real Euclidean space, (L, A_L), from which the labels are chosen. The labels describing a continuum will be the vectors $\vec{x}$ in a certain open and subset H of label space L. We will denote real physical space by P, and we will orient it in the usual way, denoting the "right handed" unimodular alternating tensor by A_P, or simply A. We use this notation:

$$\vec{\mathbf{r}}^{\,L}(\vec{x}, t) := \text{ position at time } t \text{ of the particle labelled } \vec{x}. \tag{12.2.1}$$

The label $\vec{x}$ is in the open subset H of label space L, and the particle position $\vec{\mathbf{r}}^{\,L}(\vec{x}, t)$ is in P. One special case which is easy to visualize is to take $L = P$, and to choose some fixed time t_0, and to label the particles by their positions at t_0. With this labelling scheme, called "t_0-position labelling," we have

$$\mathbf{r}^{\,L}(\vec{x}, t_0) = \vec{x}, \text{ for } t_0\text{-position labelling.} \tag{12.2.2}$$

The motion of a continuum is completely described by giving the position of every particle at every instant, i.e. by giving $\vec{\mathbf{r}}^{\,L}(\vec{x}, t)$ for all $\vec{x} \in H$ and all $t \in \mathcal{R}$. This amounts to knowing the function

$$\mathbf{r}^{\,L} : H \times \mathcal{R} \to \mathcal{P}. \tag{12.2.3}$$

This function is called the Lagrangian description of the motion of the continuum.

We denote by $K(t)$ the set in position space P occupied by the particles of the continuum at time t. We assume that particles neither fission nor coalesce; each particle retains its identity for all time. Therefore we need never give two particles the same label $\vec{x}$, and we do not do so. But then if $\vec{r} \in K(t)$ there is exactly one particle at position $\vec{r}$ at time t. We denote its label by $\vec{\mathbf{x}}^{\,E}(\vec{r}, t)$. Thus

$$\mathbf{x}^{\,E}(\vec{r}, t) := \text{ label of particle which is at position } \vec{r} \text{ at time } t. \tag{12.2.4}$$

From the definitions, it is clear that if $\vec{r} \in K(t)$ and $\vec{x} \in H$ then

$$\vec{r} = \vec{\mathbf{r}}^{\,L}(\vec{x}, t) \Leftrightarrow \vec{x} = \vec{\mathbf{x}}^{\,E}(\vec{r}, t). \tag{12.2.5}$$

But (12.2.5) says that $\vec{\mathbf{r}}^{\,L}(\cdot, t) : H \to K(t)$ is a bijection, whose inverse is $\vec{\mathbf{x}}^{\,E}(\cdot, t) : K(t) \to H$. If we know the function $\vec{\mathbf{x}}^{\,E}(\cdot, t)$, we can find the function $\vec{\mathbf{r}}^{\,L}(\cdot, t)$. For any $\vec{x} \in H$, $\vec{\mathbf{r}}^{\,L}(\vec{x}, t)$ is the unique solution $\vec{r}$ of the equation $\vec{x} = \vec{\mathbf{x}}^{\,E}(\vec{r}, t)$. Therefore, knowing the label function $\vec{\mathbf{x}}^{\,E}$ is equivalent to knowing $\vec{\mathbf{r}}^{\,L}$. Either function is a complete description of the motion of the continuum, and either can be found from the other. For any fixed t,

$$\vec{\mathbf{r}}^{\,L}(\cdot, t)^{-1} = \vec{\mathbf{x}}^{\,E}(\cdot, t). \tag{12.2.6}$$

In order to enforce that nearby particles have nearby labels we will assume not only that $\vec{\mathbf{r}}^{\,L}$ and $\vec{\mathbf{x}}^{\,E}$ are continuous, but that they are continuously differentiable as many times as needed to make our arguments work. (Usually, twice continuously differentiable will suffice.) Since (12.2.6) implies $\vec{\mathbf{r}}^{\,L}(\cdot, t) \circ \vec{\mathbf{x}}^{\,E}(\cdot, t) = I_P|K(t)$, the identity function on P, restricted to $K(t)$, and also $\vec{\mathbf{x}}^{\,E}(\cdot, t) \circ \vec{\mathbf{r}}^{\,L}(\cdot, t) = I_L|H$, the identity function on L, restricted to H, therefore the chain rule implies that if $\vec{r} = \vec{\mathbf{r}}^{\,L}(\vec{x}, t)$ or, equivalently, $\vec{x} = \vec{\mathbf{x}}^{\,E}(\vec{r}, t)$, then

$$\begin{cases} \vec{\nabla}\,\vec{\mathbf{r}}^{\,L}(\vec{x}, t) \cdot \vec{\nabla}\,\vec{\mathbf{x}}^{\,E}(\vec{r}, t) = \overleftrightarrow{I}_L \\[2em] \vec{\nabla}\,\vec{\mathbf{x}}^{\,E}(\vec{r}, t) \cdot \vec{\nabla}\,\vec{\mathbf{r}}^{\,L}(\vec{x}, t) = \overleftrightarrow{I}_P\,. \end{cases} \tag{12.2.7}$$

12.3 Physical quantities

If the motion of a continuum is to be predicted from physical laws, more properties of the continuum must be considered than simply the positions of its particles at various times. Most of the useful properties are local, i.e. they have a value at each particle at each instant. Examples are temperature, mass density, charge density, electric current density, entropy density, magnetic polarization, the electrical conductivity tensor and, in a fluid, the pressure.

A local physical property f can be described in two ways: at any instant t, we can say what is the value of f at the particle labelled $\vec{x}$, or we can say what is the value of f at the particle whose position is $\vec{r}$ at time t. The two values are the same, but they are given by different functions, which we denote by f^L and f^E. We call these the Lagrangian and the Eulerian descriptions of f. They are defined thus:

$$f^L(\vec{x},t) = \text{ value of physical quantity } f \text{ at time } t \text{ at the particle labelled } \vec{x}; \quad (12.3.1)$$

$$f^E(\vec{r},t) \;=\; \text{value of physical quantity } f \text{ at time } t \text{ at the particle whose position is } \vec{r} \text{ at time } t. \quad (12.3.2)$$

The physical quantity f can be a scalar, vector or tensor. From the definitions, obviously

$$\begin{cases} \text{if } \vec{r} = \vec{\mathbf{r}}^{\,L}(\vec{x},t) \text{ or } \vec{x} = \vec{\mathbf{x}}^{\,E}(\vec{r},t) \text{ then} \\ \\ \qquad\qquad f^L(\vec{x},t) = f^E(\vec{r},t). \end{cases} \quad (12.3.3)$$

This is equivalent to asserting

$$\begin{cases} f^L(\vec{x},t) = f^E\left[\vec{\mathbf{r}}^{\,L}(\vec{x},t), t\right] \text{ and} \\ \\ f^E(\vec{r},t) = f^L\left[\vec{\mathbf{x}}^{\,E}(\vec{r},t), t\right]. \end{cases} \quad (12.3.4)$$

These two assertions are equivalent in turn to

$$\begin{cases} f^L(\cdot,t) = f^E(\cdot,t) \circ \vec{\mathbf{r}}^{\,L}(\cdot,t) \\ \\ f^E(\cdot,t) = f^L(\cdot,t) \circ \vec{\mathbf{x}}^{\,E}(\cdot,t). \end{cases} \quad (12.3.5)$$

If we have the Lagrangian description of the motion, the function $\vec{\mathbf{r}}^{\,E}$, then we can find the labelling function $\vec{\mathbf{x}}^{\,E}$. Then using (12.3.4) or (12.3.5), we can find both f^E and f^L if we know either one of them.

So far, our discussion of "physical quantities" has been somewhat intuitive. We have found some mathematical rules which are satisfied by certain entities that we feel comfortable to call physical quantities. In the spirit of mathematical model building, it is useful to give a precise definition to what will qualify as a "physical quantity". Because of what we have learned above, we introduce

Definition 12.3.44 Given a continuum whose motion is described by $\vec{\mathbf{r}}^{\,L} : H \times \mathcal{R} \to P$, and whose labelling function is $\vec{\mathbf{x}}^{\,E}$, with $\vec{\mathbf{x}}^{\,E}(\cdot, t) : K(t) \to H$, a physical quantity f is an ordered pair of functions (f^L, f^E) with these properties:

i) The domain of f^L is $H \times \mathcal{R}$

ii) The domain of f^E is a subset of $P \times \mathcal{R}$, and the domain of $f^E(\cdot, t)$ is $K(t)$

iii) If $\vec{\mathbf{r}}^{\,L}(\vec{x}, t) = \vec{r}$, or $\vec{\mathbf{x}}^{\,E}(\vec{r}, t) = \vec{x}$, then $f^L(\vec{x}, t) = f^E(\vec{r}, t)$.

The function f^L is the *Lagrangian description* of the physical quantity f, and f^E is the *Eulerian description* of f. Property iii) is equivalent to

$$\begin{cases} f^L(\vec{x}, t) = f^E\left[\vec{\mathbf{r}}^{\,L}(\vec{x}, t), t\right] \text{ or to} \\ \\ f^E(\vec{r}, t) = f^L\left[\vec{\mathbf{x}}^{\,E}(\vec{r}, t), t\right] \text{ for all } t \in \mathcal{R}. \end{cases} \tag{12.3.6}$$

These in turn are equivalent to

$$\begin{cases} f^L(\cdot, t) = f^E(\cdot, t) \circ \mathbf{r}^{\,L}(\cdot, t) \text{ or to} \\ \\ f^E(\cdot, t) = f^L(\cdot, t) \circ \vec{\mathbf{x}}\,(\cdot, t) \text{ for all } t \in \mathcal{R}. \end{cases} \tag{12.3.7}$$

Temperature, mass density, and the other items enumerated beginning on page 144 are physical quantities in the sense of definition 12.3.44. The advantage of having such a formal definition is that it enables us to introduce new physical quantities to suit our convenience. We can define a new physical quantity f by giving either its Eulerian description f^E or its Lagrangian description f^L. The missing description can be obtained from (12.3.6) or (12.3.7), assuming that we have $\vec{\mathbf{r}}^{\,L}$, the Lagrangian description of the motion of the continuum. A somewhat trivial example is the f such that $f^L(\vec{x}, t) = 7$ for all $\vec{x}, t$. Clearly $f^E(\vec{r}, t) = 7$, and we can regard 7 (or any other constant) as a physical quantity.

A more interesting example is a physical quantity which we call "particle position", and denote by $\vec{\mathbf{r}}$. It is defined by requiring that its Eulerian description be

$$\vec{\mathbf{r}}^E(\vec{r}, t) = \vec{r} \text{ for all } t \in \mathcal{R} \text{ and } \vec{r} \in K(r). \tag{12.3.8}$$

From (12.3.6), its Lagrangian description is the $\vec{\mathbf{r}}^L$ of (12.2.1), the Lagrangian description of the motion.

Another interesting example is a physical quantity we call "particle label", and denote by $\vec{\mathbf{x}}$. It is defined by requiring that its Lagrangian description be

$$\vec{\mathbf{x}}^L(\vec{x}, t) = \vec{x} \text{ for all } (\vec{x}, t) \in H \times \mathcal{R}. \tag{12.3.9}$$

From (12.3.6), its Eulerian description is $\vec{\mathbf{x}}^E$, the label function. The fact that the physical quantity $\vec{\mathbf{x}}$ is conserved at each particle turns out to be very useful in some variational formulations of the equations of motion of a fluid, although at first glance the fact seems rather trivial.

12.4 Derivatives of physical quantities

It will be useful to introduce the following notation for the derivatives of a physical quantity $f = (f^L, f^E)$.

$$\begin{aligned} D_t f^L(\vec{x}, t) &= \lim_{\tau \to 0} \frac{f^L(\vec{x}, t+\tau) - f^L(\vec{x}, t)}{\tau} \\ &= \text{time derivative of } f^L(\vec{x}, t) \text{ at the particle label } \vec{x}. \end{aligned} \tag{12.4.1}$$

$$\begin{aligned} \vec{D} f^L(\vec{x}, t) &= \vec{\nabla}\left[f^L(\cdot, t)\right](\vec{x}) = \text{ gradient tensor of the function } f^L(\cdot, t) \\ &\quad \text{at the particle label } \vec{x}. \end{aligned} \tag{12.4.2}$$

$$\begin{aligned} \partial_t f^E(\vec{r}, t) &= \lim_{\tau \to 0} \frac{f^E(\vec{r}, t+\tau) - f^E(\vec{r}, t)}{\tau} \\ &= \text{time derivative of } f^E(\vec{r}, t) \text{ at spatial position } \vec{r}. \end{aligned} \tag{12.4.3}$$

$$\begin{aligned} \vec{\partial} f^E(\vec{r}, t) &= \vec{\nabla}\left[f^E(\cdot, t)\right](\vec{r}) = \text{ gradient tensor of the function } f^E(\cdot, t) \\ &\quad \text{at the position } \vec{r} \text{ in } P. \end{aligned} \tag{12.4.4}$$

From these definitions, (12.3.7), and the chain rule, it is clear that if $\vec{r} = \vec{\mathbf{r}}^{\,L}(\vec{x}, t)$ (or $\vec{x} = \vec{\mathbf{x}}^{\,E}(\vec{r}, t)$) then

$$\vec{D} f^L(\vec{x}, t) = \vec{D}\,\vec{\mathbf{r}}^{\,L}(\vec{x}, t) \cdot \vec{\partial} f^E(\vec{r}, t) \tag{12.4.5}$$

$$\vec{\partial} f^E(\vec{r}, t) = \vec{\partial}\,\vec{\mathbf{x}}^{\,E}(\vec{r}, t) \cdot \vec{D} f^L(\vec{x}, t). \tag{12.4.6}$$

It is also useful to relate $D_t f^L$ and $\partial_t f^E$. We will calculate the former from the latter. We assume that $\vec{r}$ and $\vec{x}$ are chosen so

$$\vec{r} = \vec{\mathbf{r}}^{\,L}(\vec{x}, t).$$

Then we define $\vec{h}(\tau)$ for any real τ by

$$\vec{h}(\tau) = \vec{\mathbf{r}}^{\,L}(\vec{x}, t+\tau) - \vec{\mathbf{r}}^{\,L}(\vec{x}, t).$$

We have

$$\left\{ \begin{array}{c} \vec{h}(\tau) = \tau D_t\,\vec{\mathbf{r}}^{\,L}(\vec{x}, t) + \tau \vec{R}_1(\tau) \text{ where} \\ \\ \vec{R}_1(\tau) \to 0 \text{ as } \tau \to 0. \end{array} \right. \tag{12.4.7}$$

We also have

$$\begin{aligned} f^L(\vec{x}, t+\tau) &= f^E\left[\vec{\mathbf{r}}^{\,L}(\vec{x}, t+\tau), t+\tau\right] = f^E\left[\vec{r} + \vec{h}(\tau), t+\tau\right] \\ &= f^E\left[\vec{r} + \vec{h}(\tau), t\right] + \tau \partial_t f^E\left[\vec{r} + \vec{h}(\tau), t\right] + \tau R_2(\tau) \end{aligned}$$

where $R_2(\tau) \to 0$ as $\tau \to 0$. Define

$$R_3(\tau) = R_2(\tau) + \partial_t f^E\left[\vec{r} + \vec{h}(\tau), t\right] - \partial_t f^E\left(\vec{r}, t\right).$$

Since $\vec{h}(\tau) \to \vec{0}$ as $\tau \to 0$ therefore

$$\left\{ \begin{array}{c} R_3(\tau) \to 0 \text{ as } \tau \to 0 \text{ and} \\ \\ f^L(\vec{x}, t+\tau) = f^E\left[\vec{r} + \vec{h}(\tau), t\right] + \tau \partial_t f^E(\vec{r}, t) + \tau R_3(\tau). \end{array} \right. \tag{12.4.8}$$

Also

$$\begin{cases} f^E(\vec{r}+\vec{h},t) = f^E(\vec{r},t) + \vec{h}\cdot\vec{\partial} f^E(\vec{r},t) + \|\vec{h}\| R_4(\vec{h}) \\ \\ \qquad\qquad \text{where } R_4(\vec{h}) \to 0 \text{ as } \vec{h} \to \vec{0}. \end{cases} \tag{12.4.9}$$

Substituting (12.4.7) in (12.4.9) gives

$$\begin{aligned} f^E(\vec{r}+\vec{h},t) &= f^E(\vec{r},t) + \tau D_t\,\vec{\mathbf{r}}^{\,L}(\vec{x},t)\cdot\vec{\partial} f^E(\vec{r},t) \\ &\quad + \tau R_5(\tau) \end{aligned} \tag{12.4.10}$$

where

$$\begin{aligned} R_5(\tau) &= \vec{R}_1(\tau)\cdot\vec{\partial} f^E(\vec{r},t) \\ &\quad + \left\| D_t\,\vec{\mathbf{r}}^{\,L}(\vec{x},t) + \vec{R}_1(\tau)\right\| R_4\left[\vec{h}(\tau)\right]. \end{aligned}$$

Thus $R_5(\tau) \to 0$ as $\tau \to 0$. Moreover, $f^E(\vec{r},t) = f^L(\vec{x},t)$, so substituting (12.4.10) in (12.4.8) gives

$$\begin{aligned} f^L(\vec{x},t+\tau) &= f^L(\vec{x},t) + \tau\left[\partial_t f^E(\vec{r},t) + D_t\,\vec{\mathbf{r}}^{\,L}(\vec{x},t)\cdot\vec{\partial} f^E(\vec{r},t)\right] \\ &\quad + \tau\left[R_3(\tau) + R_5(\tau)\right] \end{aligned} \tag{12.4.11}$$

where $R_3(\tau) + R_5(\tau) \to 0$ as $\tau \to 0$. From 12.4.11 it follows immediately that

$$\begin{cases} D_t f^L(\vec{x},t) = \partial_t f^E(\vec{r},t) + D_t\,\vec{\mathbf{r}}^{\,L}(\vec{x},t)\cdot\vec{\partial} f^E(\vec{r},t) \\ \\ \quad \text{where } \vec{r} = \vec{\mathbf{r}}^{\,L}(\vec{x},t) \qquad (\text{ or } \vec{x} = \vec{\mathbf{x}}^{\,E}(\vec{r},t)). \end{cases} \tag{12.4.12}$$

The formal mathematical relationship between labels and positions is symmetrical. That is, we can interpret $\vec{r}$ as label and $\vec{x}$ as position. Therefore, we can infer immediately from (12.3.8) that

$$\begin{aligned} \partial_t f^E(\vec{r},t) &= D_t f^L(\vec{x},t) + \partial_t\,\vec{\mathbf{x}}^{\,E}(\vec{r},t)\cdot\vec{D} f^L(\vec{x},t) \\ &\qquad \text{where } r = \vec{\mathbf{r}}^{\,L}(\vec{x},t) \qquad (\text{or } \vec{x} = \vec{\mathbf{x}}^{\,E}(\vec{r},t)). \end{aligned} \tag{12.4.13}$$

Alternatively, (12.4.13) can be computed in the same way as was (12.4.12). Note the consequence of (12.4.12, 12.4.13) that for any physical quantity f,

$$D_t\,\vec{\mathbf{r}}^{\,L}(\vec{x},t)\cdot\vec{\partial}f^E(\vec{r},t)+\partial_t\,\vec{\mathbf{x}}^{\,E}(\vec{r},t)\cdot\vec{D}f^L(\vec{x},t)=0. \tag{12.4.14}$$

The notation has become somewhat cumbersome, and can be greatly simplified by introducing some new physical quantities.

Definition 12.4.45 Let $\vec{\mathbf{r}}$ be the physical quantity "particle position" (see bottom of page 146) and let f be any physical quantity. The physical quantities $\vec{v}$, $\vec{a}$, $\vec{D}f$, $\vec{\partial}f$, D_tf, ∂_tf are defined as follows:

$$\begin{aligned}
\vec{v}^L(\vec{x},t) &= D_t\,\vec{\mathbf{r}}^{\,L}(\vec{x},t) && (12.4.15)\\
\vec{a}^L(\vec{x},t) &= D_t\vec{v}^L(\vec{x},t) && (12.4.16)\\
(\vec{D}f)^L(\vec{x},t) &= \vec{D}f^L(\vec{x},t) && (12.4.17)\\
(\vec{\partial}f)^E(\vec{r},t) &= \vec{\partial}f^E(\vec{r},t) && (12.4.18)\\
(D_tf)^L(\vec{x},t) &= D_tf^L(\vec{x},t) && (12.4.19)\\
(\partial_tf)^E(\vec{r},t) &= \partial_tf^E(\vec{r},t). && (12.4.20)
\end{aligned}$$

The names of these physical quantities are as follows:

$\vec{v}$ = particle velocity

$\vec{a}$ = particle acceleration

$\vec{D}f$ = label gradient of f

$\vec{\partial}f$ = spatial gradient of f

D_tf = substantial time derivative of f = material time derivative of f = time derivative moving with the material.

∂_tf = partial time derivative of f.

With these definitions, the relations among the derivatives of a physical quantity become

$$\vec{v} = D_t \vec{\mathbf{r}} \tag{12.4.21}$$

$$\vec{a} = D_t \vec{v} \tag{12.4.22}$$

$$\vec{D} f = \vec{D} \vec{\mathbf{r}} \cdot \vec{\partial} f \tag{12.4.23}$$

$$D_t f = \partial_t f + \vec{v} \cdot \vec{\partial} f \tag{12.4.24}$$

$$\vec{\partial} f = \vec{\partial} \vec{\mathbf{x}} \cdot \vec{D} f \tag{12.4.25}$$

$$\partial_t f = D_t f + \partial_t \vec{\mathbf{x}} \cdot \vec{D} f \tag{12.4.26}$$

The function $\vec{v}^E$ is called the Eulerian description of the motion of the continuum.[1] If the Lagrangian description is known, the Eulerian description can be obtained from (12.4.21). The converse is true, in the following sense.

Remark 12.4.56 Suppose the Eulerian description $\vec{v}^E$ of a continuum is given. Suppose that at one instant t_0, the function $\vec{\mathbf{r}}^{\,L}(\cdot, t_0)$ is known; that is, for each $\vec{x}$, the position at time t_0 of the particle labelled $\vec{x}$ is known. Then the Lagrangian description of the motion can be calculated. That is, the position $\vec{\mathbf{r}}^{\,L}(\vec{x}, t)$ of the particle labelled $\vec{x}$ can be calculated for every time t.

Proof:

By definition of $\vec{v}$, we have for $\vec{r} = \vec{\mathbf{r}}^{\,L}(\vec{x}, t)$ the relation $D_t \vec{\mathbf{r}}^{\,L}(\vec{x}, t) = \vec{v}^E(\vec{r}, t)$ or

$$D_t \vec{\mathbf{r}}^{\,L}(\vec{x}, t) = \vec{v}^E \left[\vec{\mathbf{r}}^{\,L}(\vec{x}, t), t \right]. \tag{12.4.27}$$

But (12.4.27) is an ordinary differential equation in t if $\vec{x}$ is fixed. Therefore, for any fixed particle label $\vec{x}$, we can solve (12.4.27) for $\vec{\mathbf{r}}^{\,L}(\vec{x}, t)$ if we know $\vec{\mathbf{r}}^{\,L}(\vec{x}, t_0)$ for one t_0 . QED

Corollary 12.4.43 If $\vec{v}^E$ is known, then the Lagrangian description of the motion can be calculated without further information if the particles are "t_0-position labelled", i.e. if their labels are their positions at some tme t_0.

[1] If $\partial_t \vec{v}^E = \vec{0}$, the motion is called "steady".

Proof:

With t_0-position labelling $\vec{\mathbf{r}}^L(\vec{x}, t_0) = \vec{x}$, so $\vec{\mathbf{r}}^L(\vec{x}, t_0)$ is known for all particle labels $\vec{x}$.

12.5 Rigid body motion

For this discussion we need a lemma about linearity:

Lemma 12.5.27 Suppose V and W are real vector spaces and $f : V \to W$. Suppose f has the following properties, for any **nonzero** $\vec{x}, \vec{y}, \vec{z} \in V$:

i) $f(\vec{0}) = \vec{0}$

ii) $f(-\vec{x}) = -f(\vec{x})$

iii) $f(a\vec{x}) = af(\vec{x})$ for any **positive** real a (i.e., $a \in R, a > 0$).

iv) If $\vec{x} + \vec{y} + \vec{z} = \vec{0}$ then $f(\vec{x}) + f(\vec{y}) + f(\vec{z}) = \vec{0}$.

In that case, f is linear.

Proof:

a) For any $c \in R$ and $\vec{v} \in V$, $f(c\vec{v}) = cf(\vec{v})$

a1) If $c\vec{v} = \vec{0}$, either $c = 0$ or $\vec{v} = \vec{0}$, so $f(c\vec{v}) = \vec{0} = cf(\vec{v})$.

a2) If $c\vec{v} \neq \vec{0}$ and $c > 0$, use iii) above.

a3) If $c\vec{v} \neq \vec{0}$ and $c < 0$, then $f((-c)\vec{v}) = (-c)f(\vec{v})$ by iii) above. But $f((-c)\vec{v}) = f(-c\vec{v}) = -f(c\vec{v})$ by ii) above.

Hence $-f(c\vec{v}) = -cf(\vec{v})$.

b) For any $\vec{u}, \vec{v} \in V$, $f(\vec{u} + \vec{v}) = f(\vec{u}) + f(\vec{v})$.

b1) If $\vec{u}$ or $\vec{v}$ is $\vec{0}$, one of $f(\vec{u})$ and $f(\vec{v})$ is $\vec{0}$ by i) above. E.g., if $\vec{u} = \vec{0}$, $f(\vec{u} + \vec{v}) = f(\vec{v}) = f(\vec{u}) + f(\vec{v})$.

b2) If $\vec{u}$ and $\vec{v}$ are nonzero but $\vec{u}+\vec{v}=\vec{0}$, then $\vec{v}=-\vec{u}$ so $f(\vec{u}+\vec{v})= f(\vec{0})=\vec{0}=f(\vec{u})-f(\vec{u})=f(\vec{u})+f(-\vec{u})=f(\vec{u})+f(\vec{v})$.

b3) If $\vec{u}$, $\vec{v}$ and $\vec{u}+\vec{v}$ are nonzero, let $\vec{w}=\vec{u}+\vec{v}$. Then $\vec{u}+\vec{v}+(-\vec{w})=\vec{0}$ so by iv), $f(\vec{u})+f(\vec{v})+f(-\vec{w})=\vec{0}$. By ii) this is $f(\vec{u})+f(\vec{v})-f(\vec{w})=\vec{0}$, or $f(\vec{w})=f(\vec{u})+f(\vec{v})$.

QED

The motion of a continuum is called "rigid body motion" if the distance separating every pair of particles in the continuum is independent of time. We want to study rigid body motions.

Step 1: Choose one particular particle in the continuum, and call it the pivot particle.

Step 2: Choose $t_0 \in \mathcal{R}$. Choose a reference frame for real physical space P so that its origin is at the position of the pivot particle at time t_0.

Step 3: Introduce t_0-position labelling to give the Lagrangian description of the motion, $\vec{\mathbf{r}}^L$. Let H denote the open subset of $L=P$ consisting of the particle labels. The pivot particle is labelled $\vec{0}$, so $\vec{0}\in H$, and

$$\vec{\mathbf{r}}^L(\vec{x},t_0)=\vec{x} \text{ for all } \vec{x}\in H.$$

Step 4: For each $t\in\mathcal{R}$, define a mapping $\vec{r}_t : H \to P$ by

$$\vec{r}_t(\vec{x}) = \vec{\mathbf{r}}^L(\vec{x},t) - \vec{\mathbf{r}}^L(\vec{0},t)$$

and define $\vec{R}:\mathcal{R}\to\mathcal{P}$ as $\vec{R}(t)=\vec{\mathbf{r}}^L(\vec{0},t)$. Thus

$$\vec{\mathbf{r}}^L(\vec{x},t)=\vec{R}(t)+\vec{r}_t(\vec{x}) \tag{12.5.1}$$

for all $\vec{x}\in H$ and $t\in\mathcal{R}$. Moreover, for any $t\in\mathcal{R}$ and any particle labels $\vec{x},\vec{y}\in H$, $\vec{\mathbf{r}}^L(\vec{x},t)-\vec{\mathbf{r}}^L(\vec{y},t)=\vec{r}_t(\vec{x})-\vec{r}_t(\vec{y})$, so, by the definition of a rigid body,

$$\|\vec{r}_t(\vec{x})-\vec{r}_t(\vec{y})\|=\|\vec{x}-\vec{y}\| \text{ for all } t\in\mathcal{R} \text{ and } \vec{x},\vec{y}\in H. \tag{12.5.2}$$

Also, from the definition of $\vec{R}(t)$ and (12.5.1),

$$\vec{r}_t(\vec{0}) = \vec{0} \text{ for all } t \in \mathcal{R}. \tag{12.5.3}$$

The main result implied by the above is contained in

Theorem 12.5.27 Suppose $f : H \to P$ where H is an open subset of real Euclidean vector space P. Suppose $f(\vec{0}) = \vec{0}$ and for any $\vec{x}, \vec{y} \in H$, $\|f(\vec{x}) - f(\vec{y})\| = \|\vec{x} - \vec{y}\|$. Then there is a unique $L \in \Theta(P)$ (i.e., L is linear and orthogonal) such that for all $\vec{x} \in H$

$$f(\vec{x}) = \vec{x}\cdot \overleftrightarrow{L} = L(\vec{x}). \tag{12.5.4}$$

Proof:

Lemma 1: $\|f(\vec{x})\| = \|\vec{x}\|$ for any $\vec{x} \in H$.

Proof: $\|f(\vec{x})\| = \|f(\vec{x}) - \vec{0}\| = \|f(\vec{x}) - f(\vec{0})\| = \|\vec{x} - \vec{0}\| = \|\vec{x}\|$.

Lemma 2: $f(\vec{x}) \cdot f(\vec{y}) = \vec{x} \cdot \vec{y}$ for all $\vec{x}, \vec{y} \in H$.

Proof: $\|f(\vec{x}) - f(\vec{y})\|^2 = \|\vec{x} - \vec{y}\|^2$ so

$$\|f(\vec{x})\|^2 - 2f(\vec{x}) \cdot f(\vec{y}) + \|f(\vec{y})\|^2 = \|\vec{x}\|^2 - 2\vec{x} \cdot \vec{y} + \|\vec{y}\|^2.$$

But by lemma 1, $\|f(\vec{x})\|^2 = \|\vec{x}\|^2$ and $\|f(\vec{y})\|^2 = \|\vec{y}\|^2$.

Lemma 3: Suppose $a^1, \cdots, a^N \in \mathcal{R}$, $\vec{x}_1, \cdots, \vec{x}_N \in H$, and $a^i\vec{x}_i \in H$. Then $f(a^i\vec{x}_i) = a^i f(\vec{x}_i)$.

Proof: $\|f(a^i\vec{x}_i) - a^i f(\vec{x}_i)\|^2 = \|f(a^i\vec{x}_i)\|^2 - 2a^j f(\vec{x}_j)\cdot f(a^i\vec{x}_i) + \|a^i f(\vec{x}_i)\|^2 = \|a^i\vec{x}_i\|^2 - 2a^j\,[f(\vec{x}_j) \cdot f(a^i\vec{x}_i)]\; + a^i a^j f(\vec{x}_i) \cdot f(\vec{x}_j)$. By lemma 2, this is $\|a^i\vec{x}_i\|^2 - 2a^j\,[x_j \cdot (a^i\vec{x}_i)] + a^i a^j(\vec{x}_i \cdot \vec{x}_j) = \|a^i\vec{x}_i\|^2 - a^i a^j \vec{x}_i \cdot \vec{x}_j = 0$.

Lemma 4: Choose $\delta > 0$ so small that if $\|\vec{x}\| < \delta$ then $\vec{x} \in H$. Define $L : P \to P$ as follows:

i) $L(\vec{0}) = \vec{0}$

ii) $L(\vec{x}) = (\|\vec{x}\|/\delta) f(\delta\vec{x}/\|\vec{x}\|)$ if $\vec{x} \neq \vec{0}$.

Then $L(\vec{x}) = f(\vec{x})$ for all $\vec{x} \in H$, and L satisfies conditions i)-iv) of lemma 12.5.27.

Proof: First, if $\vec{x} \in H$ then in lemma 3 take $N = 1$, $a' = \delta/\|\vec{x}\|$, $\vec{x}_1 = \vec{x}$. Lemma 3 then implies $f(\delta\vec{x}/\|\vec{x}\|) = \delta/\|\vec{x}\| f(\vec{x})$, and from ii) above, obviously $L(\vec{x}) = f(\vec{x})$. Now for i)-iv) of lemma 12.5.27.

i) is obvious

ii) $L(-\vec{x}) = (\|-\vec{x}\|/\delta)\ f(-\delta\vec{x}/\|-\vec{x}\|) = (\|\vec{x}\|/\delta) f(-\delta\vec{x}/\|\vec{x}\|)$. In lemma 3 above, take $N = 1$, $a' = -1$, $\vec{x}_1 = \delta\vec{x}/\|\vec{x}\|$. Then $f(-\delta\vec{x}/\|\vec{x}\|) = -f(\delta\vec{x}/\|\vec{x}\|)$. Thus $L(-\vec{x}) = -L(\vec{x})$.

iii) $L(a\vec{x}) = (\|a\vec{x}\|/\delta) f(\delta a\vec{x}/\|a\vec{x}\|) = a(\|\vec{x}\|/\delta) f(\delta a\vec{x}/a\|\vec{x}\|)$
$= a(\|\vec{x}\|/\delta) f(\delta\vec{x}/\|\vec{x}\|) = aL(\vec{x})$ if $a > 0$ and $\vec{x} \neq \vec{0}$.

iv) Suppose $\vec{x} + \vec{y} + \vec{z} = \vec{0}$ and $\vec{x}, \vec{y}, \vec{z}$ nonzero. Choose η so small ($\eta \in \mathcal{R}$, $\quad \eta > 0$) that $\eta\vec{x} + \eta\vec{y} + \eta\vec{z} \in H$. Then $\eta\vec{x} + \eta\vec{y} + \eta\vec{z} = \vec{0}$, so, using lemma 3 , $\vec{0} = f(\vec{0}) = f(\eta\vec{x} + \eta\vec{y} + \eta\vec{z}) = f(\eta\vec{x}) + f(\eta\vec{y}) + f(\eta\vec{z})$. Since $\eta\vec{x} \in H$, $f(\eta\vec{x}) = L(\eta\vec{x})$, and similarly for $\vec{y}, \vec{z}$. Thus $\vec{0} = L(\eta\vec{x}) + L(\eta\vec{y}) + L(\eta\vec{z})$. By iii) above, $L(\eta\vec{x}) = \eta L(\vec{x})$, etc., so $\vec{0} = \eta[L(\vec{x}) + L(\vec{y}) + L(\vec{z})]$. Thus $L(\vec{x}) + L(\vec{y}) + L(\vec{z}) = \vec{0}$.

It follows from lemma 4 above and lemma 12.5.27 that $L : P \to P$ is linear. For any $\vec{x} \neq \vec{0}$ in P, choose η so small ($\eta \in R$, $\eta > 0$) that $\eta\vec{x} \in H$. Then $\eta(\|L(\vec{x})\| - \|\vec{x}\|) = \|\eta L(\vec{x})\| - \|\eta\vec{x}\| = \|L(\eta\vec{x})\| - \|\eta\vec{x}\| = \|f(\eta\vec{x})\| - \|\eta\vec{x}\| = \|\eta\vec{x}\| - \|\eta\vec{x}\| = 0$.

Thus $\|L(\vec{x})\| = \|\vec{x}\|$ for all $\vec{x} \in P$. Therefore $L \in \Theta(P)$. QED

By theorem (12.5.27), for each $t \in R$ there is an orthogonal tensor $\overleftrightarrow{L}(t) \in P \otimes P$ such that in equation (12.5.1) we can write

$$\mathbf{\vec{r}}^{\,L}(\vec{x}, t) = \vec{R}(t) + \vec{x}\cdot \overleftrightarrow{L}(t) \qquad \text{for all } \vec{x} \in H. \tag{12.5.5}$$

Then clearly

$$\vec{D}\,\mathbf{\vec{r}}^{\,L}(\vec{x}, t) = \overleftrightarrow{L}(t) \qquad \text{for all } \vec{x} \in H. \tag{12.5.6}$$

We are assuming that $\vec{D}\,\mathbf{\vec{r}}^{\,L}(\vec{x},t)$ is continuous in $\vec{x}$ and t, so $\overleftrightarrow{L}(t)$ is a continuous function of t. At the labelling time we have $\vec{R}(t_0) = \mathbf{\vec{r}}^{\,L}(\vec{0},t_0) = \vec{0}$ and so $\vec{x}\cdot \overleftrightarrow{L}(t_0) = \mathbf{\vec{r}}^{\,L}(\vec{x},t_0) = \vec{x}$. Thus

$$\overleftrightarrow{L}(t_0) = \overleftrightarrow{I}_P \qquad\qquad \vec{R}(t_0) = \vec{0} \tag{12.5.7}$$

and det $\overleftrightarrow{L}(t_0) = +1$. But det $L(t)$ depends continuously on t since $L(t)$ does, and det $L(t)$ is always $+1$ or -1 since $L(t)$ is orthogonal. Therefore

$$\det L(t) = +1, \quad L(t) \in \Theta^+(P) \text{ for all } t \in R. \tag{12.5.8}$$

Every rigid motion can be written as in (12.5.5) with (12.5.7) and (12.5.8).

Now let us find the Eulerian description of the motion whose Lagrangian description is (12.5.5). We have

$$\vec{v}^L(\vec{x},t) = D_t\,\mathbf{\vec{r}}^{\,L}(\vec{x},t) = \delta_t\vec{R}(t) + \vec{x}\cdot\delta_t\overleftrightarrow{L}(t). \tag{12.5.9}$$

[**Note:** Since $\vec{R}$ and $\overleftrightarrow{L}$ are functions of t alone, we write their time derivatives as $\delta_t\vec{R}$ and $\delta_t\overleftrightarrow{L}$, so as not to confuse them with ∂_t or D_t.]

To obtain the Eulerian description from (12.5.9) we must express $\vec{x}$ in terms of $\vec{r}$, where as usual

$$\vec{r} = \mathbf{\vec{r}}^{\,L}(\vec{x},t), \text{ or}$$

$$\vec{r} = \vec{R}(t) + \vec{x}\cdot\overleftrightarrow{L}(t). \tag{12.5.10}$$

Since $\overleftrightarrow{L} \in \Theta^+(P)$, $\overleftrightarrow{L}^T = \overleftrightarrow{L}^{-1}$. Therefore the solution of (12.5.10) is

$$\vec{x} = \left[\vec{r} - \vec{R}(t)\right]\cdot\overleftrightarrow{L}^T. \tag{12.5.11}$$

[**Note:** It follows that $\mathbf{x}^{\,E}(\vec{r},t) = [\vec{r} - \vec{R}(t)]\cdot\overleftrightarrow{L}(t)^T$.] Substituting this (12.5.9) and using $\vec{v}^L(\vec{x},t) = \vec{v}^E(\vec{r},t)$ gives

$$\vec{v}^E(\vec{r},t) = \delta_t\vec{R}(t) + \left[\vec{r} - \vec{R}(t)\right]\cdot\overleftrightarrow{L}(t)^T\cdot\delta_t\overleftrightarrow{L}(t).$$

We introduce these abbreviations:

$$\vec{V}(t) = \delta_t\vec{R}(t) \tag{12.5.12}$$

$$\overleftrightarrow{\Omega}(t) = \overleftrightarrow{L}(t)^T \cdot \delta_t \overleftrightarrow{L}(t). \tag{12.5.13}$$

Then the Eulerian description of the motion is

$$\vec{v}^E(\vec{r}, t) = \vec{V}(t) + \left[\vec{r} - \vec{R}(t)\right] \cdot \overleftrightarrow{\Omega}(t). \tag{12.5.14}$$

Remark 12.5.57 $[\delta_t \overleftrightarrow{L}(t)]^T = \delta_t L^T(t)$.

Proof:

For any $\vec{x} \in P$, $\vec{x} \cdot \overleftrightarrow{L}(t) = \overleftrightarrow{L}^T(t) \cdot \vec{x}$, so $\vec{x} \cdot [\delta_t \overleftrightarrow{L}(t)] = [\delta_t \overleftrightarrow{L}^T(t)] \cdot \vec{x}$. But $\vec{x} \cdot [\delta_t \overleftrightarrow{L}] = (\delta_t \overleftrightarrow{L})^T \cdot \vec{x}$, so $[\delta_t \overleftrightarrow{L}(t)]^T \cdot \vec{x} = [\delta_t \overleftrightarrow{L}^T(t)] \cdot \vec{x}$ for all $\vec{x} \in P$.

Hence, we have the remark

Remark 12.5.58 $\overleftrightarrow{\Omega}(t) = -\overleftrightarrow{\Omega}(t)^T$. I.e., $\overleftrightarrow{\Omega}(t)$ is antisymmetric.

Proof:

Since $\overleftrightarrow{L}(t)$ is orthogonal, $\overleftrightarrow{L}^T(t) \cdot \overleftrightarrow{L}(t) = \overleftrightarrow{I}_P$ for all t. Hence, $(\delta_t \overleftrightarrow{L}^T) \cdot \overleftrightarrow{L} + \overleftrightarrow{L}^T \cdot (\delta_t \overleftrightarrow{L}) = \delta_t \overleftrightarrow{I}_P = \overleftrightarrow{0}$ for all t. But $\overleftrightarrow{\Omega} = \overleftrightarrow{L}^T \cdot \delta_t \overleftrightarrow{L}$ and $\overleftrightarrow{\Omega}^T = (\delta_t \overleftrightarrow{L})^T \cdot (\overleftrightarrow{L}^T)^T = (\delta_t \overleftrightarrow{L}^T) \cdot (\overleftrightarrow{L})$. Thus $\overleftrightarrow{\Omega}^T + \overleftrightarrow{\Omega} = \overleftrightarrow{0}$. QED

Now define

$$\overleftrightarrow{\Omega}(t) = \frac{1}{2} A\langle 2\rangle \overleftrightarrow{\Omega}(t) \tag{12.5.15}$$

where A is the unimodular alternating tensor we have chosen to orient real physical space P. By exercise 6,

$$\overleftrightarrow{\Omega}(t) = A \cdot \vec{\Omega}(t). \tag{12.5.16}$$

We claim that for any vector $\vec{x} \in P$ we have

$$\vec{x} \cdot \overleftrightarrow{\Omega}(t) = \vec{\Omega}(t) \times \vec{x}. \tag{12.5.17}$$

Proof:

Take components relative to a pooob. Then

$$\begin{aligned} \left(\vec{x} \cdot \overleftrightarrow{\Omega}\right)_j &= \left(\vec{x} \cdot A \cdot \vec{\Omega}\right)_j = x_i \varepsilon_{ijk} \Omega_k \\ &= \varepsilon_{kij} \Omega_k x_i = \left(A\langle 2\rangle \vec{\Omega}\vec{x}\right)_j = \left(\vec{\Omega} \times \vec{x}\right)_j . \end{aligned}$$

Thus (12.5.14) can be written

$$\vec{v}^E(\vec{r},t) = \vec{V}(t) + \vec{\Omega}(t) \times \left[\vec{r} - \vec{R}(t)\right] \qquad \left(\vec{V} = \delta_t \vec{R}\right). \tag{12.5.18}$$

The vector $\vec{\Omega}(t)$ is called the angular velocity of the motion relative to the pivot particle. Obviously $\vec{V}(t)$ is the velocity of the pivot particle, and $\vec{R}(t)$ is its position.

An obvious question not answered by the foregoing discussion is this: if we are given two arbitrary differentiable functions $\vec{R}: \mathcal{R} \to P$, $\vec{\Omega}: \mathcal{R} \to P$, and define $\vec{V} = \delta_t \vec{R}$, then (12.5.18) gives the Eulerian description of a continuum motion. Is this always a rigid body motion, or do rigid body motions produce particular kinds of functions $\vec{R}: \mathcal{R} \to P$ and $\vec{\Omega}: \mathcal{R} \to P$?

In fact, any choice of $\vec{R}$ and $\vec{\Omega}$ will make (12.5.18) the Eulerian description of a rigid body motion. To see this, use the given $\overleftrightarrow{\Omega}(t)$ as in (12.5.16). $\overleftrightarrow{\Omega}(t)$ is, of course, antisymmetric. Now define $\overleftrightarrow{L}(t)$ as follows:

$$\delta_t \overleftrightarrow{L} = \overleftrightarrow{L} \cdot \overleftrightarrow{\Omega}. \tag{12.5.19}$$

(This is obtained from 12.5.13 as if $\overleftrightarrow{L}^T = \overleftrightarrow{L}^{-1}$, but we don't assume that.) We obtain $\overleftrightarrow{L}(t)$ by solving (12.5.19) with the initial condition

$$\overleftrightarrow{L}(t_0) = \overleftrightarrow{I}_P. \tag{12.5.20}$$

Remark 12.5.59 $\overleftrightarrow{L}(t)$ obtained from (12.5.19,12.5.20) is proper orthogonal.

Proof:

$$\begin{aligned}
\delta_t \left(\overleftrightarrow{L} \cdot \overleftrightarrow{L}^T\right) &= \delta_t \overleftrightarrow{L} \cdot \overleftrightarrow{L}^T + \overleftrightarrow{L} \cdot \left(\delta_t \overleftrightarrow{L}\right)^T \\
&= \overleftrightarrow{L} \cdot \overleftrightarrow{\Omega} \cdot \overleftrightarrow{L}^T + \overleftrightarrow{L} \cdot \left(\overleftrightarrow{L} \cdot \overleftrightarrow{\Omega}\right)^T \\
&= \overleftrightarrow{L} \cdot \overleftrightarrow{\Omega} \cdot \overleftrightarrow{L}^T + \overleftrightarrow{L} \cdot \overleftrightarrow{\Omega}^T \cdot \overleftrightarrow{L}^T \\
&= \overleftrightarrow{L} \cdot \left(\overleftrightarrow{\Omega} + \overleftrightarrow{\Omega}^T\right) \cdot \overleftrightarrow{L}^T = \overleftrightarrow{L} \cdot \overleftrightarrow{O} \cdot \overleftrightarrow{L}^T = \overleftrightarrow{O}.
\end{aligned}$$

Thus, $\overleftrightarrow{L}(t) \cdot \overleftrightarrow{L}(t)^T$ is independent of t. At $t = t_0$, it is $\overleftrightarrow{I}_P$, so it is always $\overleftrightarrow{I}_P$. This proves $\overleftrightarrow{L}(t) \in \Theta(P)$. And det $\overleftrightarrow{L}(t)$ is continuous in t and always ± 1 and $+1$ at t_0, so it is always $+1$. Thus $\overleftrightarrow{L}(t) \in \Theta^+(P)$ for all t. QED

Now we can write (12.5.18) in the form (12.5.14). As we have seen in deriving (12.5.18), the motion whose Lagrangian description is (12.5.5) then has Eulerian description (12.5.14) or, equivalently (12.5.18). But as long as $L(t) \in \Theta^+(P)$ for all t, (12.5.5) is a rigid-body motion. Indeed, for any $\vec{x}$ and $\vec{y} \in P$, $\|\vec{\mathbf{r}}^L(\vec{x},t) - \vec{\mathbf{r}}^L(\vec{y},t)\| = \|(\vec{x}-\vec{y})\cdot \overleftrightarrow{L}(t)\| = \|\vec{x}-\vec{y}\|$ if $L(t) \in \Theta(P)$.

12.6 Relating continuum models to real materials

In a rock there is some jumping of atoms out of lattice positions, i.e., "self-diffusion", but crystal structures stay mostly intact. In water and air, however, individual molecules move at random. After t seconds, two water molecules which were initially neighbors are separated, on average, by about $(10^{-5}t)^{1/2}$ cm and two air molecules at sea level are separated by $(0.2t)^{1/2}$ cm. What is the physical meaning of a "particle" in water or air? We cannot mean a molecule, because then $\vec{\mathbf{r}}^L(\vec{x},t)$ would quickly become a very discontinuous function of $\vec{x}$.

In reality, continuum models are "two scale" approximations. To see what we mean, let us examine how to make a continuum model of a gas in which the individual molecules are moderately well modelled as randomly colliding very small hard spheres, with all their mass at the center. Let $B(\vec{r},\lambda)$ be the open ball of radius λ with center at $\vec{r}$. Let $M_t[B(\vec{r},\lambda)]$ be the sum of the masses of all the molecules whose centers are in $B(\vec{r},\lambda)$ at time t. Let $|B(\vec{r},\lambda)| = 4\pi\lambda^3/3 =$ volume of $B(\vec{r},\lambda)$. The average density in $B(\vec{r},\lambda)$ at time t is

$$\langle\rho\rangle_\lambda(\vec{r},t) = \frac{1}{|B(\vec{r},\lambda)|} M_t\left[B(\vec{r},\lambda)\right]. \tag{12.6.1}$$

To use a continuum model, we must be able to choose λ so large that the jumps in $\langle\rho\rangle_\lambda(\vec{r},t)$ as a function of $\vec{r}$ and t, which occur when individual molecules enter or leave $B(\vec{r},\lambda)$ as $\vec{r}$ and t vary, are very small fractions of $\langle\rho\rangle_\lambda(\vec{r},t)$. But at the same time, λ must be so small that (12.6.1) does not sample very different physical conditions inside $B(\vec{r},\lambda)$. That is, λ must be so small that $\|\delta\vec{r}\|$ can be many times λ and yet $\langle\rho\rangle_\lambda(\vec{r}+\delta\vec{r},t)$

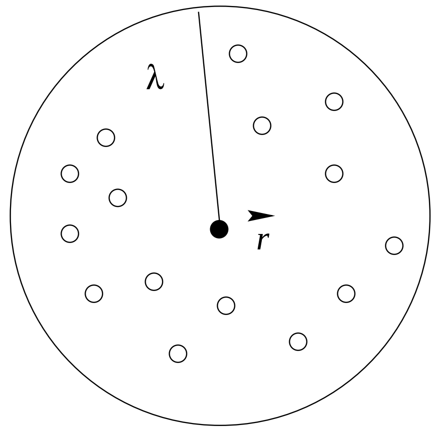


Figure 12.1:

will differ from $\langle\rho\rangle_\lambda(\vec{r},t)$ by only a very small fraction. Therefore, the average physical conditions must not change appreciably over a length λ.

For example, if we want to use a continuum approximation to study sound waves of wavelength Λ in air, we must use $\lambda << \Lambda$ in (12.6.1), and yet we must have $\lambda >>$ average distance between nearest neighbors. Actually, to do all of continuum mechanics accurately (including momentum and heat) we must have $\lambda >>$ mean free path of air molecules. We cannot find such a λ to use in (12.6.1) unless $\Lambda >>$ mean free path of air molecules. This mfp varies from 10^{-5} cm at sea level to 10 cm at 100 km altitude, and determines the shortest sound wave treatable by continuum mechanics.

In the spirit of (9.2.3), we would define the average momentum in $B(\vec{r},\lambda)$ at time t as

$$\langle\vec{p}\rangle_\lambda(\vec{r},t) = \frac{1}{|B(\vec{r},\lambda)|}\vec{P}_t\left[B(\vec{r},\lambda)\right] \tag{12.6.2}$$

where $\vec{P}_t[B(\vec{r},\lambda)]$ is the sum of the momenta $m\vec{v}$ for all the molecules in $B(\vec{r},\lambda)$ at time t. Then we *define* the average velocity at $\vec{r},t$ to be

$$\langle\vec{v}\rangle_\lambda(\vec{r},t) = \frac{\langle\vec{p}\rangle_\lambda(\vec{r},t)}{\langle\rho\rangle_\lambda(\vec{r},t)} = \frac{\vec{P}_t\left[B(\vec{r},\lambda)\right]}{M_t\left[B(\vec{r},\lambda)\right]}. \tag{12.6.3}$$

If Λ is the shortest length scale in our gas over which occur appreciable fractional changes in average physical properties, and if $\Lambda >>$ mean free path, we can choose any λ in $mfp << \lambda << \Lambda$, and model our gas as a continuum whose Eulerian velocity function is

$$\vec{v}^E(\vec{r},t) = \langle\vec{v}\rangle_\lambda(\vec{r},t). \tag{12.6.4}$$

The Lagrangian description of this continuum is obtained by choosing a labelling scheme and then integrating (12.4.27) with $\vec{v}^E$ given by (12.6.4). Usually t_0-position labelling is convenient.

Evidently in a gas, the "particles" of the continuum model are fictitious. In a solid, they can be thought of as individual molecules, at least for times short enough that self diffusion is unimportant. An example of the breakdown of the continuum model is a shock wave in air. They are considerably less thick than one mean free path. In fact, continuum models work better inside shocks than is at present understood.

Chapter 13

Conservation Laws in a Continuum

In this whole chapter we will use the following notation. (L, A_L) is oriented label space and (P, A_P) is oriented real physical space. H is the open subset of L consisting of the labels $\vec{x}$ of all the particles making up the continuum and H' is any open subset of H with piecewise smooth boundary $\partial H'$ (i.e. $\partial H'$ consists of a finite number of smooth pieces). The open set in P occupied by the particles at time t will be written $K(t)$, and the open subset of $K(t)$ consisting of the particles with labels in H' will be written $K'(t)$. As t varies, $K(t)$ and $K'(t)$ will change position in P, but will always contain the same particles. The sets $K(t)$ and $K'(t)$ are said to move with the material. The outward normal velocity $W(\vec{r}, t)$ of $\partial K'(t)$ at $\vec{r} \in \partial K'(t)$, to be used in (11.6.1), is $W(\vec{r}, t) = \hat{n}_P(\vec{r}, t) \cdot \vec{v}^E(\vec{r}, t)$, where $\vec{v}$ is the particle velocity and $\hat{n}_P(\vec{r}, t)$ is the unit outward normal to $\partial K'(t)$ at $\vec{r} \in \partial K'(t)$.

We will use $dV_L(\vec{x})$ to denote a very small open subset of H' containing the particle label $\vec{x}$. We will denote the positive numerical volume in L of this small set by the same symbol, $dV_L(\vec{x})$. We use $dV_P(\vec{r})$ to denote a very small open subset of $K'(t)$ containing the particle position $\vec{r}$; we also use $dV_P(\vec{r})$ for the positive numerical volume of this small set in P. We will always assume when using this notation for the sets that $\vec{r} = \vec{\mathbf{r}}^{\,L}(\vec{x}, t)$ and $dV_P(\vec{r}) = [\vec{\mathbf{r}}^{\,L}(\cdot, t)][dV_L(\vec{x})]$, so $dV_P(\vec{r})$ consists of the positions at time t of the particles whose labels are in $dV_L(\vec{x})$. It follows that, for the numerical volumes,

$$dV_P(\vec{r}) = \left|\det \vec{D}\, \vec{\mathbf{r}}^{\,L}(\vec{x}, t)\right| dV_L(\vec{x}). \tag{13.0.1}$$

We assume that the reader knows this Jacobian formula from calculus. (We will accidentally prove it later.) Note that ^{L} means Lagrangian description and $_L$ means label space.

13.1 Mass conservation

13.1.1 Lagrangian form

The mass of the collection of particles with labels in $dV_L(\vec{x})$ is

$$dm = \rho^E(\vec{r}, t) dV_P(\vec{r}) \tag{13.1.1}$$

where $\rho^E(\vec{r}, t)$ is the mass per unit of volume in physical space. This is our mathematical model of how mass is distributed in a continuum.

Because of (13.0.1), we can also write

$$dm = \tilde{\rho}^L(\vec{x}, t) dV_L(\vec{x}) \tag{13.1.2}$$

if we set $\tilde{\rho}^L(\vec{x}, t) = \rho^E(\vec{r}, t)|\det \vec{D}\,\vec{\mathbf{r}}^{\,L}(\vec{x}, t)]$. But $\rho^E(\vec{r}, t) = \rho^L(\vec{x}, t)$ since we are assuming $\vec{r} = \vec{\mathbf{r}}^{\,L}(\vec{x}, t)$. Therefore $\tilde{\rho}^L(\vec{x}, t) = \rho^L(\vec{x}, t)|\det \vec{D}\,\vec{\mathbf{r}}^{\,L}(\vec{x}, t)| = (\rho|\det \vec{D}\,\vec{\mathbf{r}}\,|)^L(\vec{x}, t)$, so

$$\tilde{\rho} = \rho|\det \vec{D}\,\vec{\mathbf{r}}\,|. \tag{13.1.3}$$

The physical quantity $\tilde{\rho}$ is the mass per unit of volume in label space.

The material with labels in H' always consists of the same particles, so it cannot change its mass. Therefore, dm in (13.1.2) must be independent of time. Since $dV_L(\vec{x})$ is defined in a way independent of time, it follows that $\tilde{\rho}^L(\vec{x}, t)$ must be independent of t. Therefore, for all $\vec{x} \in H$ and all t, $t_0 \in \mathcal{R}$,

$$\tilde{\rho}^L(\vec{x}, t) = \tilde{\rho}^L(\vec{x}, t_0). \tag{13.1.4}$$

An equivalent equation is $D_t\tilde{\rho}^L(\vec{x}, t) = 0$, or $(D_t\rho)^L(\vec{x}, t) = 0$, so

$$D_t\tilde{\rho} = 0. \tag{13.1.5}$$

Either equation (13.1.4) or (13.1.5) expresses the content of a physical law, the law of conservation of mass. The mathematical identity (13.1.3) makes this law useful. From (13.1.3) and (13.1.4) we deduce

$$\rho^L(\vec{x},t)\left|\det \vec{D}\,\vec{\mathbf{r}}^{\,L}(\vec{x},t)\right| = \rho^L(\vec{x},t_0)\left|\det \vec{D}\,\vec{\mathbf{r}}^{\,L}(\vec{x},t_0)\right| \tag{13.1.6}$$

for all $\vec{x} \in H$ and all $t,\ t_0 \in R$. If we use t_0-position labelling, then $\vec{\mathbf{r}}^{\,L}(\vec{x},t_0) = \vec{x}$, so $\vec{D}\,\vec{\mathbf{r}}^{\,L}(\vec{x},t_0) = \overset{\leftrightarrow}{I}_P$ and $\det \vec{D}\,\vec{\mathbf{r}}^{\,L}(\vec{x},t_0) = \det \overset{\leftrightarrow}{I}_p = \det I_P = 1$. Thus, with t_0-position labelling,

$$\rho^L(\vec{x},t)\left|\det \vec{D}\,\vec{\mathbf{r}}^{\,L}(\vec{x},t)\right| = \rho^L(\vec{x},t_0). \tag{13.1.7}$$

13.1.2 Eulerian form of mass conservation

An informal physical argument like (13.1.1, 13.1.2) is possible here, but rather confusing. The bookkeeping becomes clearer if we account for real, finite masses and volumes. The total mass in the set $K'(t)$ in physical space P at time t is

$$M^E\left[K'(t)\right] = \int_{K'(t)} dV_P(\vec{r})\rho^E(\vec{r},t). \tag{13.1.8}$$

The physical law of mass conservation is

$$\frac{d}{dt}M^E\left[K'(t)\right] = 0, \tag{13.1.9}$$

because $K'(t)$ always consists of the same material.

Using (11.6.1) with $w = \hat{n}_P \vec{v}^E(\vec{r},t)$ we have the mathematical identity

$$\frac{d}{dt}M^E\left[K'(t)\right] = \int_{K'(t)} dV_P(\vec{r})\partial_t\rho^E(\vec{r},t) + \int_{\partial K'(t)} dA_P(\vec{r})\hat{n}_P\cdot\vec{v}_P^E(\vec{r},t),$$

where $\hat{n}_P$ is the unit outward normal to $\partial K'(t)$. An application of Gauss's theorem to the surface integral gives

$$\frac{d}{dt}M^E\left[K'(t)\right] = \int_{K'(t)} dV_P(\vec{r})\left[\partial_t\rho + \vec{\partial}\cdot(\rho\vec{v})\right]^E(\vec{r},t). \tag{13.1.10}$$

Comparing (13.1.9) and (13.1.10), we see that if $K' = K'(t)$ then

$$\int_{K'} dV_P(\vec{r})\left[\partial_t\rho + \vec{\partial}\cdot(\rho\vec{v})\right]^E(\vec{r},t) = 0. \tag{13.1.11}$$

Now let K' be *any* open subset of $K(t)$, with piecewise smooth boundary $\partial K'$. Let H' be the set of labels belonging to particles which are in K' at time t. Then for this set of particles, $K'(t) = K'$, so (13.1.11) holds. In other words, (13.1.11) holds for every open subset K' of $K(t)$, as long as $\partial K'$ is piecewise smooth. Therefore, by the vanishing integral theorem, $[\partial_t \rho + \vec{\partial} \cdot (\rho \vec{v})]^E(\vec{r}, t) = 0$ for all $t \in R$ and all $\vec{r} \in K(t)$. Therefore

$$\partial_t \rho + \vec{\partial} \cdot (\rho \vec{v}) = 0. \tag{13.1.12}$$

Equation (13.1.12) is the Eulerian form of the law of mass conservation. It is called the "continuity equation".

Since $\vec{\partial} \cdot (\rho \vec{v}) = (\vec{\partial} \rho) \cdot \vec{v} + \rho(\vec{\partial} \cdot \vec{v})$ and $D_t \rho = \partial_t \rho + \vec{v} \cdot (\vec{\partial} \rho)$, the continuity equation can also be written

$$D_t \rho + \rho(\vec{\partial} \cdot \vec{v}) = 0. \tag{13.1.13}$$

By the chain rule for ordinary differentiation, $D_t \ln \rho = D_t \rho / \rho$, so

$$D_t(\ln \rho) + \vec{\partial} \cdot \vec{v} = 0. \tag{13.1.14}$$

A motion of a continuum is called incompressible if it makes $\rho^L(\vec{x}, t)$ independent of t at each particle $\vec{x}$. The density is constant at each particle, but can vary from particle to particle. A material is incompressible if it is incapable of any motion except incompressible motion. A motion is incompressible iff $D_t \rho^L(\vec{x}, t) = 0$ for all $\vec{x} \in H$ and $t \in \mathcal{R}$, i.e. iff $D_t \rho = 0$. Comparing (13.1.13), the motion is incompressible iff

$$\vec{\partial} \cdot \vec{v} = 0. \tag{13.1.15}$$

An extremely useful consequence of the continuity equation follows from

Lemma 13.1.28 Suppose U and V are Euclidean spaces, D is an open subset of U, and $\vec{f} : D \to U$ and $\vec{g} : D \to V$ are differentiable. Then

$$\vec{\partial} \cdot (\vec{f} \vec{g}) = (\vec{\partial} \cdot \vec{f}) \vec{g} + \vec{f} \cdot (\vec{\partial} \vec{g}). \tag{13.1.16}$$

Proof:

Take components relative to orthonormal bases in U and V. Then (13.1.16) is equivalent to $\partial_i(f_i g_j) = (\partial_i f_i)g_j + f_i(\partial_i g_j)$. But this is the elementary rule for the partial derivative of a product.

Remark 13.1.60 Suppose $\vec{f}$ is any physical quantity taking values in a Euclidean space V. Suppose $K'(t)$ is any open set moving with the material in a continuum (i.e., always consisting of the same particles). Then

$$\frac{d}{dt}\int_{K'(t)} dV_P(\vec{r})(\rho\vec{f})^E(\vec{r},t) = \int_{K'(t)} dV_P(\vec{r})(\rho D_t\vec{f})^E(\vec{r},t). \tag{13.1.17}$$

Proof:

By (11.6.1), we have

$$\begin{aligned}\frac{d}{dt}\int_{K'(t)} dV_P(\vec{r})\left(\rho\vec{f}\right)^E(\vec{r},t) &= \int_{K'(t)} dV_p(\vec{r})\partial_t(\rho\vec{f})^E(\vec{r},t)\\ &+\int_{\partial K'(t)} dA_P(\vec{r})\hat{n}_P\cdot(\vec{v}\rho\vec{f})^E(\vec{r},t)\end{aligned}$$

where $\hat{n}_P$ is the unit outward normal to $\partial K'(t)$ and dA_P is an element of area. Applying Gauss's theorem to the surface integral gives

$$\begin{aligned}&\frac{d}{dt}\int_{K'(t)} dV_P(\vec{r})(\rho\vec{f})^E(\vec{r},t)\\ &\quad= \int_{K'(t)} dV_P(\vec{r})\left[\partial_t(\rho\vec{f}) + \vec{\partial}\cdot\left(\rho\vec{v}\vec{f}\right)\right]^E(\vec{r},t).\end{aligned}$$

The integrand is $(\partial_t\rho)\vec{f} + \rho(\partial_t\vec{f}) + \left[\vec{\partial}\cdot(\rho\vec{v})\right]\vec{f} + \rho\vec{v}\cdot(\vec{\partial}\vec{f})$ because of lemma (13.1.28). This integrand can therefore be written $[\partial_t\rho + \vec{\partial}\cdot(\rho\vec{v})]\vec{f} + \rho[\partial_t\vec{f} + \vec{v}\cdot(\vec{\partial}\vec{f})]$. From (12.4.21-12.4.26) and (13.1.12) this is $0\vec{f} + \rho D_t\vec{f}$, which proves (13.1.17).

An alternative proof of (13.1.17) may be physically more enlightening. We use (13.0.1) to change the variable of integration in (13.1.17) to $\vec{x}$ instead of $\vec{r}$. Then

$$\int_{K'(t)} dV_P(\vec{r})\rho^E(\vec{r},t)\vec{f}^E(\vec{r},t)$$

$$= \int_{H'} dV_L(\vec{x}) \left|\det \vec{D}\, \vec{\mathbf{r}}^{\,L}(\vec{x},t)\right| \rho^L(\vec{x},t)\vec{f}^L(\vec{x},t)$$

$$\overset{\text{(use 13.1.3)}}{=} \int_{H'} dV_L(\vec{x})\tilde{\rho}^L(\vec{x})\vec{f}^L(\vec{x},t).$$

Since H' and $\tilde{\rho}^L(\vec{x})$ do not vary with t, $(dV_L(\vec{x})\tilde{\rho}^L(\vec{x}) = dm)$

$$\frac{d}{dt}\int_{K'}(t)dV_P(\vec{r})(\rho\vec{f})^E(\vec{r},t) = \int_{H'} dV_L(\vec{x})\tilde{\rho}^L(\vec{x})D_t\vec{f}^L(\vec{x},t)$$

$$= \int_{H'} dV_L(\vec{x}) \left|\det \vec{D}\, \vec{\mathbf{r}}^{\,L}(\vec{x},t)\right| \rho^L(\vec{x},t)(D_t\vec{f})^L(\vec{x},t)$$

$$= \int_{K(t)} dV_P(\vec{r})\rho^E(\vec{r},t)(D_t\vec{f})^E(\vec{r},t).$$

QED

13.2 Conservation of momentum

13.2.1 Eulerian form

We use the notation introduced on page 163. We denote by $\vec{P}^E[K'(t)]$ the total momentum at time t of the material in subset $K'(t)$ of P at time t. This material consists of the particles whose labels are in H' and whose positions at time t are in $K'(t)$. The law of conservation of momentum is

$$\frac{d}{dt}\vec{P}^E\left[K'(t)\right] = \vec{\mathcal{F}}^E\left[K'(t)\right], \tag{13.2.1}$$

where $\vec{\mathcal{F}}^E[K'(t)]$ is the total force at time t on the material in $K'(t)$. As it stands, (13.2.1) is essentially a definition of $\vec{\mathcal{F}}^E[K'(t)]$, and our task is to calculate this force and then to extract from (13.2.1) a local form of the law of conservation of momentum in the same way that (13.1.12) was extracted from (13.1.9). To do this, we also need an expression for $\vec{P}^E[K'(t)]$, but this is simple. The mass in $dV_P(\vec{r})$ is $dV_P(\vec{r})\rho^E(\vec{r},t)$ at time t, so its momentum is $dV_P(\vec{r})\rho^E(\vec{r},t)\vec{v}^E(\vec{r},t)$. Summing over all the volume elements $dV_P(\vec{r})$ in $K'(t)$ gives

$$\vec{P}^E\left[K'(t)\right] = \int_{K'(t)} dV_P(\vec{r})(\rho\vec{v})^E(\vec{r},t) \tag{13.2.2}$$

If we apply to this equation both (13.1.17) and (12.4.21-12.4.26) we obtain the identity

$$\frac{d}{dt}\vec{P}^E\left[K'(t)\right] = \int_{K'(t)} dV_P(\vec{r})(\rho\vec{a})^E(\vec{r},t) \tag{13.2.3}$$

where $\vec{a} = D_t\vec{v} = \partial_t\vec{v} + \vec{v}\cdot\vec{\partial}\vec{v} =$ particle acceleration.

Finding $\vec{\mathcal{F}}^E[K'(t)]$ is less simple. The expression for it is one of the major eighteenth century advances in mechanics. If Ψ is the gravitation potential, so that $\vec{g} = -\vec{\partial}\Psi$ is the local acceleration of gravity, then the gravitational force on the matter in $dV_P(\vec{r})$ at time t is $dV_P(\vec{r})\rho^E(\vec{r},t)\vec{g}^E(\vec{r},t)$. If $\vec{E}$ and $\vec{B}$ are the electric and magnetic fields, and ρ_Q and $\vec{J}$ are the electric charge density and electric current density then the electromagnetic force on the matter in $dV_P(\vec{r})$ at time t is $dV_P(\vec{r})(\rho_Q\vec{E} + \vec{J}\times\vec{B})^E(\vec{r},t)$. The total gravitational plus electromagnetic force on the matter in $dV_P(\vec{r})$ is

$$d\vec{\mathcal{F}}_B(\vec{r}) = dV_P(\vec{r})\vec{f}^E(\vec{r},t) \tag{13.2.4}$$

where

$$\vec{f} = \rho\vec{g} + \rho_Q\vec{E} + \vec{J}\times\vec{B}. \tag{13.2.5}$$

The total gravitational plus electromagnetic force on the matter in $K'(t)$ is

$$\vec{\mathcal{F}}_B^E\left[K'(t)\right] = \int_{K'(t)} dV_P(\vec{r})\vec{f}^E(\vec{r},t). \tag{13.2.6}$$

This is called the "body force" on $K'(t)$, and $\vec{f}$ is the body force density per unit of physical volume, the physical density of body force.

If we accept $\vec{\mathcal{F}}_B^E$ as a good model for $\vec{\mathcal{F}}^E$, the only force acting on a cubic centimeter of ocean or rock is $\rho\vec{g}$ (assuming $\vec{E} = \vec{B} = \vec{0}$), and yet neither is observed to fall at $981 cm^2/sec$. Something important is still missing in our model of $\vec{\mathcal{F}}^E$.

The physical origin of our difficulty is clear. The forces in (13.2.5) are calculated from the average distribution of molecules and charge carriers, as in Figure 12.1. In addition to these long range average forces, we expect that the molecules just outside $\partial K'(t)$, will exert forces on the molecules just inside $\partial K'(t)$, and these contribute to $\vec{\mathcal{F}}^E[K'(t)]$. Also, in gases and, to some extent, in liquids, individual molecules will cross $\partial K'(t)$, and the entering molecules may have, on average, different momenta from the exiting molecules.

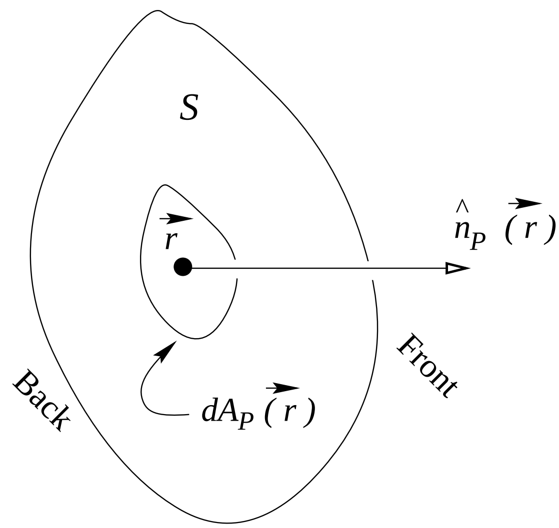


Figure 13.1:

This will result in a net contribution to the rate of change of the total momentum in $K'(t)$. Therefore, it is part of $d\vec{P}[K'(t)]/dt$ and hence, by (13.2.1), part of $\vec{\mathcal{F}}^E[K'(t)]$. Both the intermolecular force and the momentum transfer by molecular motion can be modelled as follows: Fix time t and choose a fixed $\vec{r} \in K(t)$. Choose a very small nearby plane surface S in $K(t)$ such that S passes through $\vec{r}$. Choose an even smaller surface $dA_P(\vec{r})$ which passes through $\vec{r}$ and lies in S. Arbitrarily designate one side of S as its "front", and let $\hat{n}_P(\vec{r})$ be the unit normal to S extending in front of S. Assume that S is so small that the molecular statistics do not change appreciably across S, but that S is considerably larger in diameter than the intermolecular distance or the mean free path. Then the total force exerted by the molecules just in front of $dA_P(\vec{r})$ on the molecules just behind $dA_P(\vec{r})$ will be proportional[1] to the area of $dA_P(\vec{r})$. We write this area as $dA_P(\vec{r})$. The proportionality constant is a vector, which we write $\vec{S}_{force}$. It depends on $\vec{r}$ and t, and it may also depend on the orientation of the surface S, i.e. on the unit normal $\hat{n}_P(\vec{r})$. If the material is a gas or liquid, there will also be a net transfer of momentum from front to back across $dA_P(\vec{r})$ because molecules cross $dA_P(\vec{r})$ and collide just after crossing, and the population of molecules one mean free path in front of $dA_P(\vec{r})$ may be

[1]if the linear dimension of dA is $>>$ distance between molecules

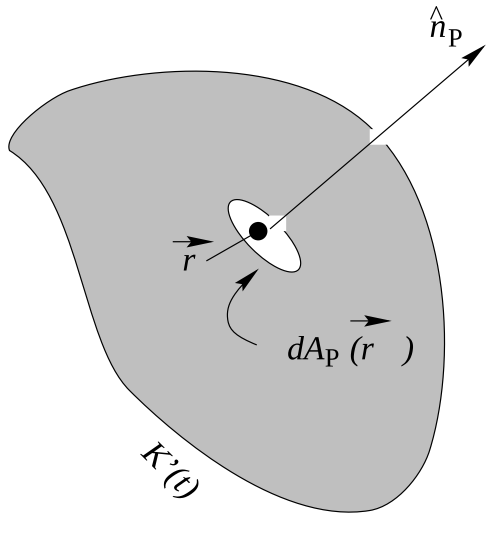


Figure 13.2:

statistically different from the population one mean free path behind $dA_P(\vec{r})$. The net rate of momentum transfer from just in front of $dA(\vec{r})$ to just behind $dA_P(\vec{r})$ will produce a time rate of change of momentum of the material just behind $dA_P(\vec{r})$; that is, it will exert a net force on that material. This force will also be proportional to $dA_P(\vec{r})$, and the proportionality constant is another vector, which we write $\vec{S}_{mfp}$. This vector also depends on $\vec{r}$,t and $\hat{n}_P(\vec{r})$. The sum $\vec{S}(\vec{r}, t, \hat{n}_P) = \vec{S}_{force}(\vec{r}, t, \hat{n}_P) + \vec{S}_{mfp}(\vec{r}, t, \hat{n}_P)$ is called the stress on the surface $(S, \hat{n}_P)$. The total force exerted by the material just in front of $dA_P(\vec{r})$ on the material just behind $dA_P(\vec{r})$ is

$$d\vec{\mathcal{F}}_S(\vec{r}) = dA_P(\vec{r})\vec{S}(\vec{r}, t, \hat{n}_P). \tag{13.2.7}$$

This is called the surface force on $dA_P(\vec{r})$. Summing (13.2.7) over all the elements of area $dA_P(\vec{r})$ on $\partial K'(t)$ gives

$$\vec{\mathcal{F}}_S^E\left[K'(t)\right] = \int_{\partial K'(t)} dA_P(\vec{r})\vec{S}\left[\vec{r}, t, \hat{n}_P(\vec{r}, t)\right] \tag{13.2.8}$$

where $\hat{n}_P(\vec{r}, t)$ is the unit outward normal to $\partial K'(t)$ at $\vec{r} \in \partial K'(t)$. The expression (13.2.8) is called the surface force on $K'(t)$. It is the total force exerted *on* the material just inside $\partial K'(t)$ by the material just outside $\partial K'(t)$. The total force on $K'(t)$ is the sum of the body and surface forces, $\vec{\mathcal{F}}^E[K'(t)] = \vec{\mathcal{F}}^E_B[K'(t)] + \vec{\mathcal{F}}^E_S[K'(t)]$, so

$$\vec{\mathcal{F}}^E\left[K'(t)\right] = \int_{K'(t)} dV_P(\vec{r})\vec{f}^E(\vec{r},t) + \int_{\partial K'(t)} dA_P(\vec{r})\vec{S}\left[\vec{r},t,\hat{n}_P(\vec{r},t)\right]. \tag{13.2.9}$$

Combining the physical law (13.2.1) with the mathematical expressions (13.2.3) and (13.2.9) gives

$$\int_{K'} dV_P(\vec{r})(\rho\vec{a} - \vec{f})^E(\vec{r},t) = \int_{\partial K'} dA_P(\vec{r})\vec{S}\left[\vec{r},t,\hat{n}_P(\vec{r})\right]. \tag{13.2.10}$$

where $K' = K'(t)$ and $\hat{n}_P(\vec{r})$ is the unit outward normal to $\partial K'$ at $\vec{r} \in \partial K'$.

As with (13.1.11), (13.2.10) must hold for every open subset K' of $K(t)$, as long as $\partial K'$ is piecewise smooth. This conclusion is forced on us if we accept the mathematical models (13.2.9) for $\vec{\mathcal{F}}^E[K'(t)]$ and (13.2.2) for $P^E[K'(t)]$.

To convert (13.2.10) to a local equation, valid for all $\vec{r} \in K(t)$ at all times t, (i.e., to "remove the integral signs") we would like to invoke the vanishing integral theorem, as we did in going from (13.1.11) to (13.1.12). The surface integral in (13.2.10) prevents this.

Even worse, (13.2.10) makes our model look mathematically self-contradictory, or internally inconsistent. Suppose that K' shrinks to a point while preserving its shape. Let λ be a typical linear dimension of K'. Then the left side of (13.2.10) seems to go to zero like λ^3, while the right side goes to zero like λ^2. How can they be equal for all $\lambda > 0$?

Cauchy resolved the apparent contradiction in 1827. He argued that the right side of (13.2.10) can be expanded in a power series in λ, and the validity of (13.2.10) for all λ shows that the first term in this power series, the λ^2 term, must vanish. In modern language, Cauchy showed that this can happened iff at every instant t, at every $\vec{r} \in K(t)$, there is a unique tensor $\overleftrightarrow{S}^E(\vec{r},t) \in P \otimes P$ such that for each unit vector $\hat{n} \in P$,

$$\vec{S}(\vec{r},t,\hat{n}) = \hat{n}\cdot \overleftrightarrow{S}^E(\vec{r},t). \tag{13.2.11}$$

If equation (13.2.11) is true, we can substitute it in the surface integral in (13.2.10) and use Gauss's theorem to write

$$\int_{\partial K'} dA_P(\vec{r})\hat{n}_P(\vec{r})\cdot \overleftrightarrow{S}^E(\vec{r},t) = \int_{K'} dV_P(\vec{r})\vec{\partial}\cdot \overleftrightarrow{S}^E(\vec{r},t).$$

Then (13.2.10) becomes

$$\int_{K'} dV(\vec{r}) \left[\rho\vec{a} - \vec{f} - \vec{\partial}\cdot \overset{\leftrightarrow}{S}\right]^E (\vec{r}, t) = 0.$$

Since this is true for all open subsets K' of $K(t)$ with piecewise smooth boundaries $\partial K'$, the vanishing integral theorem implies that the integrand vanishes for all $\vec{r} \in K(t)$ if, as we shall assume, it is continuous. Therefore

$$\rho\vec{a} = \vec{\partial}\cdot \overset{\leftrightarrow}{S} + \vec{f}. \tag{13.2.12}$$

This is the Eulerian form of the momentum equation. The tensor $\overset{\leftrightarrow}{S}^E (\vec{r}, t)$ called the Cauchy stress tensor at $(\vec{r}, t)$. The physical quantity $\overset{\leftrightarrow}{S}$ is also called the Cauchy stress tensor.

The argument which led Cauchy from (13.2.10) to (13.2.11) is fundamental to continuum mechanics, so we examine it in detail. In (13.2.10), t appears only as a parameter, so we will ignore it. Then (13.2.10) implies (13.2.11) because of

Theorem 13.2.28 (Cauchy's Theorem) Suppose U and V are Euclidean spaces, K is an open subset of U, N_U = set of all unit vectors $\hat{n} \in U$, $\vec{f} : K \to V$ is continuous, and $\vec{S} : K \times N_U \to V$ is continuous. Suppose that for any open subset K' of K whose boundary $\partial K'$ is piecewise smooth, we have

$$\int_{K'} dV_U(\vec{r})\vec{f}(\vec{r}) = \int_{\partial K'} dA_U(\vec{r})\vec{S}\left[\vec{r}, \hat{n}_U(\vec{r})\right], \tag{13.2.13}$$

where $\hat{n}_U(\vec{r})$ is the unit outward normal to $\partial K'$ at $\vec{r} \in \partial K'$. Then for each $\vec{r} \in K$ there is a unique $\overset{\leftrightarrow}{S}(\vec{r}) \in U \otimes V$ such that for all $\hat{n} \in N_U$,

$$\vec{S}(\vec{r}, \hat{n}) = \hat{n}\cdot \overset{\leftrightarrow}{S}(\vec{r}). \tag{13.2.14}$$

[This theorem is true for any value of $\dim U \geq 2$. We give the proof only for $\dim U = 3$, the case of interest to us. For other values of $\dim U$, a proof can be given in exactly the same way except that $\vec{u}_1 \times \vec{u}_2$ must be replaced by the vector in Exercise 7.]

Proof:

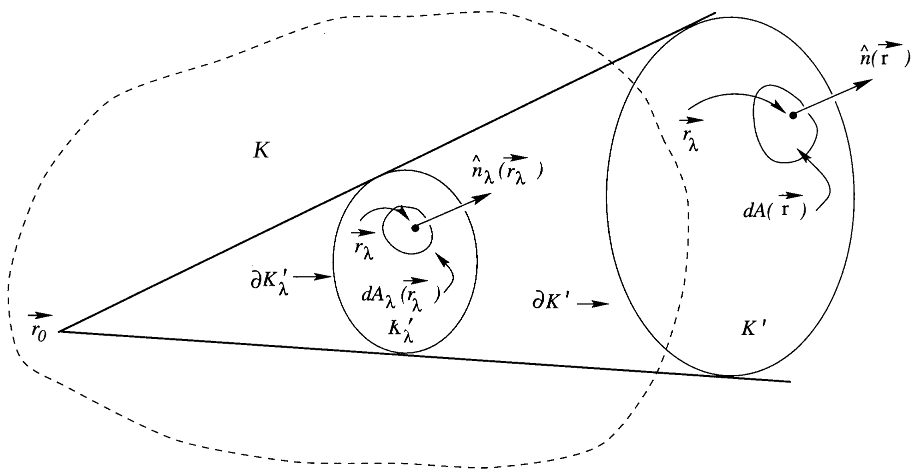


Figure 13.3:

Two lemmas are required.

Lemma 13.2.29 Suppose $\vec{f}$ and $\vec{S}$ satisfy the hypotheses of theorem 13.2.28. Let $\vec{r}_0$ be any point in K and let K' be *any* open bounded (i.e., there is a real M such that $\vec{r} \in K' \Rightarrow \|\vec{r}\| \leq M$) subset of U, with piecewise smooth boundary $\partial K'$. We don't need $K' \subseteq K$. Then

$$\int_{\partial K'} dA_U(\vec{r})\vec{S}\,[\vec{r}_0, \hat{n}_U(\vec{r})] = \vec{0}_V \tag{13.2.15}$$

if $\hat{n}_U(\vec{r})$ is the unit outward normal to $\partial K'$ at $\vec{r} \in \partial K'$ and $dA_U(\vec{r})$ is the element of area on $\partial K'$.

Proof of Lemma 13.2.29: For any real λ in $0 < \lambda < 1$, define $\vec{\mathbf{r}}_\lambda : U \to U$ by $\vec{\mathbf{r}}_\lambda(\vec{r}) = \vec{r}_0 + \lambda(\vec{r} - \vec{r}_0)$ for all $\vec{r} \in U$. Since $\vec{r}_\lambda(\vec{r}) - \vec{r}_0 = \lambda(\vec{r} - \vec{r}_0)$, $\vec{\mathbf{r}}_\lambda$ shrinks U uniformly toward $\vec{r}_0$ by the factor λ. The diagram above is for $\lambda \approx 1/2$. Define $K'_\lambda = \vec{\mathbf{r}}_\lambda(K')$ so $\partial K'_\lambda = \vec{\mathbf{r}}_\lambda(\partial K')$. Choose $\vec{r} \in \partial K'$ and let $\vec{r}_\lambda = \vec{\mathbf{r}}_\lambda(\vec{r})$. Let $dA(\vec{r})$ denote a small nearly plane patch of surface in $\partial K'$, with $\vec{r} \in dA(\vec{r})$, and use $dA(\vec{r})$ both as the name of this set and as the numerical value of its area. Let the set $dA_\lambda(\vec{r}_\lambda)$ be defined as $\vec{\mathbf{r}}_\lambda[dA(\vec{r})]$, and denote its area also by $dA_\lambda(\vec{r}_\lambda)$. Then by geometric similarity

$$dA_\lambda(\vec{r}_\lambda) = \lambda^2 dA(\vec{r}). \tag{13.2.16}$$

Let $\hat{n}(\vec{r})$ be the *unit* outward normal to $\partial K'$ at $\vec{r}$, and let $\hat{n}_\lambda(\vec{r}_\lambda)$ be the *unit* outward normal to $\partial K'_\lambda$ at $\vec{r}_\lambda$. By similarity, $\hat{n}(\vec{r})$ and $\hat{n}_\lambda(\vec{r}_\lambda)$ point in the same direction. Being unit vectors, they are equal:

$$\hat{n}_\lambda(\vec{r}_\lambda) = \hat{n}(\vec{r}). \tag{13.2.17}$$

Since $\vec{r}_0$ is fixed, it follows that

$$\int_{\partial K'_\lambda} dA_\lambda(\vec{r}_\lambda)\vec{S}\left[\vec{r}_0, \hat{n}_\lambda(\vec{r}_\lambda)\right] = \lambda^2 \int_{\partial K'} dA(\vec{r})\vec{S}\left[\vec{r}_0, \hat{n}(\vec{r})\right]. \tag{13.2.18}$$

If λ is small enough, $K'_\lambda \subseteq K$. Then, by hypothesis, we have (13.2.13) with K' and $\partial K'$ replaced by K'_λ and $\partial K'_\lambda$. Therefore

$$\begin{aligned}\int_{K'_\lambda} dV(\vec{r})\vec{f}(\vec{r}) &= \int_{K'_\lambda} dA_\lambda(\vec{r}_\lambda)\left\{\vec{S}\left[\vec{r}_\lambda, \hat{n}_\lambda(\vec{r}_\lambda)\right] - \vec{S}\left[\vec{r}_0, \hat{n}_\lambda(\vec{r}_\lambda)\right]\right\} \\ &\quad + \int_{\partial K'_\lambda} dA_\lambda(\vec{r}_\lambda)\vec{S}\left[\vec{r}_0, \hat{n}_\lambda(\vec{r}_\lambda)\right].\end{aligned}$$

From (13.2.18) it follows that

$$\begin{aligned}\int_{\partial K'} dA(\vec{r})\vec{S}\left[\vec{r}_0, \hat{n}(\vec{r})\right] &= \frac{1}{\lambda^2}\int_{K'_\lambda} dV(\vec{r})\vec{f}(\vec{r}) \\ &\quad + \frac{1}{\lambda^2}\int_{\partial K'_\lambda} dA_\lambda(\vec{r}_\lambda) \\ &\quad \left\{\vec{S}\left[\vec{r}_0, \hat{n}_\lambda(\vec{r}_\lambda)\right] - \vec{S}\left[\vec{r}_\lambda, \hat{n}_\lambda(\vec{r}_\lambda)\right]\right\}.\end{aligned} \tag{13.2.19}$$

Let $m_{\vec{S}}(\lambda)$ = maximum value of $\|\vec{S}(\vec{r}_0, \hat{n}) - \vec{S}(\vec{r}, \hat{n})\|$ for all $\vec{r} \in \partial K'_\lambda$ and all $\hat{n} \in N_U$.

Let $m_{\vec{f}}(\lambda)$ = maximum value of $|\vec{f}(\vec{r})|$ for all $\vec{r} \in K'_\lambda$.

Let $|\partial K'_\lambda|$ = area of $\partial K'_\lambda$, $|\partial K'|$ = area of $\partial K'$.

Let $|K'_\lambda|$ = volume of K'_λ, $|K'|$ = volume of K'.

Then $|\partial K'_\lambda| = \lambda^2 |\partial K'|$ and $|K'_\lambda| = \lambda^3 |K'|$, so (10.2.3) and (13.2.19) imply

$$\left\|\int_{\partial K'} dA(\vec{r})\vec{S}\left[\vec{r}_0, \hat{n}(\vec{r})\right]\right\| \le \lambda |K'| m_{\vec{f}}(\lambda) + |\partial K'| m_{\vec{S}}(\lambda). \tag{13.2.20}$$

As $\lambda \to 0$, $m_{\vec{f}}(\lambda)$ remains bounded (in fact $\to \|f(\vec{r}_0)\|$) and $m_{\vec{S}}(\lambda) \to 0$ because $\vec{S} : K \times N_U \to V$ is continuous. Therefore, as $\lambda \to 0$, the right side of (13.2.20) $\to 0$. Inequality (13.2.20) is true for all sufficiently small $\lambda > 0$, and the left side is non-negative and independent of λ. Therefore the left side must be 0. This proves (13.2.15) and hence proves lemma 13.2.29.

We also need

Lemma 13.2.30 Suppose $\vec{S} : N_U \to V$. Suppose that for any open set K' with piecewise smooth boundary $\partial K'$, $\vec{S}$ satisfies

$$\int_{\partial K'} dA(\vec{r})\vec{S}\left[\hat{n}(\vec{r})\right] = \vec{0}_V \tag{13.2.21}$$

where $\hat{n}(\vec{r})$ is the unit outward normal to $\partial K'$ as $\vec{r} \in \partial K'$. Suppose $F : U \to V$ is defined as follows:

$$F(\vec{0}_U) = \vec{0}_V \text{ and if } \vec{u} \neq \vec{0}_U, \quad F(\vec{u}) = \|\vec{u}\|\vec{S}\left(\frac{\vec{u}}{\|\vec{u}\|}\right). \tag{13.2.22}$$

Then F is linear.

Proof of Lemma 13.2.30: a) $F(-\vec{u}) = -F(\vec{u})$ for all $\vec{u} \in U$. To prove this, it suffices to prove

$$\vec{S}(-\hat{n}) = -\vec{S}(\hat{n}) \quad \text{for all } \hat{n} \in N_U. \tag{13.2.23}$$

Let K' be the flat rectangular box shown at upper right. For this box, (13.2.21) gives

$$L^2\vec{S}(\hat{n}) + L^2\vec{S}(-\hat{n}) + \varepsilon L\left[\vec{S}(\hat{n}_1) + \vec{S}(-\hat{n}_1) + \vec{S}(\hat{n}_2) + \vec{S}(-\hat{n}_2)\right] = \vec{0}_V.$$

Hold L fixed and let $\varepsilon \to 0$. Then divide by L^2 and (13.2.23) is the result.

b) If $c \in \mathcal{R}$ and $\vec{u} \in U$, $F(c\vec{u}) = cF(\vec{u})$.

i) If $c = 0$ or $\vec{u} = \vec{0}_U$, this is obvious from $F(\vec{0}_U) = \vec{0}_V$.

ii) If $c > 0$ and $\vec{u} \neq \vec{0}_U$, $F(c\vec{u}) = \|c\vec{u}\|\|\vec{S}(c\vec{u}/\|c\vec{u}\|) = c\|\vec{u}\|\vec{S}(c\vec{u}/c\|\vec{u}\|) = c\|\vec{u}\|\vec{S}(\vec{u}/\|\vec{u}\|) = cF(\vec{u})$.

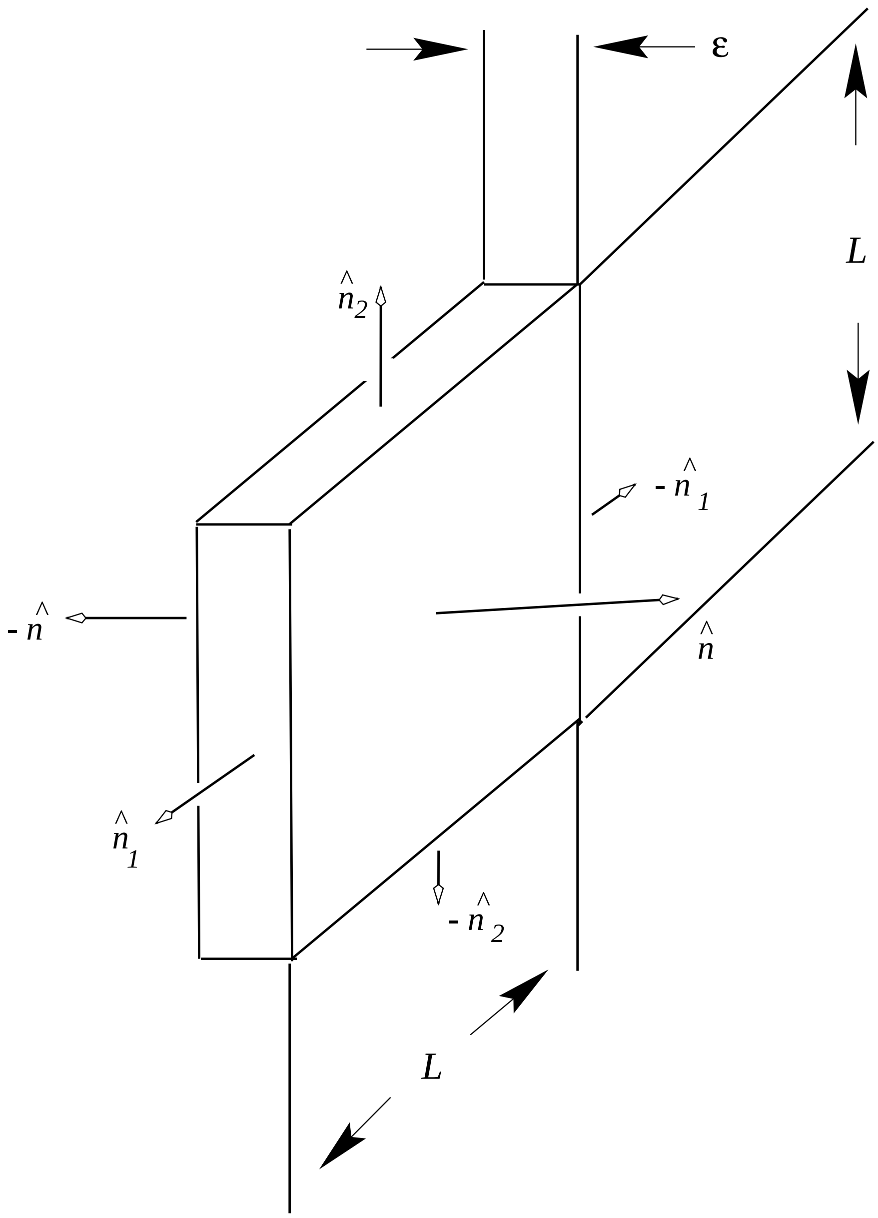


Figure 13.4:

iii) If $c < 0$, $F(c\vec{u}) = -F(-c\vec{u})$ by a) above. But $-c > 0$ so $F(-c\vec{u}) = (-c)F(\vec{u})$ by ii). Then $F(c\,\vec{u}) = -(-c)F(\vec{u}) = cF(\vec{u})$.

c) $F(\vec{u}_1 + \vec{u}_2) = F(\vec{u}_1) + F(\vec{u}_2)$ for all $\vec{u}_1$, $\vec{u}_2 \in U$.

i) If $\vec{u}_1 = \vec{0}_U$, $F(\vec{u}_1 + \vec{u}_2) = F(\vec{u}_2) = \vec{0}_V + F(\vec{u}_2) = F(\vec{u}_1) + F(\vec{u}_2)$.

ii) If $\vec{u}_1 \neq \vec{0}_U$ and $\vec{u}_2 = c\vec{u}_1$ then $F(\vec{u}_1 + \vec{u}_2) = F[(1+c)\vec{u}_1] = (1+c)F(\vec{u}_1) = F(\vec{u}_1) + cF(\vec{u}_1) = F(\vec{u}_1) + F(c\vec{u}_1) = F(\vec{u}_1) + F(\vec{u}_2)$.

iii) If $\{\vec{u}_1, \vec{u}_2\}$ is linearly independent, let $\vec{u}_3 = -\vec{u}_1 - \vec{u}_2$. We want to prove $F(-\vec{u}_3) = F(\vec{u}_1) + F(\vec{u}_2)$, or $-F(\vec{u}_3) = F(\vec{u}_1) + F(\vec{u}_2)$, or

$$F(\vec{u}_1) + F(\vec{u}_2) + F(\vec{u}_3) = \vec{0}_V. \tag{13.2.24}$$

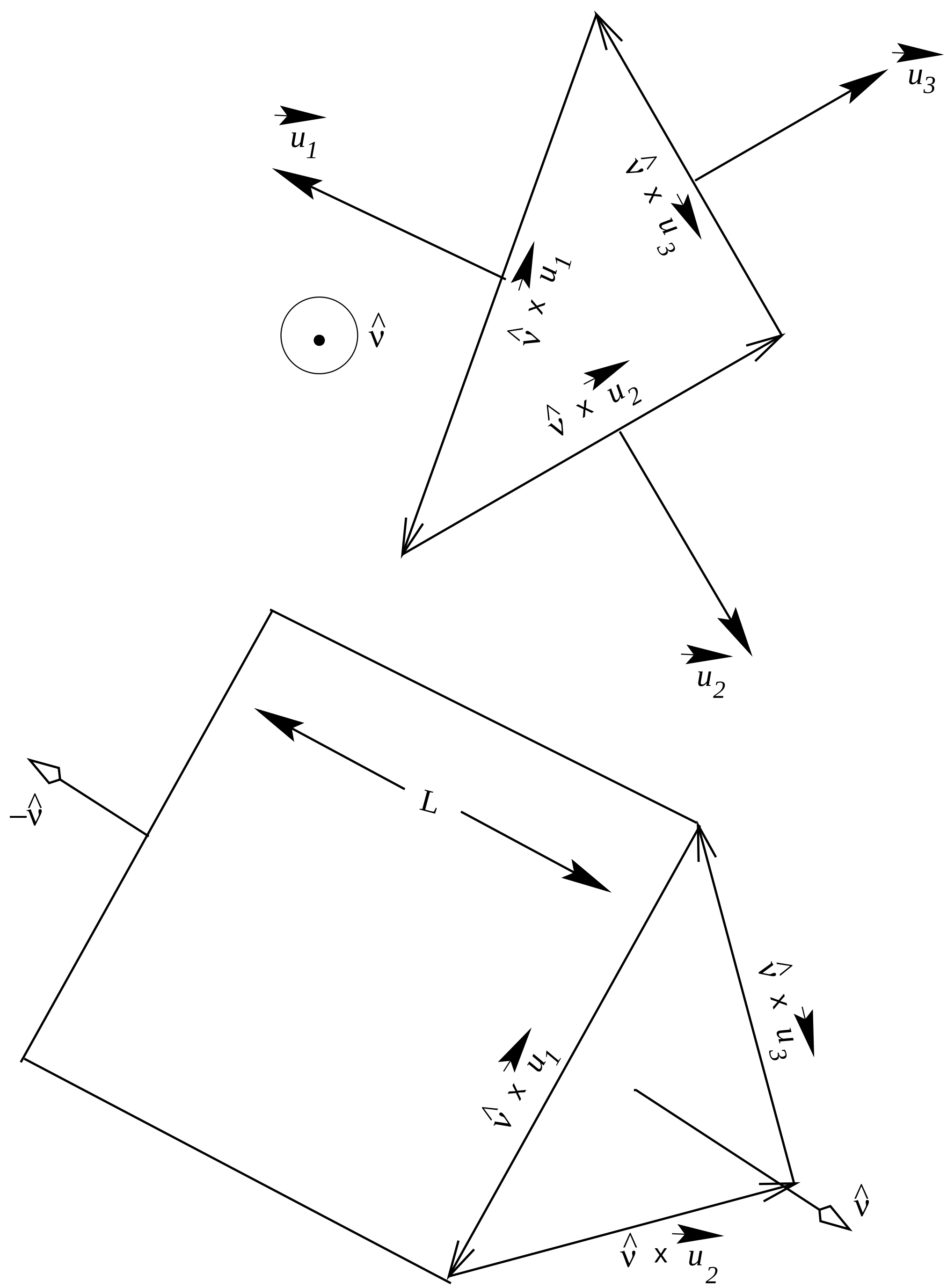
To prove (13.2.24) note that since $\vec{u}_1$, $\vec{u}_2$ are linearly independent, we can define the unit vector $\hat{\nu} = (\vec{u}_1 \times \vec{u}_2)/\|\vec{u}_1 \times \vec{u}_2\|$. We place the plane of this paper so that it contains $\vec{u}_1$ and $\vec{u}_2$, and $\hat{\nu}$ points out of the paper. The vectors $\vec{u}_1$, $\vec{u}_2$, $\vec{u}_3$ form the three sides of a nondegenerate triangle in the plane of the paper. $\hat{\nu} \times \vec{u}_i$ is obtained by rotating $\vec{u}_i$ $90°$ counterclockwise. If we rotate the triangle with sides $\vec{u}_1$, $\vec{u}_2$, $\vec{u}_3$ $90°$ counterclockwise, we obtain a triangle with sides $\hat{\nu} \times \vec{u}_1$, $\hat{\nu} \times \vec{u}_2$, $\hat{\nu} \times \vec{u}_3$. The length of side $\hat{\nu} \times \vec{u}_i$ is $\|\hat{\nu} \times \vec{u}_i\| = \|\vec{u}_i\|$, and $\vec{u}_i$ is perpendicular to that side and points out of the triangle. Let K' be the right cylinder whose base is the triangle with sides $\hat{\nu} \times \vec{u}_i$ and whose generators perpendicular to the base have length L. The base and top of the cylinder have area $A = \|\vec{u}_1 \times \vec{u}_2\|/2$ and their unit outward normals are $\vec{\nu}$ and $-\vec{\nu}$. The three rectangular faces of K' have areas $L\|\vec{u}_i\|$ and unit outward normals $\vec{u}_{(i)}/\|\vec{u}_{(i)}\|$. Applying (13.2.21) to this K' gives

$$A\vec{S}(\hat{\nu}) + A\vec{S}(-\hat{\nu}) + \sum_{i=1}^{3} L\|\vec{u}_i\|\vec{S}(\vec{u}_i/\|\vec{u}_i\|) = \vec{0}_V.$$

But $\vec{S}(\hat{\nu}) = -\vec{S}(-\hat{\nu})$ so dividing by L and using (13.2.22) gives (13.2.23).

Corollary 13.2.44 (to Lemma 13.2.30.) Under the hypotheses of lemma 13.2.30, there is a unique $\overleftrightarrow{S} \in U \otimes V$ such that for all $\hat{n} \in N_U$

$$\vec{S}(\hat{n}) = \hat{n} \cdot \overleftrightarrow{S} . \tag{13.2.25}$$



$\vec{u}_1$
$\vec{u}_3$
$\hat{\nu}$
$\hat{\nu} \times \vec{u}_1$
$\hat{\nu} \times \vec{u}_2$
$\hat{\nu} \times \vec{u}_3$
$\vec{u}_2$
L
$-\hat{\nu}$
$\hat{\nu} \times \vec{u}_1$
$\hat{\nu} \times \vec{u}_3$
$\hat{\nu}$
$\hat{\nu} \times \vec{u}_2$

Figure 13.5:

Proof:

Existence. Take $\overleftrightarrow{S}=\overleftrightarrow{F}$, the tensor in $U\otimes V$ corresponding to the $F\in\mathcal{L}(U\to V)$ defined by 13.2.22. Then $\hat{n}\cdot\overleftrightarrow{S}=\hat{n}\cdot\overleftrightarrow{F}=F(\hat{n})=\vec{S}(\hat{n})$.

Uniqueness. If $\hat{n}\cdot\overleftrightarrow{S}_1=\hat{n}\cdot\overleftrightarrow{S}_2$ for all $\hat{n}\in N_U$ then$(c\hat{n})\cdot\overleftrightarrow{S}_1=(c\hat{n})\cdot\overleftrightarrow{S}_2$ for all $\hat{n}\in N_U$ and $c\in\mathcal{R}$. But every $\vec{\mu}\in U$ is $c\hat{n}$ for some $c\in\mathcal{R}$ and $\hat{n}\in N_U$, so $\vec{u}\cdot\overleftrightarrow{S}_2$ for all $\vec{u}\in U$. Hence $\overleftrightarrow{S}_1=\overleftrightarrow{S}_2$.

Now we return to the proof of Cauchy's theorem (theorem 13.2.28). If $\vec{r}_0$ is any fixed point in K_1 then by lemma 13.2.29, the function $\vec{S}(\vec{r}_0,\cdot):N_U\to V$ satisfies the hypothesis of lemma 13.2.30. Therefore by corollary 13.2.44 there is a unique $\overleftrightarrow{S}(\vec{r}_0)\in U\otimes V$ such that for any unit vector $\hat{n}\in U$, $[\vec{S}(\vec{r}_0,\cdot)](\hat{n})=\hat{n}\cdot\overleftrightarrow{S}(\vec{r}_0)$. But $[\vec{S}(\vec{r}_0,\cdot)](\hat{n})=\vec{S}(\vec{r}_0,\hat{n})$ so we have (13.2.14).

Having completed the proof of Cauchy's theorem, we have proved that if (13.2.10) holds for all K' then Cauchy's stress tensor $\overleftrightarrow{S}^E(\vec{r},t)$ exists and satisfies (13.2.11). And as we have seen, (13.2.11) leads automatically to the Eulerian momentum equation (13.2.12). It is important to have a clear physical picture of what the existence of a Cauchy stress tensor $\overleftrightarrow{S}^E(\vec{r},t)$ means. If dA_P is any small nearly plane surface in the material in physical space P at time t, and if $\vec{r}\in dA_P$, choose one side of dA_P to be its front, and let $\hat{n}_P$ be the unit normal extending in front of dA_P. Then the force $d\vec{\mathcal{F}}_S$ exerted by the material just in front of dA *on* the material just behind dA is, at time t,

$$d\vec{\mathcal{F}}_S=dA_P\,\hat{n}_P\cdot\overleftrightarrow{S}^E(\vec{r},t). \tag{13.2.26}$$

The stress, or force per unit area, exerted on the material just behind dA by the material just in front is

$$\vec{S}(\vec{r},t,\hat{n}_P)=\hat{n}_P\cdot\overleftrightarrow{S}^E(\vec{r},t). \tag{13.2.27}$$

If K' is any open subset of $K(t)$, with piecewise smooth boundary $\partial K'$, then at time t the total surface force on the matter in K' is, according to (13.2.8),

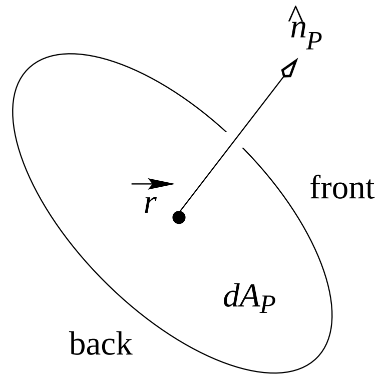


Figure 13.6:

and (13.2.27),

$$\vec{F}_S^E[K'] = \int_{\partial K'} dA_P(\mathbf{\vec{r}})\hat{n}_P(\vec{r})\cdot \overleftrightarrow{S}^E(\vec{r},t). \tag{13.2.28}$$

Using Gauss's theorem, we can write this as

$$\vec{\mathcal{F}}_S^E[K'] = \int_{K'} dV_P(\vec{r})\vec{\partial}\cdot \overleftrightarrow{S}^E(\vec{r},t). \tag{13.2.29}$$

The total force on the matter in K' is, according to (13.2.9) and (13.2.29),

$$\vec{\mathcal{F}}^E[K'] = \int_{K'} dV_P(\vec{r})\left(\vec{f} + \vec{\partial}\cdot \overleftrightarrow{S}\right)^E(\vec{r},t). \tag{13.2.30}$$

If the particles are all motionless, $\mathbf{\vec{r}}^L(\vec{x},t)$ is independent of time, so $\vec{v}^L(\vec{x},t) = D_t\mathbf{\vec{r}}^L(\vec{x},t) = \vec{0}$, and $\vec{a}^L(\vec{x},t) = D_t\vec{v}^L(\vec{x},t) = \vec{0}$, and the momentum equation (13.2.12) reduces to

$$\vec{0} = \vec{\partial}\cdot \overleftrightarrow{S} + \vec{f}, \tag{13.2.31}$$

the "static equilibrium equation." In this case, we see from (13.2.30) that the surface force on the matter in K' exactly balances the body force, for every open subset K' of $K(t)$.

If we have both $\vec{a} = \vec{0}$ and $\vec{f} = \vec{0}$, then

$$\vec{\partial}\cdot \overleftrightarrow{S} = \vec{0}. \tag{13.2.32}$$

The physical quantity $\overleftrightarrow{S}$ is called a "static stress field" if it satisfies (13.2.32). An example is an iron doughnut in orbit around the sun (to cancel gravity).

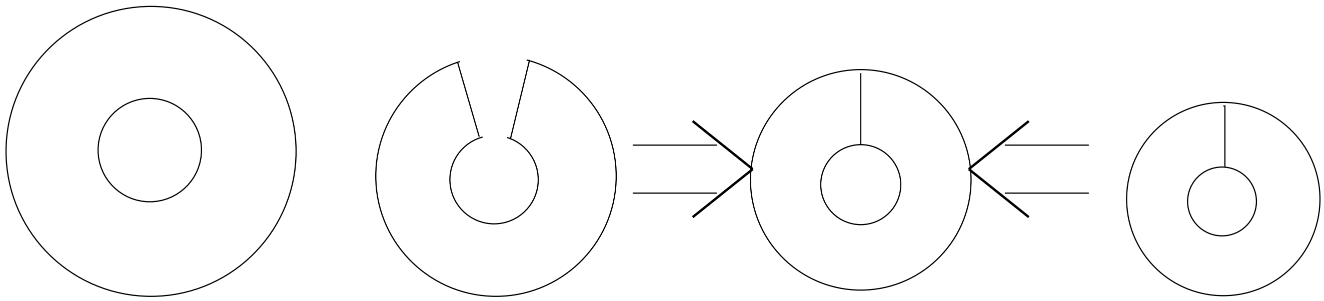


Figure 13.7:

Cut out a slice, close the gap by pressure, weld, and remove the pressure, as in figure 13.7 The resulting doughnut will contain a nonzero static stress field. The net force on every lump of iron in the doughnut vanishes because the body force vanishes and the surface force sums to $\vec{0}$.

13.2.2 Shear stress, pressure and the stress deviator

We want to discuss some of the elementary properties of stress and the Cauchy stress tensor. In figure (13.6), choosing which side of dA is the "front" is called "orienting dA", and the pair $(dA, \hat{n})$ is an oriented surface. Two oriented surfaces can be produced from dA, namely $(dA, \hat{n})$, and $(dA, -\hat{n})$ depending on which side of dA we call the front.

We want to study the state of stress at one instant t at one position $\vec{r} \in K(t)$, so we will omit $\vec{r}, t$ and also the superscript ^{E}. We will work only with Eulerian descriptions at the single point $\vec{r}$ and time t. Thus the stress $\vec{S}(\vec{r}, t, \hat{n})$ on small nearly plane oriented area $(dA, \hat{n})$ will be written $\vec{S}(\hat{n})$, and $\overset{\leftrightarrow}{S}^E(\vec{r}, t)$ will be written $\overset{\leftrightarrow}{S}$.

Definition 13.2.46

$$S_n(\hat{n}) := \hat{n} \cdot \vec{S}(\hat{n}) = \hat{n} \cdot \overset{\leftrightarrow}{S} \cdot \hat{n} \tag{13.2.33}$$

$$\vec{S}_S(\hat{n}) := \vec{S}(\hat{n}) - \hat{n} S_n(\hat{n}) = \hat{n} \cdot \overset{\leftrightarrow}{S} - \hat{n} \left[\hat{n} \cdot \overset{\leftrightarrow}{S} \cdot \hat{n} \right] \tag{13.2.34}$$

$$p_n(\hat{n}) := -S_n(\hat{n}) = -\hat{n} \cdot \overset{\leftrightarrow}{S} \cdot \hat{n}. \tag{13.2.35}$$

The vector $\hat{n} S_n(\hat{n})$ is called the normal stress acting on the oriented surface $(dA, \hat{n})$, the vector $\vec{S}_S(\hat{n})$ is called the tangential or shear stress acting on $(dA, \hat{n})$, and $p_n(\hat{n})$ is the pressure acting on $(dA, \hat{n})$.

Corollary 13.2.45

$$\hat{n} \cdot \vec{S}_S(\hat{n}) = 0 \tag{13.2.36}$$

$$\vec{S}(\hat{n}) = \hat{n} S_n(\hat{n}) + \vec{S}_S(\hat{n}) = -\hat{n} p_n(\hat{n}) + \vec{S}_S(\hat{n}). \tag{13.2.37}$$

The shear stress is always perpendicular to $\hat{n}$, parallel to dA. The normal stress acts oppositely to $\hat{n}$ if $p_n(\hat{n}) > 0$.

To visualize $\vec{S}$ as a function of $\hat{n}$, note that the set N_P of unit vectors $\hat{n} \in P$ is precisely $\partial B(\vec{0}, 1)$, the spherical surface of radius 1 centered on $\vec{0}$ in P. Thus $\vec{S}$ attaches to each $\hat{n}$ on the unit sphere N_P a vector $\vec{S}(\hat{n})$ whose radial part is $\hat{n} S_N(\hat{n})$ and whose tangential part is $\vec{S}_S(\hat{n})$. This suggests

Definition 13.2.47 The average value of $p_n(\hat{n})$ for all $\hat{n} \in N_P$ is called the average or mean pressure at $\vec{r}, t$. It is written $\langle p_n \rangle(\vec{r}, t)$. We do not write the $(\vec{r}, t)$ in this chapter. Thus

$$\langle p_n \rangle = \frac{1}{4\pi} \int_{N_P} dA(\hat{n}) p_n(\hat{n}) \tag{13.2.38}$$

where $dA(\hat{n})$ is the element of area on N_P.

Remark 13.2.61 $\langle p_n \rangle = -1/3 \; tr \overleftrightarrow{S}$

Proof:

Let $\hat{y}_1$, $\hat{y}_2$, $\hat{y}_3$ be an orthonormal basis for P, and take components relative to this basis. Then

$$p_n(\hat{n}) = -\hat{n} \cdot \overleftrightarrow{S} \cdot \hat{n} = -n_i S_{ij} n_j = -S_{ij} n_i n_j.$$

By the symmetry of N_P,

$$\int_{N_P} dA(\hat{n}) n_i n_j = 0 \text{ if } i \neq j \text{ and}$$

$$\int_{N_P} dA(\hat{n}) n_1^2 = \int_{N_P} dA(\hat{n}) n_2^2 = \int_{N_P} dA(\hat{n}) n_3^2 = \int_{N_P} dA(\hat{n}) \frac{(n_1^2 + n_2^2 + n_3^2)}{3} = \frac{4\pi}{3}.$$

Therefore,

$$\int_{N_P} dA(\hat{n}) n_i n_j = \frac{4\pi}{3}\delta_{ij} \text{ and}$$

$$\begin{aligned} \langle p_n \rangle &= -\frac{1}{4\,\pi} S_{ij} \int_{N_P} dA(\hat{n}) n_i n_j = -\frac{1}{3}\,\delta_{ij} S_{ij} \\ &= -\frac{1}{3} S_{ii} = -\frac{1}{3} tr \overleftrightarrow{S} . \end{aligned}$$

Only part of $\overleftrightarrow{S}$ produces the shear stress $\vec{S}_S(\hat{n})$. To see this we need a brief discussion of tensors in $U \otimes U$, where U is any Euclidean space.

Lemma 13.2.31 $tr \overleftrightarrow{I}_U = \dim U$.

Proof:

Let $\hat{y}_1, \cdots, \hat{y}_n$ be an orthonormal basis for U. Then $\overleftrightarrow{I}_U = \hat{y}_i \hat{y}_i$ and $tr \overleftrightarrow{I}_U = \hat{y}_i \cdot \hat{y}_i = n = \dim U$.

Lemma 13.2.32 If $\overleftrightarrow{S} \in U \otimes U$, $tr \overleftrightarrow{S} = \overleftrightarrow{I}_U \langle 2 \rangle \overleftrightarrow{S}$.

Proof:

With $\hat{y}_i$ as above, $tr \overleftrightarrow{S} = S_{ii} = \delta_{ij} S_{ij} = (\overleftrightarrow{I}_U)_{ij} S_{ij} = \overleftrightarrow{I}_U \langle 2 \rangle \overleftrightarrow{S}$.

Definition 13.2.48 $\overleftrightarrow{S} \in U \otimes U$ is "traceless", or a "deviator" if $tr \overleftrightarrow{S} = 0$.

Corollary 13.2.46 $\overleftrightarrow{S}$ is a deviator $\Leftrightarrow$ $\overleftrightarrow{S}$ is orthogronal to all isotropic tensors in $U \otimes U$.

Proof:

The isotropic tensors are the scalar multiples of $\overleftrightarrow{I}_U$. But $\overleftrightarrow{S}$ is orthogonal to $\overleftrightarrow{I}_U \Leftrightarrow \overleftrightarrow{S} \langle 2 \rangle \overleftrightarrow{I}_U = \hat{0}$. Now use lemma 13.2.32.

Remark 13.2.62 If $\overleftrightarrow{S} \in U \otimes U$, there is exactly one isotropic tensor $\sigma \overleftrightarrow{I}_U$ and one deviator $\overleftrightarrow{S}_D$ such that

$$\overleftrightarrow{S} = \sigma \overleftrightarrow{I}_U + \overleftrightarrow{S}_D . \qquad (13.2.39)$$

Proof:

Uniqueness. Since $tr\ \overleftrightarrow{S}_D = 0$, (13.2.39) implies $tr\ \overleftrightarrow{S} = \sigma\, tr\ \overleftrightarrow{I}_U = \sigma(\dim U)$. Thus σ is uniquely determined by $\overleftrightarrow{S}$. Then, from (13.2.39), the same is true of $\overleftrightarrow{S}_D$.

Existence. Choose $\sigma = (tr\ \overleftrightarrow{S})/\dim U$ and define $\overleftrightarrow{S}_D$ by (13.2.39). Then $tr\ \overleftrightarrow{S} = \sigma(tr\ \overleftrightarrow{I}_U) + tr\ \overleftrightarrow{S}_D = tr\ \overleftrightarrow{S} + tr\ \overleftrightarrow{S}_D$ so $tr\ \overleftrightarrow{S}_D = 0$.

Definition 13.2.49 In (13.2.39), $\sigma\ \overleftrightarrow{I}_U$ is the isotropic part of $\overleftrightarrow{S}$, and $\overleftrightarrow{S}_D$ is the deviatoric part of $\overleftrightarrow{S}$. If $U = P$ and $\overleftrightarrow{S}$ is a Cauchy stress tensor, $\overleftrightarrow{S}_D$ is the stress deviator.

Remark 13.2.63 Suppose $\sigma\ \overleftrightarrow{I}_P$ and $\overleftrightarrow{S}_D$ are the isotropic and deviatoric parts of Cauchy stress tensor $\overleftrightarrow{S}$. Then

i) $\langle p_n \rangle = -\sigma$;

ii) If $\overleftrightarrow{S} = \sigma\ \overleftrightarrow{I}_P$ (i.e., $\overleftrightarrow{S}_D = \overleftrightarrow{0}$) then $p_n(\hat{n}) = -\sigma$ for all $\hat{n}$ and $\overleftrightarrow{S}_S\ (\hat{n}) = \vec{0}$ for all $\hat{n}$;

iii) $\overleftrightarrow{S}$ and $\overleftrightarrow{S}_D$ produce the same shear stress for all $\hat{n}$ (i.e., the shear stress is due entirely to the deviatoric part of $\overleftrightarrow{S}$).

Proof:

i) By (13.2.39) $tr\ \overleftrightarrow{S} = 3\sigma$. By remark (13.2.61), $tr\ \overleftrightarrow{S} = -3\langle p_n \rangle$.

ii) Substitute $\sigma\ \overleftrightarrow{I}_P$ for $\overleftrightarrow{S}$ in (13.2.35) and (13.2.34).

iii) From (13.2.34), if $\hat{n}$ is fixed, $\vec{S}_S(\hat{n})$ depends linearly on $\overleftrightarrow{S}$. By (13.2.39), $\vec{S}_S(\hat{n})$ is the sum of a contribution from $\sigma\ \overleftrightarrow{I}_P$ and one from $\overleftrightarrow{S}_D$. By ii) above, the contribution from $\sigma\ \overleftrightarrow{I}_P$ vanishes.

Remark 13.2.64 If a Cauchy stress tensor produces no shear stress for any $\hat{n}$, it is isotropic.

Proof:

The hypothesis is, from (13.2.34), that $\hat{n}\cdot \overleftrightarrow{S}= \hat{n}S_n(\hat{n})$ for all unit vectors $\hat{n}$. Let $\hat{y}_1$, $\hat{y}_2$, $\hat{y}_3$ be an orthonormal basis for P. Then setting $\hat{n} = \hat{y}_i$ gives $\hat{y}_i\cdot \overleftrightarrow{S}= \hat{y}_{(i)}S_n(\hat{y}_{(i)})$. But $\overleftrightarrow{S}=\overleftrightarrow{I}_P \cdot \overleftrightarrow{S}= (\hat{y}_i\hat{y}_i)\cdot \overleftrightarrow{S}= \hat{y}_i(\hat{y}_i\cdot \overleftrightarrow{S}) = \sum \hat{y}_i\hat{y}_iS_n(\hat{y}_i)$. That is

$$\overleftrightarrow{S}= \sum_{i=1}^{3} \hat{y}_i\hat{y}_iS_n(\hat{y}_i).$$

But then $\overleftrightarrow{S}^T=\overleftrightarrow{S}$, so $\hat{n}\cdot \overleftrightarrow{S}\cdot\hat{y}_i = \hat{y}_i\cdot \overleftrightarrow{S}\cdot\hat{n}$, and thus $\hat{n}S_n(\hat{n})\cdot\hat{y}_i = \hat{y}_iS_n(\hat{y}_i)\cdot\hat{n}$, or $S_n(\hat{n})(\hat{n}\cdot\hat{y}_i) = S_n(\hat{y}_i)(\hat{n}\cdot\hat{y}_i)$. This is true for any $\hat{n}$, so we may choose an $\hat{n}$ with $\hat{n}\cdot\hat{y}_i \neq 0$ for all of $i = 1, 2, 3$. Then $S_n(\hat{n}) = S_n(\hat{y}_1) = S_n(\hat{y}_2) = S_n(\hat{y}_3)$, so $\overleftrightarrow{S}= S_n(\hat{n})\hat{y}_i\hat{y}_i = S_n(\hat{n})\overleftrightarrow{I}_P$. QED

Fluid Motion:

As an application of Cauchy stress tensors and the momentum equation, we consider certain elementary fluid problems.

Definition 13.2.50 A fluid is a material which cannot support shear stresses when it is in static equilibrium. A non-viscous fluid is a material which can never support shear stresses.

Corollary 13.2.47 In a non-viscous fluid, or in a fluid in static equilibrium, the Cauchy stress tensor is isotropic, and $\overleftrightarrow{S}^E(\vec{r},t) = -\langle p_n\rangle(\vec{r},t)\overleftrightarrow{I}_P$.

Proof:

From remark (13.2.64), $\overleftrightarrow{S}^E(\vec{r},t) = \sigma(\vec{r},t)\overleftrightarrow{I}_P$. From remark 13.2.63, i), $\sigma = -\langle p_n\rangle$.

Definition 13.2.51 In a non-viscous fluid, or in a fluid in static equilibrium, $\langle p_n\rangle$ is called simply "the" pressure, written p. Thus

$$\overleftrightarrow{S}= -p\overleftrightarrow{I}_P . \tag{13.2.40}$$

Remark 13.2.65 $\vec{\partial}\cdot(p\overleftrightarrow{I}) = \vec{\partial}p$.

Proof:

By exercise 11a, $\vec{\partial} \cdot (p \overset{\leftrightarrow}{I}_P) = \vec{\partial} p \cdot \overset{\leftrightarrow}{I}_P + p\vec{\partial} \cdot \overset{\leftrightarrow}{I}_P$. But $\vec{\partial} \cdot \overset{\leftrightarrow}{I}_P = \overset{\leftrightarrow}{I}_P \langle 2 \rangle \vec{\partial} \overset{\leftrightarrow}{I}_P = 0$ because $\overset{\leftrightarrow}{I}_P$ is constant and $\vec{\partial} \overset{\leftrightarrow}{I}_P = 0$. Also $\vec{\partial} p \cdot \overset{\leftrightarrow}{I}_P = \vec{\partial} p$. QED.

Therefore, the Eulerian momentum equation in a non-viscous fluid is

$$\rho \vec{a} = -\vec{\partial} p + \vec{f}. \tag{13.2.41}$$

Since $\vec{a} = D_t \vec{u} = \partial_t \vec{u} + \vec{u} \cdot \vec{\partial} \vec{u}$, this is often written

$$\rho \left(\partial_t \vec{u} + \vec{u} \cdot \vec{\partial} \vec{u} \right) = -\vec{\partial} p + \vec{f}. \tag{13.2.42}$$

In a fluid in static equilibrium, $\rho \vec{a} = \vec{0}$, so

$$\vec{\partial} p = \vec{f}. \tag{13.2.43}$$

Therefore, $\vec{\partial} \times \vec{f} = \vec{0}$. We have

Remark 13.2.66 Static equilibrium in a fluid subjected to body force $\vec{f}$ is impossible unless $\vec{\partial} \times \vec{f} = \vec{0}$.

Remark 13.2.67 If a fluid is in static equilibrium in a gravitational field, then the pressure and density of the fluid are constant on gravitational equipotential surfaces which are arcwise connected.

Proof:

Let Ψ be the gravitational potential. Then $\vec{f} = -\rho \vec{\partial} \Psi$, so (13.2.43) shows that $\vec{\partial} p$ is a scalar multiple of $\vec{\partial} \Psi$. And $\vec{\partial} \times \vec{f} = -\vec{\partial} \rho \times \vec{\partial} \Psi$. Since $\vec{\partial} \times \vec{f} = \vec{0}$, $\vec{\partial} \rho$ is also a scalar multiple of $\vec{\partial} \Psi$. Therefore, if C is any curve lying in a gravitational equipotential surface, and if $\hat{\tau}(\vec{r})$ is the unit tangent vector to C at $\vec{r}$, then $\vec{\tau}(\vec{r}) \cdot \vec{\partial} \rho(\vec{r}) = \hat{\tau}(\vec{r}) \cdot \vec{\partial} p(\vec{r}) = 0$. Therefore, by theorem 11.2.23, ρ and p are the same at both ends of C.

13.3 Lagrangian form of conservation of momentum

We use the notation introduced on page 163. If $\vec{x}$ is a point in label space L, we use $dV_L(\vec{x})$ to stand for a small set containing $\vec{x}$ and also for the volume of that set. The mass per unit volume in label space is $\tilde{\rho}^L(\vec{x})$, independent of time, so the total mass of the particles whose labels are in $dV_L(\vec{x})$ is $\tilde{\rho}^L(\vec{x})dV_L(\vec{x})$. Their velocity at time t is $\vec{v}^L(\vec{x},t)$, so their momentum is $dV_L(\vec{x})\tilde{\rho}^L(\vec{x})\vec{v}^L(\vec{x},t)$. The total momentum of the particles whose labels are in open subset H' of H is, at time t,

$$\vec{P}_t^L[H'] = \int_{H'} dV_L(\vec{x})\tilde{\rho}^L(\vec{x})\vec{v}^L(\vec{x},t) = \int_{H'} dV_L(\vec{x})(\tilde{\rho}\vec{v})^L(\vec{x},t).$$

Let $\vec{\mathcal{F}}_t^L[H']$ denote the total force exerted on those particles at time t. Then the law of conservation of momentum requires

$$\frac{d}{dt}\vec{P}_t^L[H'] = \vec{\mathcal{F}}_t^L[H']. \tag{13.3.1}$$

Since the set H' does not vary with time, we also have the mathematical identity

$$\frac{d}{dt}\vec{P}_t^L[H'] = \int_{H'} dV_L(\vec{x})\tilde{\rho}^L(\vec{x})D_t\vec{v}^L(\vec{x},t),$$

or

$$\frac{d}{dt}\vec{P}_t^L[H'] = \int_{H'} dV_L(\vec{x})(\tilde{\rho}\vec{a})^L(\vec{x},t). \tag{13.3.2}$$

The particles whose labels are in H' occupy the set $K'(t)$ in physical space P at time t. Therefore $\vec{P}_t^L[H'] = \vec{P}^E[K'(t)]$, and (13.3.2) should agree with (13.2.3). That it does can be seen by changing the variables of integration from $\vec{r}$ to $\vec{x}$ in (13.2.3) by means of (11.7.6) and then using (13.1.3).

Let $\vec{\mathcal{F}}_{B,t}^L[H']$ denote the total body force on the particle with labels in H'. Then $\vec{\mathcal{F}}_{B,t}^L[H'] = \vec{\mathcal{F}}_B^E[K'(t)]$. By changing the variables of integration in (13.2.6) from $\vec{r}$ to $\vec{x}$ via (11.7.6), we find

$$\vec{F}_{B,t}^L[H'] = \int_{H'} dV_L(\vec{x})\left|\det \vec{D}\,\vec{\mathbf{r}}^{\,L}(\vec{x},t)\right|\vec{f}^L(\vec{x},t),$$

or

$$\vec{\mathcal{F}}_{B,t}^L[H'] = \int_{H'} dV_L(\vec{x})\vec{\tilde{f}}^L(\vec{x},t) \tag{13.3.3}$$

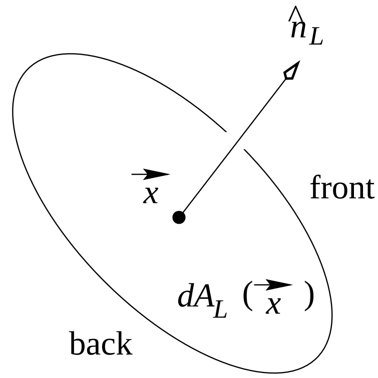


Figure 13.8:

where

$$\tilde{\vec{f}} = \left|\det \vec{D}\,\vec{\mathbf{r}}\right| \vec{f}. \tag{13.3.4}$$

The vector $\tilde{\vec{f}}$ is the body force density per unit of label-space volume. Formula (13.3.4) can be seen physically as follows. The body force on the particles in the small set $dV_P(\vec{r})$ is $dV_P(\vec{r})\vec{f}^E(\vec{r},t)$ at time t. But these are the particles whose labels are in $dV_L(\vec{x})$, so by (13.0.1) this force is $dV_L(\vec{x})\tilde{\vec{f}}^L(\vec{x},t)$ with $\tilde{\vec{f}}$ given by (13.3.4).

The surface force at time t exerted by the particles whose labels are just outside $\partial H'$ on the particles whose labels are just inside $\partial H'$ we will denote by $\vec{\mathcal{F}}^L_{S,t}[H']$. The particles with labels just outside (inside) $\partial H'$ are those whose positions at time t are just outside (inside) $\partial K'(t)$, so

$$\vec{\mathcal{F}}^L_{S,t}\left[H'\right] = \vec{\mathcal{F}}^E_S\left[K'(t)\right]. \tag{13.3.5}$$

Then $\vec{\mathcal{F}}^L_t\left[H'\right] = \vec{\mathcal{F}}^L_{B,t}\left[H'\right] + \vec{\mathcal{F}}^L_{S,t}\left[H'\right]$. It remains to study the surface force $\vec{\mathcal{F}}^L_{S,t}[H']$.

Let $dA_L(\vec{x})$ be any small nearly plane patch of surface in label space L containing the label $\vec{x} \in H$. Choose one side of dA_L to be the front, and let $\hat{n}_L$ be the unit normal to $dA_L(\vec{x})$ on its front side. Figure 13.8 is the same as figure (13.6) except for the labelling. Now, however, the picture is of a small patch of surface in label space, not physical space. The force $d\vec{\mathcal{F}}_S$ exerted by the particles with labels just in front of dA_L on the particles with labels just behind dA_L is proportional to dA_L, as long as the orientation of that small surface does not change, i.e., as long as $\hat{n}_L$ is fixed. We denote the vector "constant" of proportionality by $\tilde{\vec{S}}$. Then $d\vec{\mathcal{F}}^L_S = \tilde{\vec{S}}dA_L$. The vector $\tilde{\vec{S}}$ will change if $\hat{n}_L$ changes, and it

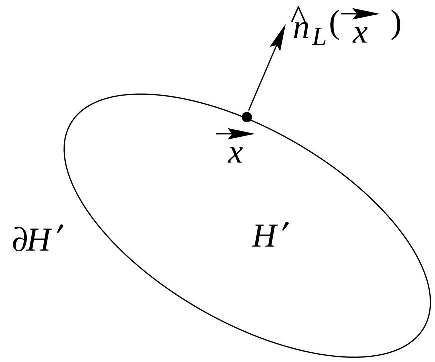


Figure 13.9:

also will depend on $\vec{x}$ and t, so at time t at label $\vec{x}$

$$d\vec{\mathcal{F}}_S^L = \vec{\tilde{S}}(\vec{x}, t, \hat{n}_L) dA_L(\vec{x}). \tag{13.3.6}$$

(Compare (13.2.7.) The vector $\vec{\tilde{S}}$ we call the "label stress". The total surface force on the particles in H' is then

$$\vec{\mathcal{F}}_{S,t}^L [H'] = \int_{\partial H'} dA_L(\vec{x}) \vec{\tilde{S}} [\vec{x}, t, \hat{n}_L(\vec{x})] \tag{13.3.7}$$

where $\hat{n}_L(\vec{x})$ is the unit outward normal to $\partial H'$ at $\vec{x} \in \partial H'$. Now we can combine (13.3.3) and (13.3.7) to obtain

$$\vec{\mathcal{F}}_t^L [H'] = \int_{H'} dV_L(\vec{x}) \tilde{f}^L(\vec{x}, t) + \int_{\partial H'} dA_L(\vec{x}) \vec{\tilde{S}} [\vec{x}, t, \hat{n}_L(\vec{x})] \,. \tag{13.3.8}$$

Inserting this and (13.3.2) in (13.3.1) gives

$$\int_{H'} dV_L(\vec{x}) \left(\tilde{\rho} \vec{a} - \vec{\tilde{f}} \right)^L (\vec{x}, t) = \int_{\partial H'} dA_L(\vec{x}) \vec{\tilde{S}} [\vec{x}, t, \hat{n}_L(\vec{x})] \,. \tag{13.3.9}$$

Equation (13.3.9) is true for every open subset H' of H for which $\partial H'$ is piecewise smooth. It is mathematically of the same form as (13.2.10) and the same application of Cauchy's theorem (13.2.28) shows that there is a unique tensor $\overset{\leftrightarrow}{\tilde{S}}{}^L (\vec{x}, t) \in L \otimes P$ such that for every unit vector $\hat{n}_L \in L$,

$$\vec{\tilde{S}}(\vec{x}, t, \hat{n}_L) = \hat{n}_L \cdot \overset{\leftrightarrow}{\tilde{S}}{}^L (\vec{x}, t). \tag{13.3.10}$$

The tensor $\overset{\leftrightarrow}{\tilde{S}}{}^L (\vec{x}, t)$ is the Piola-Kirchoff or body stress tensor. We would prefer to call it the label stress tensor, but history prevents us. Proceeding in the way which led from

(13.2.10) to (13.2.12) we note that by Gauss's theorem

$$\int_{\partial H'} dA_L(\vec{x})\hat{n}_L \cdot \overset{\leftrightarrow}{S}^L (\vec{x},t) = \int_{H'} dV_L(\vec{x})\vec{D}\cdot \overset{\leftrightarrow}{S}^L (\vec{x},t)$$

so

$$\int_{H'} dV_L(\vec{x}) \left(\tilde{\rho}\vec{a} - \vec{\tilde{f}} - \vec{D}\cdot \overset{\leftrightarrow}{\tilde{S}}\right)^L (\vec{x},t) = 0.$$

Since this is true for all open $H' \subseteq H$ with $\partial H'$ piecewise smooth, the vanishing integral theorem gives

$$\tilde{\rho}\vec{a} = \vec{\tilde{f}} + \vec{D}\cdot \overset{\leftrightarrow}{\tilde{S}}. \tag{13.3.11}$$

This is the Lagrangian form of the momentum equation. Equation (13.3.6) can now be written

$$d\vec{\mathcal{F}}_S^L = dA_L(\vec{x})\hat{n}_L \cdot \overset{\leftrightarrow}{\tilde{S}}^L (\vec{x},t), \tag{13.3.12}$$

where $d\vec{\mathcal{F}}_S^L$, $dA_L(\vec{x})$ and $\hat{n}_L$ refer to figure 13.8.

Equations (13.3.11) and (13.2.12) refer to the same motion. The relation between $\tilde{\rho}$ and ρ is (13.1.3). The relation between $\vec{\tilde{f}}$ and $\vec{f}$ is (13.3.4), which can also be written

$$\frac{\vec{\tilde{f}}}{\tilde{\rho}} = \frac{\vec{f}}{\rho}, \tag{13.3.13}$$

both sides of this equation being the force per unit mass on the material. But what is the relation between $\overset{\leftrightarrow}{\tilde{S}}$ and $\overset{\leftrightarrow}{S}$?

Suppose that $dA_L(\vec{x})$ is a small nearly plane surface in label space L, and that we have chosen one of its sides as the front. Let $\hat{n}_L$ be the unit normal to $dA_L(\vec{x})$ on its front side. At time t, let $\vec{r} = \vec{\mathbf{r}}^L(\vec{x},t)$. Let $dA_P(\vec{r})$ be the small nearly plane surface in physical space P consisting of all positions at time t of particles whose labels are in $dA_L(\vec{x})$. Let $dA_L(\vec{x})$ and $dA_P(\vec{r})$ denote both the sets of points which make up the small surfaces and also the areas of those sets. *Choose the front side of $dA_P(\vec{r})$ so that particles are just in front of $dA_P(\vec{r})$ if their labels are just in front of $dA_L(\vec{x})$.* Let $\hat{n}_P$ be the unit normal to $dA_P(\vec{r})$ on its front side. Then the force exerted on the particles just behind $dA_P(\vec{r})$ by the particles just in front can be written either as $dA_P(\vec{r})\hat{n}_P\cdot \overset{\leftrightarrow}{S}^E (\vec{r},t)$ or as $dA_L(\vec{x})\hat{n}_L \cdot \overset{\leftrightarrow}{\tilde{S}}^L (\vec{x},t)$. Therefore, these two expressions must be equal:

$$dA_P(\vec{r})\hat{n}_P \cdot \overset{\leftrightarrow}{S}^E (\vec{r},t) = dA_L(\vec{x})\hat{n}_L \cdot \overset{\leftrightarrow}{\tilde{S}}^L (\vec{x},t) \tag{13.3.14}$$

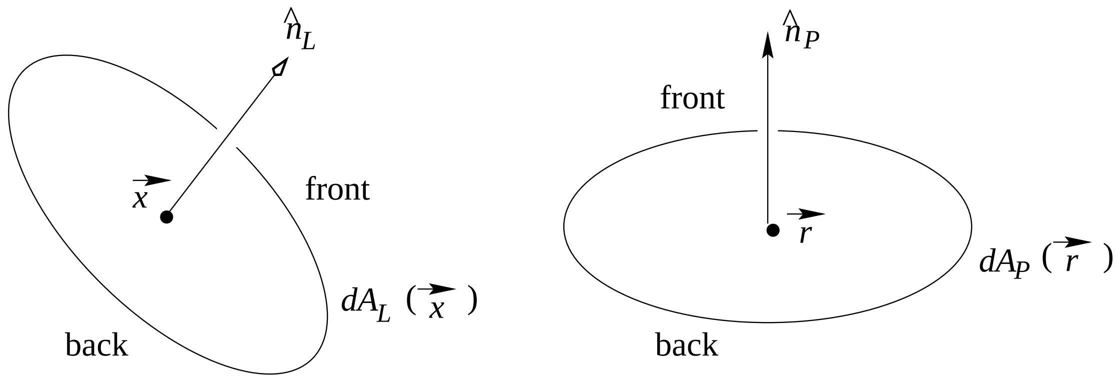


Figure 13.10:

if

$$\vec{r} = \vec{\mathbf{r}}^{\,L}(\vec{x}, t) \tag{13.3.15}$$

and

$$dA_P(\vec{r}) = \left[\vec{\mathbf{r}}^{\,L}(\cdot, t)\right][dA_L(\vec{x})]\,. \tag{13.3.16}$$

Notice that in (13.3.14), dA_P and dA_L are positive real numbers, areas, while in (13.3.16) they stand for sets, the small surfaces whose areas appear in (13.3.14).

Our program is to express $dA_P(\vec{r})\hat{n}_P$ in terms of $dA_L(\hat{x})\hat{n}_L$ in (13.3.14), to cancel $dA_L(\hat{x})\hat{n}_L$, and thus to obtain the relation between $\overset{\leftrightarrow}{S}$ and $\tilde{\overset{\leftrightarrow}{S}}$. Equation 13.3.14 holds, whatever the shape of $dA_L(\vec{x})$. It is convenient to take it to be a small parallelogram whose vertices are $\vec{x}$, $\vec{x}+\vec{\xi}_1$, $\vec{x}+\vec{\xi}_2$, $\vec{x}+\vec{\xi}_1+\vec{\xi}_2$. Then $dA_P(\vec{r})$ will be a very slightly distorted parallelogram with vertices

$$\begin{aligned}
\vec{r} &= \vec{\mathbf{r}}^{\,L}(\vec{x}, t)\\
\vec{r}+\vec{\rho}_1 &= \vec{\mathbf{r}}^{\,L}(\vec{x}+\vec{\xi}_1, t)\\
\vec{r}+\vec{\rho}_2 &= \vec{\mathbf{r}}^{\,L}(\vec{x}+\vec{\xi}_2, t)\\
\vec{r}+\vec{\rho}_1+\vec{\rho}_2 &= \vec{\mathbf{r}}^{\,L}(\vec{x}+\vec{\xi}_1+\vec{\xi}_2, t).
\end{aligned}$$

Correct to first order in the small length $\|\vec{\xi}_i\|$, we have

$$\vec{r}+\vec{\rho}_i = \vec{\mathbf{r}}^{\,L}(\vec{x}, t) + \vec{\xi}_i \cdot \vec{D}\,\vec{\mathbf{r}}^{\,L}(\vec{x}, t),$$

so

$$\vec{\rho}_i = \vec{\xi}_i \cdot \vec{D}\,\vec{\mathbf{r}}^{\,L}(\vec{x}, t). \tag{13.3.17}$$

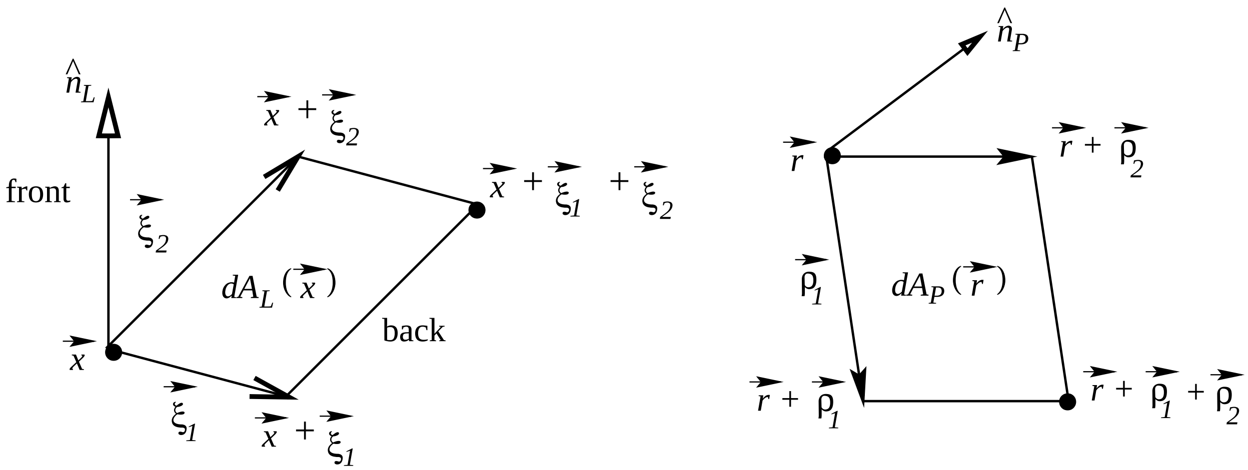


Figure 13.11:

We naturally take $\vec{\xi}_1$ and $\vec{\xi}_2$ in different directions, so $\|\vec{\xi}_1 \times \vec{\xi}_2\| \neq 0$ (recall that L and P are oriented 3-spaces, so cross products are defined). We number $\vec{\xi}_1$ and $\vec{\xi}_2$ in the order which ensures that $\vec{\xi}_1 \times \vec{\xi}_2$ points into the region in front of $dA_L(\vec{x})$, so $\vec{\xi}_1 \times \vec{\xi}_2 = \|\vec{\xi}_1 \times \vec{\xi}_2\|\hat{n}_L$. But $\|\vec{\xi}_1 \times \vec{\xi}_2\| = dA_L(\vec{x})$, so

$$dA_L(\vec{x})\hat{n}_L = \vec{\xi}_1 \times \vec{\xi}_2. \tag{13.3.18}$$

Also, clearly, $\|\vec{\rho}_1 \times \vec{\rho}_2\| = dA_P(\vec{r})$, and $\vec{\rho}_1 \times \vec{\rho}_2$ is perpendicular to $dA_P(\vec{r})$, so $\vec{\rho}_1 \times \vec{\rho}_2 = \pm\|\vec{\rho}_1 \times \vec{\rho}_2\|\hat{n}_P$. Thus

$$dA_P(\vec{r})\hat{n}_P = \pm\vec{\rho}_1 \times \vec{\rho}_2. \tag{13.3.19}$$

But which sign is correct? We have to work out which is the front side of $dA_P(\vec{r})$, because that is where we put $\hat{n}_P$. If $\vec{\xi}_3 \in L$ is small, $\vec{x} + \vec{\xi}_3$ is in front of $dA_L(\vec{x}) \Leftrightarrow \vec{\xi}_3 \cdot \hat{n}_L > 0$, and hence $\Leftrightarrow$

$$\vec{\xi}_3 \cdot \left(\vec{\xi}_1 \times \vec{\xi}_2\right) > 0. \tag{13.3.20}$$

The particle with label $\vec{x} + \vec{\xi}_3$ has position $\vec{r} + \vec{\rho}_3$, with $\vec{\rho}_3$ given by (13.3.17). Then

$$\left\{\begin{array}{l} \vec{\rho}_3 \cdot (\vec{\rho}_1 \times \vec{\rho}_2) = A_P\left(\vec{\rho}_3, \vec{\rho}_1, \vec{\rho}_2\right) = A_P\left(\vec{\rho}_1, \vec{\rho}_2, \vec{\rho}_3\right) \\ \\ \quad = A_P\left(\vec{\xi}_1 \cdot \vec{D}\,\vec{\mathbf{r}}^{\,L}, \vec{\xi}_2 \cdot \vec{D}\,\vec{\mathbf{r}}^{\,L}, \vec{\xi}_3 \cdot \vec{D}\,\vec{\mathbf{r}}^{\,L}\right) \\ \\ \quad = \left[\det \vec{D}\,\vec{\mathbf{r}}^{\,L}(\vec{x}, t)\right] A_L\left(\vec{\xi}_1, \vec{\xi}_2, \vec{\xi}_3\right) = \left[\det \vec{D}\,\vec{\mathbf{r}}^{\,L}(\vec{x}, t)\right] \vec{\xi}_3 \cdot \left(\vec{\xi}_1 \times \vec{\xi}_2\right). \end{array}\right. \tag{13.3.21}$$

Thus the position $\vec{r}+\vec{\rho}_3$ is in front of $dA_P(\vec{r}) \Leftrightarrow \vec{\rho}_3\cdot(\vec{\rho}_1\times\vec{\rho}_2)$ has the same sign as $\det \vec{D}\,\vec{\mathbf{r}}^{\,L}(\vec{x},t)$. Let $\operatorname{sgn} c$ stand for the sign of the real number c. That is, $\operatorname{sgn} c = +1$ if $c>0$, $=-1$ if $c<0$, and $=0$ if $c=0$. Then position $\vec{r}+\vec{\rho}_3$ is in front of $dA_P(\vec{r}) \Leftrightarrow$

$$\operatorname{sgn} \vec{\rho}_3\cdot(\vec{\rho}_1\times\vec{\rho}_2) = \operatorname{sgn}\, \det \vec{D}\,\vec{\mathbf{r}}^{\,L}(\vec{x},t). \tag{13.3.22}$$

We have chosen the front of $dA_P(\vec{r})$, and the direction of $\hat{n}_P$, so that $\vec{r}+\varepsilon\hat{n}_P$ is in front of $dA_P(\vec{r})$ when $\varepsilon>0$. Therefore $\operatorname{sgn} \hat{n}_P\cdot(\vec{\rho}_1\times\vec{\rho}_2) = \operatorname{sgn}\, \det \vec{D}\,\vec{\mathbf{r}}^{\,L}(\vec{x},t)$. Thus

$$\vec{\rho}_1\times\vec{\rho}_2 = \|\vec{\rho}_1\times\vec{\rho}_2\|\hat{n}_P \operatorname{sgn}\, \det \vec{D}\,\vec{\mathbf{r}}^{\,L}(\vec{x},t)$$

and

$$dA_P(\vec{r})\hat{n}_P = (\vec{\rho}\times\vec{\rho}_2)\ \operatorname{sgn}\, \det \vec{D}\,\vec{\mathbf{r}}^{\,L}(\vec{x},t). \tag{13.3.23}$$

Using (13.3.18) and (13.3.23), we hope to relate to $dA_P\hat{n}_P$ and $dA_L\hat{n}_L$, so we try to relate $\vec{\rho}_1\times\vec{\rho}_2$ and $\vec{\xi}_1\times\vec{\xi}_2$. The two ends of (13.3.21) give

$$\vec{\rho}_3\cdot(\vec{\rho}_1\times\vec{\rho}_2) = \vec{\xi}_3\cdot\left(\vec{\xi}_1\times\vec{\xi}_2\right)\det \vec{D}\,\vec{\mathbf{r}}_L(\vec{x},t) \tag{13.3.24}$$

for any $\vec{\xi}_1$, $\vec{\xi}_2$, $\vec{\xi}_3\in L$, if $\vec{\rho}_1$, $\vec{\rho}_2$, $\vec{\rho}_3$ are given by (13.3.17).

Substituting (13.3.18) and (13.3.23) in (13.3.24) and using $c \operatorname{sgn} c = |c|$, we find

$$\vec{\rho}_3\cdot[dA_P(\vec{r})\hat{n}_P] = \vec{\xi}_3\cdot[dA_L(\vec{x})\hat{n}_L]\left|\det \vec{D}\,\vec{\mathbf{r}}^{\,L}(\vec{x},t)\right|.$$

But $\vec{\rho}_3 = \vec{\xi}_3\cdot\vec{D}\,\vec{\mathbf{r}}^{\,L}(\vec{x},t)$, so

$$\vec{\xi}_3\cdot\vec{D}\,\vec{\mathbf{r}}^{\,L}(\vec{x},t)\cdot[dA_P(\vec{r})\hat{n}_P] = \vec{\xi}_3\cdot[dA_L(\vec{x})\hat{n}_L]\left|\det \vec{D}\,\vec{\mathbf{r}}^{\,L}(\vec{x},t)\right|.$$

Since this is true for any $\vec{\xi}_3\in L$, we must have

$$\vec{D}\,\vec{\mathbf{r}}^{\,\vdots}(\vec{x},t)\cdot[dA_P(\vec{r})\hat{n}_P] = dA_L(\vec{x})\hat{n}_L\left|\det \vec{D}\,\vec{\mathbf{r}}^{\,L}(\vec{x},t)\right|$$

or, letting ^{T} stand for transpose,

$$dA_P(\vec{r})\hat{n}_P\cdot\left[\vec{D}\,\vec{\mathbf{r}}^{\,L}(\vec{x},t)\right]^T = dA_L(\vec{x})\hat{n}_L\left|\det \vec{D}\,\vec{\mathbf{r}}^{\,L}(\vec{x},t)\right|. \tag{13.3.25}$$

If we dot $\overset{\leftrightarrow}{\tilde{S}}{}^L(\vec{x},t)$ on the right of each side of (13.3.25), we obtain

$$
\begin{aligned}
& dA_P(\vec{r})\hat{n}_P \cdot \vec{D}\,\vec{\mathbf{r}}^{\,L}(\vec{x},t)^T \cdot \overset{\leftrightarrow}{\tilde{S}}{}^L(\vec{x},t) \\
& = dA_L(\vec{x})\hat{n}_L \cdot \overset{\leftrightarrow}{\tilde{S}}{}^L(\vec{x},t)\left|\det \vec{D}\,\vec{\mathbf{r}}^{\,L}(\vec{x},t)\right|.
\end{aligned} \tag{13.3.26}
$$

But (see 13.3.14)

$$
d\vec{\mathcal{F}}_S = dA_L(\vec{x})\hat{n}_L \cdot \overset{\leftrightarrow}{\tilde{S}}{}^L(\vec{x},t) = dA_P(\vec{r})\hat{n}_P \cdot \overset{\leftrightarrow}{S}{}^E(\vec{r},t)
$$

where $\vec{r} = \vec{\mathbf{r}}^{\,L}(\vec{x},t)$. Thus

$$
dA_L(\vec{x})\hat{n}_L \cdot \overset{\leftrightarrow}{\tilde{S}}{}^L(\vec{x},t) = dA_P(\vec{r})\hat{n}_P \cdot \overset{\leftrightarrow}{S}{}^L(\vec{x},t).
$$

Substituting this on the right side of (13.3.26) gives

$$
\begin{aligned}
& [dA_P(\vec{r})\hat{n}_P] \cdot \left[\vec{D}\,\vec{\mathbf{r}}^{\,L}(\vec{x},t)^T \cdot \overset{\leftrightarrow}{\tilde{S}}{}^L(\vec{x},t)\right] \\
& = [dA_P(\vec{r})\hat{n}_P] \cdot \left[\overset{\leftrightarrow}{S}{}^L(\vec{x},t)\left|\det \vec{D}\,\vec{\mathbf{r}}^{\,L}(\vec{x},t)\right|\right].
\end{aligned} \tag{13.3.27}
$$

We have proved (13.3.27) when dA_P is a small parallelogram, as in Figure (13.3.14). But $\vec{\rho}_1$ and $\vec{\rho}_2$ can be arbitrary as long as they are small. Hence so can $dA_P(\vec{r})\hat{n}_P = \vec{\rho}_1 \times \vec{\rho}_2$. Thus (13.3.27) is true when $dA_P(\vec{r})\hat{n}_P$ is any small vector in P. Since (13.3.27) is linear in $dA_P(\vec{r})\hat{n}_P$, it is true whatever the size of that vector. Therefore (13.3.27) implies

$$
\vec{D}\,\vec{\mathbf{r}}^{\,L}(\vec{x},t)^T \cdot \overset{\leftrightarrow}{\tilde{S}}{}^L(\vec{x},t) = \overset{\leftrightarrow}{S}{}^L(\vec{x},t)\left|\det \vec{D}\,\vec{\mathbf{r}}^{\,L}(\vec{x},t)\right|,
$$

or

$$
\left(\vec{D}\,\vec{\mathbf{r}}\right)^T \cdot \overset{\leftrightarrow}{\tilde{S}} = \overset{\leftrightarrow}{S}\left|\det \vec{D}\,\vec{\mathbf{r}}\right|. \tag{13.3.28}
$$

This gives $\overset{\leftrightarrow}{S}$ in terms of $\overset{\leftrightarrow}{\tilde{S}}$. By (12.2.7, 13.3.24), $\vec{D}\,\vec{\mathbf{r}}^{\,L}(\vec{x},t)$ has an inverse, $\vec{\partial}\,\vec{\mathbf{x}}^{\,E}(\vec{r},t)$, where $\vec{r} = \vec{\mathbf{r}}^{\,L}(\vec{x},t)$, so $\vec{D}\,\vec{\mathbf{r}}^{\,L}(\vec{x},t)^T$ has the inverse $\vec{\partial}\,\vec{\mathbf{x}}^{\,E}(\vec{r},t)^T$. Dotting this on the left in (13.3.28) gives

$$
\overset{\leftrightarrow}{\tilde{S}} = \left(\vec{\partial}\,\vec{\mathbf{x}}\right)^T \cdot \overset{\leftrightarrow}{S}\left|\det \vec{\partial}\,\vec{\mathbf{x}}\right|^{-1} \tag{13.3.29}
$$

where we use $\det(\vec{D}\,\vec{\mathbf{r}}\,)^{-1} = (\det \vec{D}\vec{r})^{-1}$.

By appealing to (13.3.1), we can write (13.3.28) and (13.3.29) as

$$\left(\frac{\overset{\leftrightarrow}{S}}{\rho}\right) = \left(\vec{D}\,\vec{\mathbf{r}}\right)^T \cdot \left(\frac{\tilde{\overset{\leftrightarrow}{S}}}{\tilde{\rho}}\right) \tag{13.3.30}$$

$$\left(\frac{\tilde{\overset{\leftrightarrow}{S}}}{\tilde{\rho}}\right) = \left(\vec{\partial}\,\vec{\mathbf{x}}\right)^T \cdot \left(\frac{\overset{\leftrightarrow}{S}}{\rho}\right). \tag{13.3.31}$$

13.4 Conservation of angular momentum

13.4.1 Eulerian version

Refer again to the notation on page 163. All torques and angular momenta are calculated relative to the origin $\vec{O}_P$ in physical space P. Let $\vec{\mathcal{L}}^E[K'(t)]$ denote the total angular momentum of the particles in $K'(t)$, while $\vec{\mathcal{M}}^E[K'(t)]$ denotes the total torque on them. The law of conservation of angular momentum is

$$\frac{d}{dt}\vec{\mathcal{L}}^E\left[K'(t)\right] = \vec{\mathcal{M}}^E\left[K'(t)\right]. \tag{13.4.1}$$

Our problem is to express this law locally.

The momentum of the particles in $dV_P(\vec{r})$ is $dV_P(\vec{r})\rho^E(\vec{r},t)\vec{v}^E(\vec{r},t)$ and this produces, relative to $\vec{O}_P$, the angular momentum

$$\vec{r} \times \left[dV_P(\vec{r})(\rho\vec{v})^E(\vec{r},t)\right] = dV_P(\vec{r})(\vec{r} \times \rho\vec{v})^E(\vec{r},t).$$

It is possible that the individual atoms in $dV_P(\vec{r})$ might have angular momentum; this might be from the orbits of electrons not in closed shells, from the spins of unpaired electrons, or from nuclear spin. Orbital angular momentum is quantized in integer multiples of $\hbar$, while spin angular momentum is in units of $\hbar/2$. Here $\hbar = h/2\pi$ and h is Planck's constant, so $\hbar = 1.0544 \times 1^{-34}$ joule sec. If the atoms are not randomly aligned, there may be an intrinsic atomic angular momentum of $\vec{l}^E(\vec{r},t)$ per kilogram at position $\vec{r}$ at time t. Then the intrinsic angular momentum in $dV_P(\vec{r})$ is $dV_P(\vec{r})(\rho\vec{l})^E(\vec{r},t)$, and the total angular momentum in $dV_P(\vec{r})$ is

$$d\vec{\mathcal{L}}^E = dV_P(\vec{r})\left[\rho\left(\vec{r} \times \vec{v} + \vec{l}\right)\right]^E(\vec{r},t).$$

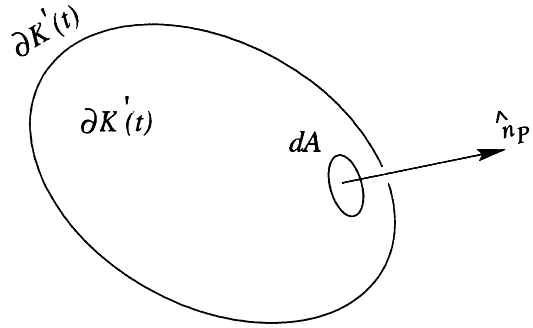


Figure 13.12:

The total angular momentum in $K'(t)$ is

$$\mathcal{L}^E\left[K'(t)\right] = \int_{K'(t)} dV_P(\vec{r})\left[\rho\left(\vec{r}\times\vec{v}+\vec{l}\right)\right]^E(\vec{r},t).$$

From (13.1.12),

$$\frac{d}{dt}\mathcal{L}^E\left[K'(t)\right] = \int_{K'(t)} dV_P(\vec{r})\left[\rho D_t\left(\vec{r}\times\vec{v}+\vec{l}\right)\right]^E(\vec{r},t).$$

Now

$$D_t(\vec{r}\times\vec{v}) = (D_t\vec{r})\times\vec{v}+\vec{r}\times D_t\vec{v} = \vec{v}\times\vec{v}+\vec{r}\times\vec{a} = \vec{r}\times\vec{a},$$

so

$$\frac{d}{dt}\mathcal{L}^E\left[K'(t)\right] = \int_{K'(t)} dV_P(\vec{r})\left[\rho\left(\vec{r}\times\vec{a}+D_t\vec{l}\right)\right]^E(\vec{r},t). \tag{13.4.2}$$

To calculate the torque on the particles in $K'(t)$ we note that the torque exerted by the body force $dV_P(\vec{r})\vec{f}^E(\vec{r},t)$ is $\vec{r}\times dV_P(\vec{r})\vec{f}^E(\vec{r},t)$, or $dV_P(\vec{r})(\vec{r}\times\vec{f})^E(\vec{r},t)$. The torque exerted by the surface force $dA_P(\vec{r})\vec{S}(\vec{r},t,\hat{n}_P)$ is $\vec{r}\times dA_P(\vec{r})\vec{S} = -dA_P(\vec{r})\vec{S}\times\vec{r} = -dA_P(\vec{r})[\hat{n}\cdot \overleftrightarrow{S}^E(\vec{r},t)]\times\vec{r}$. If $\overleftrightarrow{Q}\in P\otimes P$ is a polyad, $(\hat{n}\cdot\overleftrightarrow{Q})\times\vec{r} = \hat{n}\cdot[\overleftrightarrow{Q}\times\vec{r}]$, so this is true for all $\overleftrightarrow{Q}$. (Exercise 11b gives a definition of $\vec{r}\times\overleftrightarrow{Q}$, and $\overleftrightarrow{Q}\times\vec{r}$ is defined similarly.) Thus

$$\vec{r}\times dA_P(\vec{r})\vec{S}(\vec{r},t,\hat{n}_P) = -dA_P(\vec{r})\hat{n}_P\cdot\left[\overleftrightarrow{S}\times\vec{r}\right]^E(\vec{r},t).$$

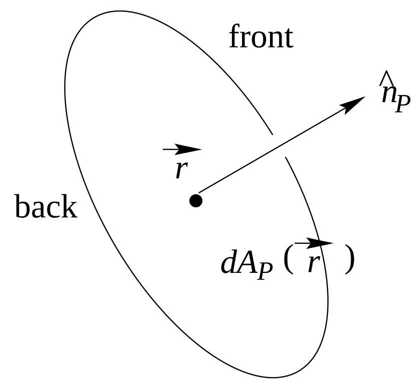


Figure 13.13:

Therefore, the torque about $\vec{O}_P$ which is exerted on the particles in $K'(t)$ by the body and surface force acting on them is

$$\int_{K'(t)} dV_P(\vec{r}) \left(vecr \times \vec{f} \right)^E (\vec{r}, t) - \int_{\partial K'(t)} dA_P(\vec{r}) \hat{n}_P \cdot \left(\overleftrightarrow{S} \times \vec{r} \right)^E (\vec{r}, t).$$

By applying Gauss's theorem to the surface integral, we can write this torque as

$$\int_{K'(t)} dV_P(\vec{r}) \left[\vec{r} \times \vec{f} - \vec{\partial} \cdot \left(\overleftrightarrow{S} \times \vec{r} \right) \right]^E (\vec{r}, t). \tag{13.4.3}$$

In addition to this torque, there may be a torque acting on each atom in $dV_P(\vec{r})$. This would be true, for example, if the material were a solid bar of magnetized iron placed in a magnetic field $\vec{B}$. If the magnetization density was $\vec{M}$, there would be a torque $dV_P(\vec{r})\vec{M} \times \vec{B}$ acting on $dV_P(\vec{r})$ and not included in (13.4.3). In general, we might want to allow for an intrinsic body torque of $\vec{m}$ joules/meter^3, so that (13.4.3) must be supplemented by a term

$$\int_{K'(t)} dV_P(\vec{r}) \vec{m}^E (\vec{r}, t).$$

Finally, in a magnetized iron bar, atoms just outside $\partial K'(t)$ exert a torque on atoms just inside $\partial K'(t)$, so there is a "torque stress" acting on $\partial K'(t)$. The torque exerted on the material just behind $dA_P(\vec{r})$ by the material just in front is proportional to dA_P as long as the orientation of that small patch of surface, i.e., its unit normal, stays fixed. We write the proportionality constant as $\vec{M}(\vec{r}, t, \hat{n}_P)$. The torque exerted on the material just behind dA_P by the material just in front, other than that due to $\hat{n}_P \cdot \overleftrightarrow{S}^E (\vec{r}, t)$, is

$\vec{M}(\vec{r},t,\hat{n}_P)dA_P(\vec{r})$, so we must also supplement (13.4.3) by a term

$$\int_{\partial K'(t)} dA_P(\vec{r})\vec{M}\left(\vec{r},t,\hat{n}_P(\vec{r},t)\right),$$

where $\hat{n}_P$ is the outward unit normal to $\partial K'(t)$. Thus we finally have

$$\begin{aligned}\vec{\mathcal{M}}^E\left[K'(t)\right] &= \int_{K'(t)} dV_P(\vec{r})\left[\vec{r}\times\vec{f}-\vec{\partial}\cdot\left(\overleftrightarrow{S}\times\vec{r}\right)+\vec{m}\right]^E(\vec{r},t)\\ &\quad+\int_{\partial K'(t)} dA_P(\vec{r})\vec{M}\left[\vec{r},t,\hat{n}_P(\vec{r},t)\right].\end{aligned}\tag{13.4.4}$$

If we substitute (13.4.2) and (13.4.4) in (13.4.1), we obtain

$$\begin{aligned}&\textstyle\int_{K'(t)} dV_P(\vec{r})\left[\rho\left(\vec{r}\times\vec{a}+D_t\vec{l}\right)-\vec{r}\times\vec{f}+\vec{\partial}\cdot\left(\overleftrightarrow{S}\times\vec{r}\right)-\vec{m}\right]^E(\vec{r},t)\\ &\qquad=\textstyle\int_{\partial K'(t)} dA_P(\vec{r})\vec{M}\left[\vec{r},t,\hat{n}_P(\vec{r},t)\right].\end{aligned}\tag{13.4.5}$$

This equation has the same mathematical form as (13.2.10), and leads via Cauchy's theorem to the same conclusion. At $\vec{r},t$ there is a unique tensor $\overleftrightarrow{M}^E(\vec{r},t)\in P\otimes P$, which we call the (Cauchy) torque-stress tensor, such that for every unit vector $\hat{n}_P\in P$ we have

$$\vec{M}\left(\vec{r},t,\hat{n}_P\right)=\hat{n}_P\cdot\overleftrightarrow{M}^E(\vec{r},t,).\tag{13.4.6}$$

Substituting this on the right in (13.4.5), applying Gauss's theorem to the surface integral, and the vanishing integral theorem to the volume integral, gives

$$\rho\left(\vec{r}\times\vec{a}+D_t\vec{l}\right)-\vec{r}\times\vec{f}+\vec{\partial}\cdot\left(\overleftrightarrow{S}\times\vec{r}\right)-\vec{m}=\vec{\partial}\cdot\vec{M}.\tag{13.4.7}$$

Before we accept this as the Eulerian angular momentum equation, we note that it can be greatly simplified by using the momentum equation (13.2.12). We have $\rho(\vec{r}\times\vec{a})=\vec{r}\times\vec{f}$ $=\vec{r}\times\left(\rho\vec{a}-\vec{f}\right)=\vec{r}\times(\vec{\partial}\cdot\overleftrightarrow{S})$, so (13.4.7) is

$$\rho D_t\vec{l}+\vec{r}\times\left(\vec{\partial}\cdot\overleftrightarrow{S}\right)+\vec{\partial}\cdot\left(\overleftrightarrow{S}\times\vec{r}\right)=\vec{m}+\vec{\partial}\cdot\overleftrightarrow{M}.\tag{13.4.8}$$

This can be still further simplified. Let $\hat{y}_1$, $\hat{y}_2$, $\hat{y}_3$ be a poob in P, $(\hat{x}_1,\hat{x}_2,\hat{x}_3)$ a pooob in L, and take components relative to these bases. Then

$$\left[\vec{r}\times(\vec{\partial}\cdot\overleftrightarrow{S})\right]_i=\varepsilon_{ijk}r_j(\vec{\partial}\cdot\overleftrightarrow{S})_k=\varepsilon_{ijk}r_j\partial_l S_{lk}.$$

Also, for $\vec{u}$, $\vec{v}$, $\vec{w} \in P$, $(\vec{u}\vec{v}) \times \vec{w} = \vec{u}(\vec{v} \times \vec{w})$ (this is how we use the lifting theorem to define $\overset{\leftrightarrow}{Q} \times \vec{w}$ for $\overset{\leftrightarrow}{Q} \in P \otimes P$ and $\vec{w} \in P$) so

$$\begin{aligned}
[(\vec{u}\vec{v}) \times \vec{w}]_{ij} &= [\vec{u}\,(\vec{v} \times \vec{w})]_{ij} = u_i\,(\vec{v} \times \vec{w})_j \\
&= u_i \varepsilon_{jkl} \nu_k w_l \\
&= (\vec{u}\vec{v})_{ik}\, \varepsilon_{jkl} w.
\end{aligned}$$

Therefore for any $\overset{\leftrightarrow}{Q} \in P \otimes P$,

$$\left[\overset{\leftrightarrow}{Q} \times \vec{w}\right]_{ij} = Q_{ik} \varepsilon_{jkl} w_l.$$

Therefore

$$\begin{aligned}
\left(\overset{\leftrightarrow}{S} \times \vec{r}\right)_{li} &= S_{lk} \varepsilon_{ikj} r_j \\
\left[\vec{\partial} \cdot \left(\overset{\leftrightarrow}{S} \times \vec{r}\right)\right]_i &= \partial_l \left(\overset{\leftrightarrow}{S} \times \vec{r}\right)_{li} = \partial_l \left[S_{lk} \varepsilon_{ikj} r_j\right] \\
&= \varepsilon_{ikj} \partial_l \left(S_{lk} r_j\right) \\
&= \varepsilon_{ikj} \left[\left(\partial_l S_{lk}\right) r_j + S_{lk} \partial_l r_j\right].
\end{aligned}$$

Referring to page 199 and using $\varepsilon_{ijk} + \varepsilon_{ijk} = 0$, we get

$$\left[\vec{r} \times \left(\vec{\partial} \cdot \vec{S}\right) + \vec{\partial} \cdot \left(\overset{\leftrightarrow}{S} \times \vec{r}\right)\right]_i = \varepsilon_{ikj} S_{lk} \partial_l r_j.$$

But $\partial_l r_j = \delta_{lj}$, so this is $\varepsilon_{ikj} S_{jk} = -\epsilon_{ijk} S_{jk} = -\left(A_P \langle 2 \rangle \overset{\leftrightarrow}{S}\right)_i$. Thus

$$\vec{r} \times \left(\vec{\partial} \cdot \overset{\leftrightarrow}{S}\right) + \vec{\partial} \cdot \left(\overset{\leftrightarrow}{S} \times \vec{r}\right) = -A_P \langle 2 \rangle \overset{\leftrightarrow}{S}. \tag{13.4.9}$$

Thus (13.4.8) is

$$\rho D_t \vec{l} = A_P \langle 2 \rangle \overset{\leftrightarrow}{S} + \vec{m} + \vec{\partial} \cdot \overset{\leftrightarrow}{M}. \tag{13.4.10}$$

This is the Eulerian form of the angular momentum equation. In non-magnetic materials it is believed that $\vec{l}$, $\vec{m}$ are $\overset{\leftrightarrow}{M}$ and all negligible, and that the essential content of (13.4.10) is

$$A_P \langle 2 \rangle \overset{\leftrightarrow}{S} = \vec{0}. \tag{13.4.11}$$

That is, $\varepsilon_{ijk} S_{jk} = 0$. Multiply by ε_{ilm} and sum on i, and this becomes (see page D-9)

$$(\delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl})\, S_{jk} = 0,$$

or

$$S_{lm} - S_{ml} = 0,$$

or

$$\overleftrightarrow{S}^T = \overleftrightarrow{S}\,. \tag{13.4.12}$$

The conservation of angular momentum requires that the Cauchy stress tensor be symmetric, if intrinsic angular momentum, body torque and torque stress are all negligible.

13.4.2 Consequences of $\overleftrightarrow{S}^T = \overleftrightarrow{S}$

In the absence of intrinsic angular momentum, torque and torque stress, the Cauchy stress tensor is symmetric. This greatly simplifies visualizing the stress at a point $\vec{r}$ at time t. We will abbreviate $\overleftrightarrow{S}^E(\vec{r}, t)$ as $\overleftrightarrow{S}$ in this chapter, because we are looking at one time t and one position $\vec{r}$ in physical space. We denote by $\vec{S}(\hat{n})$ the stress on the oriented small area $(dA, \hat{n})$ at $\vec{r}, t$ (i.e. we write $\vec{S}(\vec{r}, t, \hat{n}_P)$ as $\vec{S}(\hat{n})$, and $\hat{n}_P$ as $\hat{n}$, dA_P as dA). We denote by S the linear operator on P corresponding to the tensor $\overleftrightarrow{S}$. Then for any unit vector $\hat{n}$,

$$\vec{S}(\hat{n}) = \hat{n} \cdot \overleftrightarrow{S} = S(\hat{n}). \tag{13.4.13}$$

Since $S : P \to P$ is a symmetric operator, P has an orthonormal basis $\hat{x}, \hat{y}, \hat{z}$ consisting of eigenvectors of S. That is, there are scalars a, b, c such that

$$S(\hat{x}) = a\hat{x}, \quad S(\hat{y}) = b\hat{y}, \quad S(\hat{z}) = c\hat{z}. \tag{13.4.14}$$

Then $\hat{x} \cdot \overleftrightarrow{S} = a\hat{x}$, $\hat{y} \cdot \overleftrightarrow{S} = b\hat{y}$, $\hat{z} \cdot \overleftrightarrow{S} = c\hat{z}$ so

$$\overleftrightarrow{S} = \overleftrightarrow{I}_P \cdot \overleftrightarrow{S} = (\hat{x}\hat{x} + \hat{y}\hat{y} + \hat{z}\hat{z}) \cdot \overleftrightarrow{S} = \hat{x}\,(a\hat{x}) + \hat{y}\,(b\hat{y}) + \hat{z}\,(c\hat{z})$$

so

$$\overleftrightarrow{S} = a\hat{x}\hat{x} + b\hat{y}\hat{y} + c\hat{z}\hat{z}. \tag{13.4.15}$$

The unit vectors $\hat{x}$, $\hat{y}$, $\hat{z}$ are called "principal axes of stress" and a, b, c are "principal stresses." By relabelling, if necessary, we can always assume that $a \geq b \geq c$. Also if we change the sign of any of $\hat{x}, \hat{y}, \hat{z}$ both (13.4.14) and (13.4.15) remain unchanged. Therefore we may assume that $(\hat{x}, \hat{y}, \hat{z})$ is positively oriented.

For any vector $\vec{n} \in P$ we write

$$\vec{n} = x\hat{x} + y\hat{y} + z\hat{z}. \tag{13.4.16}$$

The unit sphere N_P consists of those $\vec{n} \in P$ for which

$$x^2 + y^2 + z^2 = 1. \tag{13.4.17}$$

The pressure $p_n(\hat{n})$ acting on $(dA, \hat{n})$ is $p_n(\hat{n}) = -\hat{n} \cdot \overset{\leftrightarrow}{S} \cdot \hat{n} = -(ax^2 + by^2 + cz^2)$. It can be thought of as a function of position $\hat{n} = x\hat{x} + y\hat{y} + z\hat{z}$ on the spherical surface (13.4.17). In the earth, $p_n > 0$, so it is convenient to introduce the principal pressures A, B, C, defined by

$$A = -a, \quad B = -b, \quad C = -c. \tag{13.4.18}$$

Then $A \leq B \leq C$. The cases where equality occurs are special and are left to the reader. We will examine the usual case,

$$A < B < C. \tag{13.4.19}$$

We want to try to visualize $p_n(\hat{n})$ and the shear stress $\vec{S}_S(\hat{n})$ acting on $(dA, \hat{n})$ as functioning of $\hat{n}$, i.e., as functions on the spherical surface (13.4.17).

We have

$$p_n(\hat{n}) = Ax^2 + By^2 + Cz^2 \tag{13.4.20}$$

$$\vec{S}(\hat{n}) = \vec{S}_S(\hat{n}) - p_n(\hat{n})\hat{n} \tag{13.4.21}$$

and, since $\vec{S}(\hat{n}) = \hat{n} \cdot \overset{\leftrightarrow}{S}$,

$$\vec{S}(\hat{n}) = -Ax\hat{x} - By\hat{y} - Cz\hat{z}. \tag{13.4.22}$$

First consider p_n. Its value is unchanged if we replace any of x, y, z by its negative in (13.4.16). Therefore p_n is symmetric under reflections in the coordinate planes $x = 0$,

$y = 0$, $z = 0$. Therefore it suffices to study $p_n(\hat{n})$ when $\hat{n}$ is in the first octant of (13.4.17), i.e., when

$$x \geq 0, \quad y \geq 0, \quad z \geq 0. \tag{13.4.23}$$

If p is any constant, then the level line $p_n(\hat{n}) = p$ on N_p satisfies

$$Ax^2 + By^2 + Cz^2 = p \tag{13.4.24}$$

$$x^2 + y^2 + z^2 = 1. \tag{13.4.25}$$

Hence, it satisfies

$$(A - p)\, x^2 + (B - p)\, y^2 + (C - p)\, z^2 = 0. \tag{13.4.26}$$

This equation and (13.4.17) have no real solutions (x, y, z) unless

$$A \leq p \leq C, \tag{13.4.27}$$

so the values of $p_n(\hat{n})$ all lie in the interval (13.4.27). The value $p = A$ requires $y = z = 0$, $x = \pm 1$, so $p_n(\hat{n}) = A$ only for $\hat{n} = \pm\hat{x}$. The value $p = C$ requires $x = y = 0$, so $p_n(\hat{n}) = C$ only for $\hat{n} = \pm\hat{z}$. The value of $p = B$ makes (13.4.25) into $x\sqrt{B - A} = \pm z\sqrt{C - B}$, and these two planes intersect the sphere (13.4.17) in great circles, one of which lies in (13.4.23). For any value of p in $A < p < C$, (13.4.25) and (13.4.26) imply all three of the following:

$$\begin{cases} (C - A)\, x^2 + (C - B)\, y^2 = C - p \\ (B - A)\, x^2 - (C - B)\, z^2 = B - p \\ (B - A)\, y^2 + (C - A)\, z^2 = p - A. \end{cases} \tag{13.4.28}$$

Therefore the projection of the level curve (13.4.25), (13.4.17) onto the xy or yz plane is part or all of an ellipse, and its projection on the xz plane is part of a hyperbola with asymptotes $x\sqrt{B - A} = \pm z\sqrt{C - B}$. The level lines $p_n(\hat{n}) = B$ on N_p, with $p = B$, are the great circles where N_p intersects the planes $x\sqrt{B - A} = \pm z\sqrt{C - B}$. The level lines of p_n on N_P are the dashed curves in figure 13.14. Next, consider the shear stress $\vec{S}_S(\hat{n})$. At any $\hat{n} \in N_P$, this is a vector tangent to N_P. It is defined by (13.4.21), and from (13.4.22) $\vec{S}(\hat{n}) = -1/2\, \vec{\partial} p_n(\hat{n})$ where p_n is to be considered as a function defined for *all* $\vec{n} \in P$ by (13.4.20), even when $\|\vec{n}\| \neq 1$. At a point $\hat{n} \in N_P$, let $\hat{\tau}(\hat{n})$ be the unit

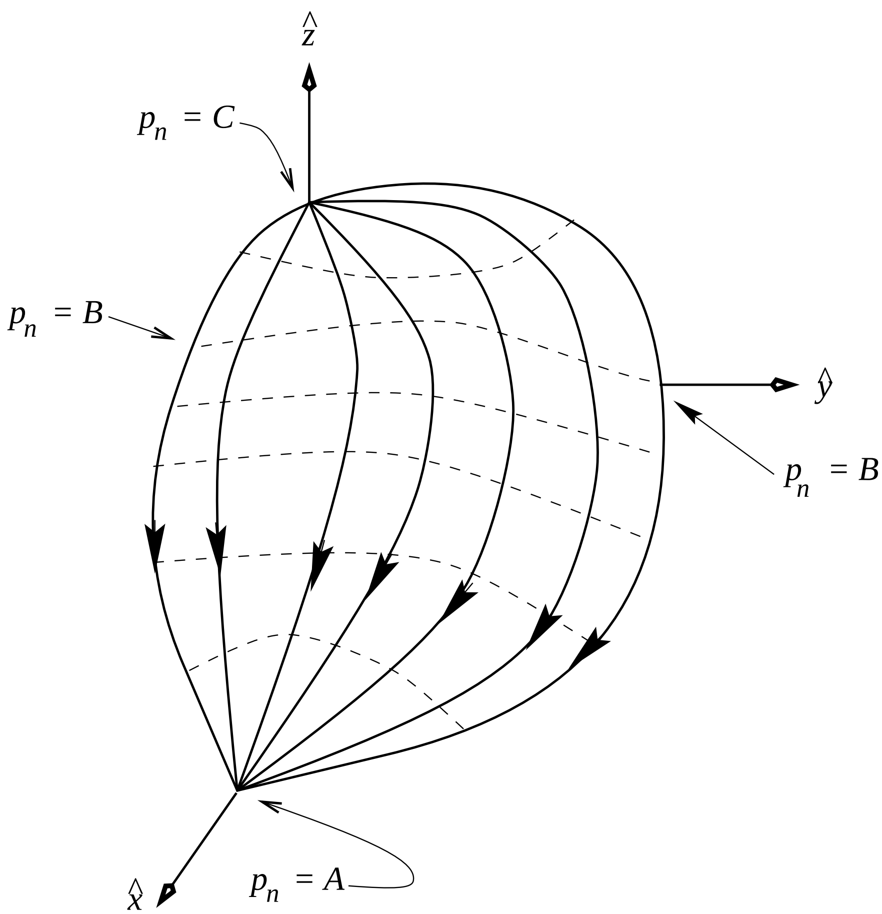


Figure 13.14:

vector tangent to the level curve of p_n passing through $\hat{n}$. Then $\hat{n}\cdot\hat{\tau}(\hat{n}) = 0$, and since $\hat{\tau}(\hat{n})$ is tangent to a curve lying in the level surface $p_n(\vec{n}) =$ constant in P, it follows that $\hat{\tau}(\hat{n})\cdot\vec{\partial}p_n(\hat{n}) = 0$. Therefore, from (13.4.29), we have

$$\hat{\tau}(\hat{n})\cdot\vec{S}(\hat{n}) = 0$$

and

$$\hat{\tau}(\hat{n})\cdot\hat{n} = 0.$$

Then from (13.4.21), $\hat{\tau}(\hat{n})\cdot\vec{S}_S(\hat{n}) = 0$. Thus the shear stress $\vec{S}_S(\hat{n})$ at $\hat{n}\in N_P$ is perpendicular to the level line of p_n which passes through $\hat{n}$. The "lines of force" of the shear stress on N_P are the curves which are everywhere tangent to $\vec{S}_S(\hat{n})$. They are the solid curves in figure 13.14, and are everywhere perpendicular to the level lines of p_n. These "lines of force" give the direction of $\vec{S}_S(\hat{n})$ everywhere on N_P.

Finally, we would like to find $\|\vec{S}_S(\hat{n})\|$ on N_P, which we abbreviate as

$$\sigma(\hat{n}) = \|\vec{S}_S(\hat{n})\|. \tag{13.4.29}$$

Again we consider only the first octant, $x\geq 0$, $y\geq 0$, $z\geq 0$. The other octants are obtained by reflection in the three coordinate planes $x=0$, $y=0$, $z=0$. First consider the edges of the first octant, where $x=0$ or $y=0$ or $z=0$. On the quarter circle $\hat{n} = \hat{x}\cos\theta + \hat{y}\sin\theta$, $0\leq\theta\leq\pi/2$, subjecting (13.4.20) and (13.4.21), (13.4.22) to a little algebra yields

$$\vec{S}_S(\hat{n}) = \sin\theta\cos\theta\,(\hat{x}\cos\theta - \hat{y}\sin\theta) \text{ on } \hat{n} = \hat{x}\cos\theta + \hat{y}\sin\theta. \tag{13.4.30}$$

Therefore, from (13.4.29),

$$\sigma(\hat{n}) = \frac{1}{2}(B-A)\sin 2\theta \text{ if } \hat{n} = \hat{x}\cos\theta + \hat{y}\sin\theta. \tag{13.4.31}$$

This reaches a maximum value of $(B-A)/2$ at $\theta = \pi/4$, half way between $\hat{x}$ and $\hat{y}$ on the quarter circle. And $\sigma(\hat{x}) = \sigma(\hat{y}) = 0$. Similarly

$$\sigma(\hat{n}) = 1/2(C-A)\sin 2\theta \text{ if } \hat{n} = \hat{x}\cos\theta + \hat{z}\sin\theta \tag{13.4.32}$$

$$\sigma(\hat{n}) = 1/2(C-B)\sin 2\theta \text{ if } \hat{n} = \hat{y}\cos\theta + \hat{z}\sin\theta. \tag{13.4.33}$$

To see what happens to $\sigma(\hat{n})$ when $x>0$, $y>0$, $z>0$, we appeal to

Lemma 13.4.33 On any level line of p_n in the first octant on N_P, $\sigma(\hat{n})$ decreases monotonically as y increases.

Proof:

From (13.4.21), $\|\vec{S}(\hat{n})\|^2 = \sigma(\hat{n})^2 + p_n(\hat{n})^2$ so

$$\sigma(\hat{n})^2 = \|\vec{S}(\hat{n})\|^2 - p_n(\hat{n})^2. \tag{13.4.34}$$

We can parametrize the level curve $p_n(\hat{n}) = p$ on N_P by the value of y at $\hat{n}$, namely $\hat{y} \cdot \hat{n} = y$. We want to show that

$$\frac{d}{dy}\sigma(\hat{n}) < 0 \text{ on the curve } p_n(\hat{n}) = p \text{ on } N_P{}'{}_s \text{ first octant.} \tag{13.4.35}$$

On that level curve, of course $p_n(\hat{n})$ is constant, so (13.4.34) shows that (13.4.35) is equivalent to

$$\frac{d}{dy}\|\vec{S}(\hat{n})\|^2 < 0 \text{ on the curve } p_n(\hat{n}) = p \text{ on } N_p{}'{}_s \text{ first octant.} \tag{13.4.36}$$

The vector $d\hat{n}/dy$ on $p_n(\hat{n}) = p$ is tangent to that level curve. Hence $\hat{n} \cdot d\hat{n}/dy = 0$ and $\vec{\partial}p(\hat{n}) \cdot d\hat{n}/dy = 0$, so $\vec{S}(\hat{n}) \cdot d\hat{n}/dy = 0$. Hence, for some λ

$$\frac{d\hat{n}}{dy} = \lambda \hat{n} \times \vec{S}(\hat{n}).$$

Now $d\hat{n}/dy = \hat{x}(dx/dy) + \hat{y} + \hat{z}(dz/dy)$ on the level curve. Hence

$$\hat{y} \cdot \frac{d\hat{n}}{dy} = 1, \text{ so } 1 = \lambda\, \hat{y} \cdot \hat{n} \times \vec{S}(\hat{n}).$$

But $\hat{n} = \hat{x}x + \hat{y}y + \hat{z}z$, $-\vec{S}(\hat{n})Ax\hat{x} + By\hat{y} + Cz\hat{z}$, so

$$-\hat{y} \cdot \hat{n} \times \vec{S}(\hat{n}) = \begin{vmatrix} 0 & 1 & 0 \\ x & y & z \\ Ax & By & Cz \end{vmatrix} = -xz(C - A).$$

Thus $\lambda = 1/xz(C - A)$, and

$$\frac{d\hat{n}}{dy} = \frac{1}{xz(C - A)}\hat{n} \times \vec{S}(\hat{n}). \tag{13.4.37}$$

By the chain rule, $\frac{d}{dy}\|\vec{S}(\hat{n})\|^2 = \frac{d\hat{n}}{dy} \cdot \vec{\partial}\|\vec{S}(\hat{n})\|^2$. From (13.4.22),

$$\begin{aligned} \|\vec{S}(\vec{n})\|^2 &= A^2x^2 + B^2y^2 + C^2z^2, \text{ so} \\ \vec{\partial}\|\vec{S}(\vec{n})\|^2 &= 2A^2x\hat{x} + 2B^2y\hat{y} + 2C^2z\hat{z}. \end{aligned}$$

Therefore

$$\begin{aligned} \frac{d}{dy}\|\vec{S}(\hat{n})\|^2 &= \frac{2}{xz(C-A)}\left[\hat{n}\times\vec{S}(\hat{n})\right]\cdot\left(A^2x\hat{x} + B^2y\hat{y} + C^2z\hat{z}\right) \\ &= \frac{2}{xz(C-A)}\begin{vmatrix} x & y & z \\ -Ax & -By & -Cz \\ A^2x & B^2y & C^2z \end{vmatrix} = -\frac{2y}{(C-A)}\begin{vmatrix} 1 & 1 & 1 \\ A & B & C \\ A^2 & B^2 & C^2 \end{vmatrix} \\ &= -2y(C-B)(B-A) < 0. \end{aligned}$$

This proves lemma 13.4.30 .

QED.

From (13.4.31, 13.4.32, 13.4.33) and lemma 13.4.33, we infer that the largest value of $\sigma(\hat{n})$, the magnitude of the shear stress, is $(C-A)/2$, and this occurs at $\hat{n} = (\pm\hat{x}\pm\hat{z})/\sqrt{2}$, and for no other $\hat{n}$.

It is of some interest to see which pairs $(p_n(\hat{n}), \sigma(\hat{n}))$ are possible on N_P. On the three quarter-circles in (13.4.31, 13.4.32, 13.4.33) we have

$$\begin{pmatrix} p_n(\hat{n}) \\ \sigma_n(\hat{n}) \end{pmatrix} = \begin{pmatrix} \frac{A+B}{2} - \frac{B-A}{2}\cos 2\theta \\ \frac{B-A}{2}\sin 2\theta \end{pmatrix} \text{ if } \hat{n} = \hat{x}\cos\theta + \hat{y}\sin\theta \quad (z=0) \tag{13.4.38}$$

$$\begin{pmatrix} p_n(\hat{n}) \\ \sigma_n(\hat{n}) \end{pmatrix} = \begin{pmatrix} \frac{A+C}{2} - \frac{C-A}{2}\cos 2\theta \\ \frac{C-A}{2}\sin 2\theta \end{pmatrix} \text{ if } \hat{n} = \hat{x}\cos\theta + \hat{z}\sin\theta \quad (y=0) \tag{13.4.39}$$

$$\begin{pmatrix} p_n(\hat{n}) \\ \sigma_n(\hat{n}) \end{pmatrix} = \begin{pmatrix} \frac{B+C}{2} - \frac{C-B}{2}\cos 2\theta \\ \frac{C-B}{2}\sin 2\theta \end{pmatrix} \text{ if } \hat{n} = \hat{y}\cos\theta + \hat{z}\sin\theta. \quad (x=0) \tag{13.4.40}$$

For $z = 0$, the possible pairs $(p_n(\hat{n}), \sigma(\hat{n}))$ are given by (13.4.38) and trace out the seimicircle marked $z = 0$ in Figure 13.15. For $y = 0$, the possible pairs are given by

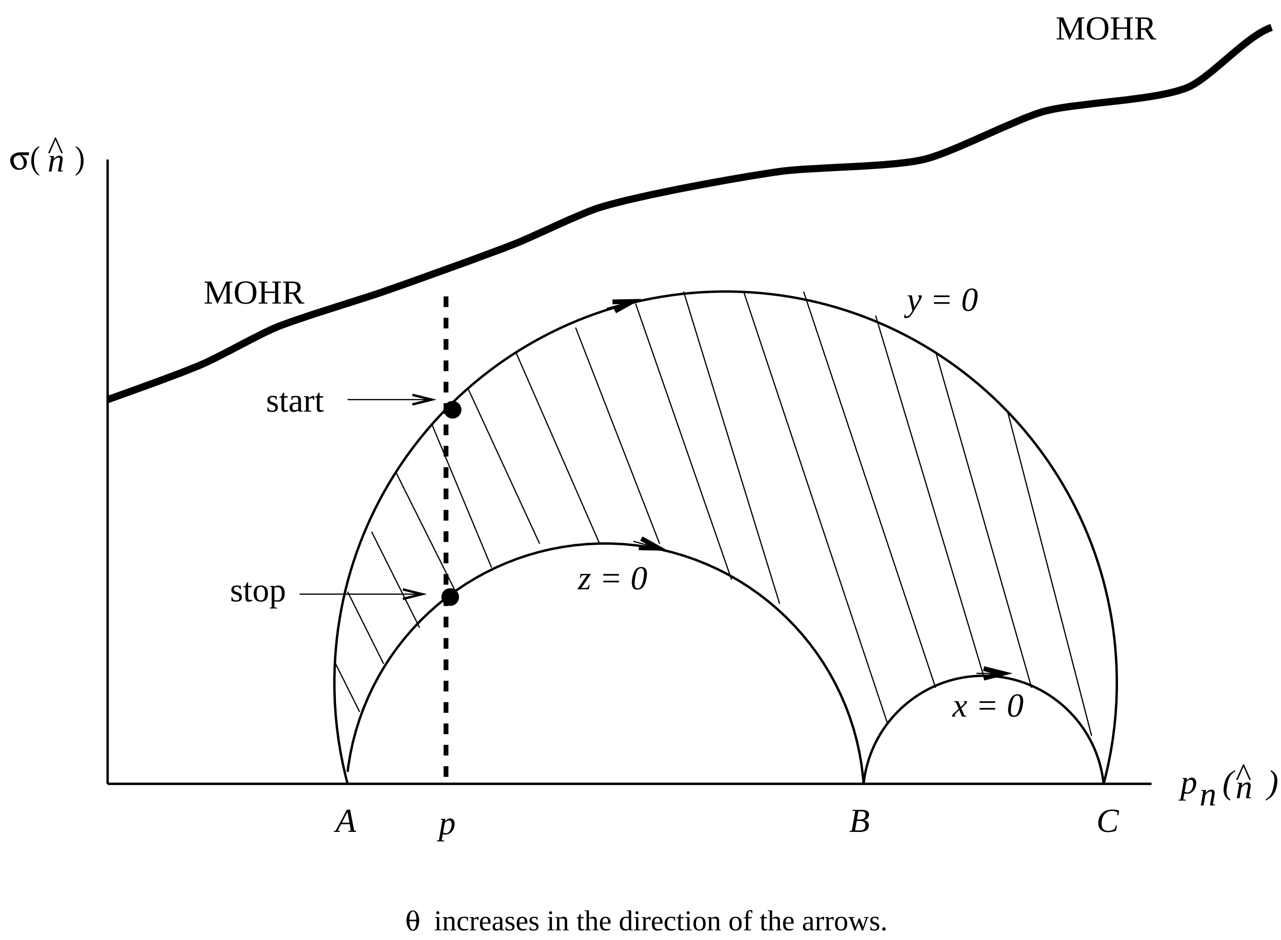

MOHR
σ(n̂)
MOHR
y = 0
start
stop
z = 0
x = 0
p_n(n̂)
A
p
B
C
θ increases in the direction of the arrows.

Figure 13.15:

(13.4.39) and trace out the semicircle marked $y = 0$ in Figure 13.15. For $x = 0$, use the third semicircle in Figure 13.4.38.

In figure 13.14, suppose we start at a point with $y = 0$ and move along the level line of p_n till we hit either $x = 0$ or $z = 0$. Then in the (σ, p_n) plane we will start at the point marked "start" in Figure 13.15, and we will remain on the vertical dashed line $p_n = p$. According to lemma 13.4.33, as we increase y on the level curve $p_n - p$ in Figure (13.14), we will decrease $\sigma(\hat{n})$, so we move down the dashed line in Figure 13.15 until we strike the point marked "stop." Therefore, the possible pairs $(p_n(\hat{n}), \sigma(\hat{n}))$ which can occur on N_P are precisely the points in the shaded region in Figure 13.15.

Coulomb suggested that a rock will break when the maximum value of $\sigma(\hat{n})$ exceeds a critical value characteristic of that rock. Navier pointed out that when two plates are pressed together, the greater this pressure the harder it is to slide one over the other. Navier suggested that the fracture criterion is that there be an orientation $\hat{n}$ for which $\sigma(\hat{n}) > \sigma_c + \mu p_n(\hat{n})$, where σ_c and μ are constants characteristic of the rock (when $p(\hat{n}) = 0$, Navier's fracture criterion reduces to Coulomb's). The constant μ is a sort of internal coefficient of friction. Mohr suggested that Coulomb and Navier had over-simplified the problem, and that the true fracture criterion is $\sigma(n) > f[p(\hat{n})]$, where f is a function characteristic of the rock, which must be measured empirically, and about which one can say in advance only that

$$f(p) > 0, \qquad df/dp > 0. \tag{13.4.41}$$

Even with such a general criterion, much can be said. The curve $\sigma = f(p)$ is marked "MOHR" in Figure 13.15. If A, B, C are moved so as to produce an overlap of the shaded region with the Mohr curve, the rock will break at first contact of the curve and the shaded region. This first contact will always occur on the plane $y = 0$, at a value of θ in 13.4.32 which is between 0 and $\pi/4$, so the normal to the plane of fracture will lie in the xz plane, closer to $\hat{x}$ than to $\hat{z}$. The circles in Figure 13.15 are called Mohr circles.

13.5 Lagrangian Form of Angular Momentum Conservation

Refer again to the notation introduced on page 163. If $\vec{\mathcal{L}}_t^L(H')$ is the angular momentum of the collection of particles with labels in H', and $\vec{\mathcal{M}}_t^L(H')$ is the total torque on them, the law of conservation of angular momentum is

$$\frac{d}{dt}\vec{\mathcal{L}}_t^L(H') = \mathcal{M}_t^L(H'). \tag{13.5.1}$$

Using the notation introduced on page 195 and page 164,

$$\vec{\mathcal{L}}_t^L(H') = \int_{H'} dV_L(\vec{x}) \left[\tilde{\rho}\vec{l} + \tilde{\rho}\vec{r}\times\vec{v}\right]^L(\vec{x},t)$$

so

$$\frac{d}{dt}\vec{\mathcal{L}}_t^L(H') = \int_{H'} dV_L(\vec{x}) \left[\tilde{\rho}\left(D_t\vec{l} + \vec{r}\times\vec{a}\right)\right]^L(\vec{x},t)\,. \tag{13.5.2}$$

Moreover,

$$\begin{aligned}\mathcal{M}_t^L(H') \quad &= \quad \int_{H'} dV_L(\vec{x}) \left(\vec{r}\times\vec{\tilde{f}} + \vec{m}\right)^L(\vec{x},t) \\ &\quad + \int_{H'} dA_L(\vec{x}) \left[\vec{r}\times\left(\hat{n}_L\cdot\overset{\leftrightarrow}{\tilde{S}}{}^L\right) + \vec{\tilde{M}}\left(\vec{x},t,\hat{n}_L(\vec{x})\right)\right]\end{aligned}$$

where $\hat{n}_L(\vec{x})$ is the unit outward normal to $\partial H'$ at $\vec{x}\in\partial H'$, $\vec{\tilde{m}}$ is body torque per unit of volume in label space, and $\vec{\tilde{M}}(\vec{x},t,\hat{n}_L(\vec{x}))dA_L$ is the force exerted by the material with labels just in front of $(dA_L,\hat{n}_L)$ on the material with labels just behind $(dA_L,\hat{n}_L)$. Again we have $\vec{r}\times(\hat{n}_L\cdot\overset{\leftrightarrow}{\tilde{S}}) = -(\hat{n}_L\cdot\overset{\leftrightarrow}{\tilde{S}})\times\vec{r} = -\hat{n}_L\cdot(\overset{\leftrightarrow}{\tilde{S}}\times\vec{r})$, so using Gauss's theorem,

$$\mathcal{M}_t^L(H') \quad = \quad \int_{H'} dV_L(\vec{x}) \left[\vec{r}\times\vec{\tilde{f}} + \vec{\tilde{m}} - \vec{D}\cdot\left(\overset{\leftrightarrow}{\tilde{S}}\times\vec{r}\right)\right]^L(\vec{x},t) \tag{13.5.3}$$

$$+\int_{H'} dA_L(\vec{x})\,\vec{\tilde{M}}\left[\vec{x},t,\hat{n}_L(\vec{x})\right]. \tag{13.5.4}$$

Therefore

$$\begin{aligned}&\textstyle\int_{H'} dV_L(\vec{x}) \left[\tilde{\rho}\left(D_t\vec{l} + \vec{r}\times\vec{a}\right) - \vec{r}\times\vec{\tilde{f}} - \vec{\tilde{m}} + \vec{D}\cdot\left(\overset{\leftrightarrow}{\tilde{S}}\times\vec{r}\right)\right] \\ &\qquad = \textstyle\int_{\partial H'} dA_L(\vec{x})\,\vec{\tilde{M}}\left[\vec{x},t,\hat{n}_L(\vec{x})\right].\end{aligned} \tag{13.5.5}$$

Applying Cauchy's theorem to (13.5.5) shows that for each $(\vec{x}, t)$ there is a unique $\overset{\leftrightarrow}{\tilde{M}}{}^L(\vec{x}, t) \in L \otimes P$ such that for any unit vector $\hat{n} \in L$,

$$\overset{\leftrightarrow}{\tilde{M}}[\vec{x}, t, \hat{n}] = \hat{n} \cdot \overset{\leftrightarrow}{\tilde{M}}{}^L(\vec{x}, t). \tag{13.5.6}$$

If we substitute this on the right in (13.5.5) and use Gauss's theorem and the vanishing integral theorem, we obtain

$$\tilde{\rho} D_t \vec{l} + \tilde{\rho} \vec{r} \times \vec{a} - \vec{r} \times \vec{\tilde{f}} - \vec{\tilde{m}} + \vec{D} \cdot \left(\overset{\leftrightarrow}{\tilde{S}} \times \vec{r} \right) = \vec{D} \cdot \overset{\leftrightarrow}{\tilde{M}}.$$

Using the Lagrangian momentum equation (13.3.11), we can write this as

$$\tilde{\rho} D_t \vec{l} + \vec{r} \times \left(\vec{D} \cdot \overset{\leftrightarrow}{\tilde{S}} \right) + \vec{D} \cdot \left(\overset{\leftrightarrow}{\tilde{S}} \times \vec{r} \right) = \vec{D} \cdot \overset{\leftrightarrow}{\tilde{M}} + \tilde{m}.$$

Finally, by the sort of argument which led to (13.4.9), we can show

$$\vec{r} \times \left(\vec{D} \cdot \overset{\leftrightarrow}{\tilde{S}} \right) + \vec{D} \cdot \left(\overset{\leftrightarrow}{\tilde{S}} \times \vec{r} \right) = -A_P \langle 2 \rangle \left[\left(\vec{D} \vec{r} \right)^T \cdot \overset{\leftrightarrow}{\tilde{S}} \right],$$

so the Lagrangian form of the angular momentum equation is

$$\tilde{\rho} D_t \vec{l} = \vec{D} \cdot \overset{\leftrightarrow}{\tilde{M}} + \overset{\leftrightarrow}{\tilde{m}} + A_P \langle 2 \rangle \left[\left(\vec{D} \vec{r} \right)^T \cdot \overset{\leftrightarrow}{\tilde{S}} \right]. \tag{13.5.7}$$

The relation between $\overset{\leftrightarrow}{\tilde{M}}$ and $\overset{\leftrightarrow}{M}$ is the same as that between $\overset{\leftrightarrow}{\tilde{S}}$ and $\overset{\leftrightarrow}{S}$. If the intrinsic angular momentum $\vec{l}$, body torque $\tilde{m}$, and torque stress $\overset{\leftrightarrow}{\tilde{M}}$ are all negligible, (13.5.7) reduces to

$$A_P \langle 2 \rangle \left[\left(\vec{D} \vec{r} \right)^T \cdot \overset{\leftrightarrow}{\tilde{S}} \right] = 0. \tag{13.5.8}$$

A glance at (13.3.30) shows that (13.5.8) is the same as (13.4.11), so we learn nothing new from (13.5.7) in this case.

13.6 Conservation of Energy

13.6.1 Eulerian form

The velocity $\vec{v}^E(\vec{r}, t)$ imparts to the matter in $dV_P(\vec{r})$ a kinetic energy $dV_P(\vec{r}) \rho^E(\vec{r}, t) \frac{1}{2} \| \vec{v}^E(\vec{r}, t) \|^2$. In addition, that material has kinetic energy of molecular motion, and the potential energy of the intermolecular forces. The sum of these latter is called the internal energy. We denote by $U^E(\vec{r},t)$ the internal energy per kilogram in the matter at $\vec{r},t$. Then the total energy in dV_P is

$$dV_P(\vec{r})\left[\rho\left(\frac{1}{2}v^2+U\right)\right]^E(\vec{r},t)\,.$$

The total energy in $K'(t)$ is

$$\mathcal{E}^E\left[K'(t)\right]=\int_{K'(t)}dV_P(\vec{r})\left[\rho\left(\frac{1}{2}v^2+U\right)\right]^E(\vec{r},t)\,. \tag{13.6.1}$$

The law of conservation of energy is

$$\frac{d}{dt}\mathcal{E}^E\left[K'(t)\right]=\mathcal{W}^E\left[K'(t)\right] \tag{13.6.2}$$

where $\mathcal{W}^E[K'(t)]$ is the rate at which energy is being added to the matter in $K'(t)$, i.e., the rate at which work is being done on that matter.

The body force $\vec{f}$ does work on $dV_P(\vec{r})$ at the rate of $dV_P(\vec{r})(\vec{v}\cdot\vec{f})^E(\vec{r},t)$ watts. If $\vec{r}\in\partial K'(t)$ then the surface stress $\hat{n}_P\cdot\overset{\leftrightarrow}{S}{}^E(\vec{r},t)$ does work on the particles just inside $\partial K'(t)$ at the rate of $dA_P(\vec{r})\hat{n}_P\cdot\overset{\leftrightarrow}{S}{}^E(\vec{r},t)\cdot\vec{v}^E(\vec{r},t)$ watts. Thus the purely mechanical component of $\mathcal{W}^E[K'(t)]$ is

$$\int_{K'(t)}dV_P\left(\vec{r}\right)\left(\vec{v}\cdot\vec{f}\right)^E(\vec{r},t)+\int_{\partial K'(t)}dA_P(\vec{r})\hat{n}_P\cdot\left(\overset{\leftrightarrow}{S}\cdot\vec{v}\right)^E(\vec{r},t)\,.$$

Using Gauss's theorem, this can be written

$$\int_{K'(t)}dV_P\left(\vec{r}\right)\left[\vec{v}\cdot\vec{f}+\vec{\partial}\cdot\left(\overset{\leftrightarrow}{S}\cdot\vec{v}\right)\right]^E(\vec{r},t)\,. \tag{13.6.3}$$

In addition to this mechanical method of adding energy to $K'(t)$ there are non-mechanical methods. Examples are

i) ohmic "heating" at a rate of $dV_P(\vec{r})\|\vec{J}\|^2/\chi$ watts, where $\vec{J}=$ electric current density and $\chi=$ electrical conductivity.

ii) "heating" by absorption of light (e.g., sunlight in muddy water).

iii) radioactive "heating".

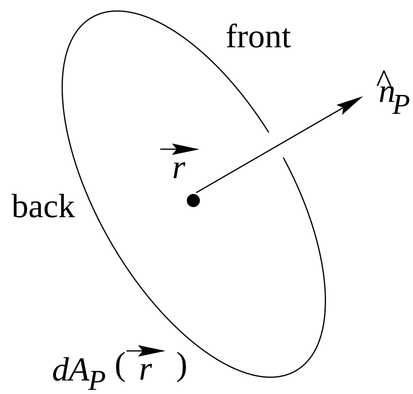


Figure 13.16:

All these "heating" mechanisms together will add energy to $dV_P(\vec{r})$ at a total rate of $dV_P(\vec{r})h^E(\vec{r},t)$, where h^E is the sum of the heating rates per unit volume. The resulting contribution to $\mathcal{W}^E[K'(t)]$ is

$$\int_{K'(t)} dV_P(\vec{r})\, h^E(\vec{r},t) \text{ watts.} \tag{13.6.4}$$

Finally, energy can leak into $K'(t)$ through the boundary, $\partial K'(t)$. The physical mechanisms for this are molecular collision and diffusion, but that does not concern us. We are interested only in the possibility that if $dA_P(\vec{r})$ is a small nearly plane surface containing $\vec{r}$, with a designated front to which a unit normal $\hat{n}_P$ is attached, the energy can leak across $dA_P(\vec{r})$ from front to back at a rate proportional to dA_P as long as the orientation of dA_P is not varied, i.e., as long as $\hat{n}_P$ is fixed. We write the constant of proportionality as $-H(\vec{r},t,\hat{n}_P)$, so that energy flows from front to back across $dA(\vec{r})$ at a net rate of $-H(\vec{r},t,\hat{n}_P)dA_P(\vec{r})$ watts. This "heat flow" contributes to $\mathcal{W}^E[K'(t)]$ the term

$$-\int_{\partial K'(t)} dA_P(\vec{r})\, H(\vec{r},t,\hat{n}_P) \text{ watts.}$$

The minus sign in the definition of the proportionality constant is a convention established for two centuries, and we must accept it. Indeed, there is aesthetic reason to put a minus sign in the definition of $\vec{S}$, to avoid the $-p\overleftrightarrow{I}$ term. Our choice of signs for $\vec{S}$ and $\overleftrightarrow{S}$ was also dictated by history, and it is unfortunate that the same choice was not made in both cases.

The sum of all the foregoing rates at which energy is added to the matter in $K'(t)$ gives

$$\mathcal{W}^E\left[K'(t)\right] = \int_{K'(t)} dV_P(\vec{r}) \left[\vec{v}\cdot\vec{f} + \vec{\partial}\cdot\left(\overleftrightarrow{S}\cdot\vec{v}\right) + h\right]^E(\vec{r},t) \quad (13.6.5)$$
$$- \int_{\partial K'(t)} dA_P\left(\vec{r}\right) H\left(\vec{r},t,\hat{n}_P\right). \quad (13.6.6)$$

From (13.6.1) and (13.1.12),

$$\frac{d\mathcal{E}^E[K'(t)]}{dt} = \int_{K'(t)} dV_P(\vec{r}) \left[\rho D_t\left(\frac{1}{2}v^2 + U\right)^E\right](\vec{r},t). \quad (13.6.7)$$

Substituting (13.6.6) and (13.6.7) in (13.6.2) gives

$$\int_{K'(t)} dV_P\left(\vec{r}\right)\left[\rho D_t\left(\frac{1}{2}v^2 + U\right) - \vec{v}\cdot\vec{f} - \vec{\partial}\cdot\left(\overleftrightarrow{S}\cdot\vec{v}\right) - h\right]^E(\vec{r},t) \quad (13.6.8)$$
$$= -\int_{\partial K'(t)} dA\left(\vec{r}\right) H\left(\vec{r},t,\hat{n}\right). \quad (13.6.9)$$

Applying Cauchy's theorem 13.2.28 to (13.6.9) shows that at each $(\vec{r},t)$ there is a unique vector $\vec{H}^E(\vec{r},t) \in P$ such that for all unit vectors $\hat{n} \in P$

$$H\left(\vec{r},t,\hat{n}\right) = \hat{n}\cdot\vec{H}^E\left(\vec{r},t\right). \quad (13.6.10)$$

The vector $\vec{H}^E(\vec{r},t)$ is called the heat-flow vector. Heat flows from front to back across $(dA_P,\hat{n}_P)$ at the rate of $-dA_P(\vec{r})\hat{n}_P\cdot\vec{H}^E(\vec{r},t)$ watts. Inserting (13.6.10) in (13.6.9) and applying Gauss's theorem converts (13.6.9) to

$$0 = \int_{K'(t)} dV_P\left(\vec{r}\right)\left[\rho D_t\left(\frac{1}{2}v^2 + U\right) - \vec{v}\cdot\vec{f} - \vec{\partial}\cdot\left(\overleftrightarrow{S}\cdot\vec{v}\right) - h + \vec{\partial}\cdot\vec{H}\right]^E(\vec{r},t). \quad (13.6.11)$$

Since $K'(t)$ can be any open subset of $K(t)$ with piecewise smooth boundary, the vanishing integral theorem gives

$$\rho D_t\left(\frac{1}{2}v^2 + U\right) + \vec{\partial}\cdot\vec{H} = h + \vec{v}\cdot\vec{f} + \vec{\partial}\cdot\left(\overleftrightarrow{S}\cdot\vec{v}\right). \quad (13.6.12)$$

This equation can be simplified. We have

$$D_t\left(v^2\right) = D_t\left(\vec{v}\cdot\vec{v}\right) = 2\vec{v}\cdot D_t\vec{v} = 2\vec{v}\cdot\vec{a}.$$

Also, taking components relative to an orthonormal basis,

$$\vec{\partial}\cdot\left(\overleftrightarrow{S}\cdot\vec{v}\right) = \partial_i\left(S_{ij}v_j\right) = \left(\partial_i S_{ij}\right)v_j + S_{ij}\partial_i v_j \tag{13.6.13}$$

$$= \left(\vec{\partial}\cdot\overleftrightarrow{S}\right)\cdot\vec{v} + \overleftrightarrow{S}\langle 2\rangle\vec{\partial}\vec{v}. \tag{13.6.14}$$

Therefore (13.6.12) is

$$\rho\vec{a}\cdot\vec{v} + \rho D_t U + \vec{\partial}\cdot\vec{H} = h + \vec{f}\cdot\vec{v} + \left(\vec{\partial}\cdot\overleftrightarrow{S}\right)\cdot\vec{v} + \overleftrightarrow{S}\langle 2\rangle\vec{\partial}\vec{v}.$$

But the momentum equation says $\rho\vec{a} = \vec{f} + \vec{\partial}\cdot\overleftrightarrow{S}$, so those terms cancel, leaving only

$$\rho D_t U + \vec{\partial}\cdot\vec{H} = h + \overleftrightarrow{S}\langle 2\rangle\vec{\partial}\vec{v}. \tag{13.6.15}$$

This is the Eulerian form of the internal energy equation, or heat equation.

13.6.2 Lagrangian form of energy conservation

The argument should be familiar by now. The total energy of the material with labels in H' is, at time t,

$$\mathcal{E}_t^L\left[H'\right] = \int_{H'} dV_L\left(\vec{x}\right)\left[\tilde{\rho}\left(\frac{1}{2}v^2 + U\right)\right]^L\left(\vec{x}, t\right). \tag{13.6.16}$$

The rate at which energy is added is

$$\mathcal{W}_t^L\left[H'\right] = \int_{H'} dV_L\left(\vec{x}\right)\left[\tilde{\vec{f}}\cdot\vec{v} + \vec{D}\cdot\left(\tilde{\overleftrightarrow{S}}\cdot\vec{v}\right) + \hbar\right]^L\left(\vec{x}, t\right) \tag{13.6.17}$$

$$- \int_{\partial H'} dA_L\left(\vec{x}\right)\ \not{H}\left(\vec{x}, t, \hat{n}_L(\vec{x})\right). \tag{13.6.18}$$

The law of conservation of energy is

$$\frac{d}{dt}\mathcal{E}_t^L\left[H'\right] = \mathcal{W}_t^L\left[H'\right] \tag{13.6.19}$$

and from (13.6.16) we have the identity

$$\frac{d}{dt}\mathcal{E}_t^L\left[H'\right] = \int_{H'} dV_L(\vec{x})\left[\tilde{\rho}D_t\left(\frac{1}{2}v^2 + U\right)\right]^L(\vec{x}, t). \tag{13.6.20}$$

Substituting (13.6.18) and (13.6.20) in (13.6.19) gives

$$\int_{H'} dV_L\left(\vec{x}\right)\left[\tilde{\rho}D_t\left(\frac{1}{2}v^2 + U\right) - \tilde{\vec{f}}\cdot\vec{v} - \vec{D}\cdot\left(\tilde{\overleftrightarrow{S}}\cdot\vec{v}\right) - \hbar\right]^L\left(\vec{x}, t\right)$$

$$= - \int_{\partial H'} dA_L(\vec{x})\ \not{H}\left(\vec{x}, t, \hat{n}_L(\vec{x})\right). \tag{13.6.21}$$

At any $(\vec{x}, t)$, Cauchy's theorem 13.2.28 assures us of the existence of a unique vector $\vec{\not{H}}^L(\vec{x}, t) \in L$ such that for any unit vector $\hat{n} \in L$, $\not{H}(\vec{x}, t, \hat{n}) = \hat{n}\cdot \hat{\not{H}}^L(\vec{x}, t)$. Substituting this on the right in (13.6.21), and applying Gauss's theorem and the vanishing integral theorem gives

$$\tilde{\rho} D_t \left(\frac{1}{2}v^2 + U\right) - \vec{\tilde{f}} \cdot \vec{v} - \vec{D} \cdot \left(\overleftrightarrow{S} \cdot \vec{v}\right)\ \hbar = -\vec{D}\cdot \vec{\not{H}}.$$

But $D_t v^2 = 2\vec{v} \cdot D_t\vec{v} = \partial\vec{v} \cdot \vec{a}$, and

$$\vec{D} \cdot \left(\overleftrightarrow{\tilde{S}} \cdot \vec{v}\right) = \left(\vec{D} \cdot \overleftrightarrow{\tilde{S}}\right) \cdot \vec{v} + \overleftrightarrow{\tilde{S}}\langle 2\rangle \vec{D}\vec{v}$$

so

$$\tilde{\rho}\vec{a} \cdot \vec{v} + \tilde{\rho} D_t U = \vec{\tilde{f}} \cdot \vec{v} - \left(\vec{D} \cdot \overleftrightarrow{\tilde{S}}\right) \cdot \vec{v} - \overleftrightarrow{S}\langle 2\rangle \vec{D}\vec{v} - \hbar = -\vec{D}\cdot \vec{\not{H}}.$$

The Lagrangian momentum equation, $\tilde{\rho}\vec{a} - \vec{\tilde{f}} - \vec{D} \cdot \overleftrightarrow{\tilde{S}} = \vec{0}$, permits some cancellation and leaves us with

$$\tilde{\rho} D_t U + \vec{D}\cdot \vec{\not{H}} = \hbar + \overleftrightarrow{\tilde{S}}\langle 2\rangle \vec{D}\vec{v}. \tag{13.6.22}$$

This is the Lagrangian form of the internal energy equation.

Using (13.3.25) and the identity

$$dA_P \hat{n} \cdot \vec{H} = dA_L \hat{n}_L\cdot \vec{\not{H}},$$

one proves

$$\frac{\vec{H}}{\rho} = \left(\vec{D}\vec{r}\right)^T \cdot \frac{\vec{\not{H}}}{\tilde{\rho}} \tag{13.6.23}$$

in the same manner as (13.3.30) was proved.

13.7 Stuff

Mass, momentum, angular momentum and energy are all examples of a general mathematical object which we call a "stuff." The idea has produced some confusion, so we discuss it here.

Definition 13.7.52 A "stuff" is an ordered pair of functions ($\vec{\boldsymbol{\phi}}^E$, $\overleftrightarrow{\mathbf{F}}^E$) with these properties: V is a Euclidean space, and for each $t \in \mathcal{R}$, $K(t)$ is an open subset of physical space P; and

$$\vec{\boldsymbol{\phi}}^E(\cdot, t) : K(t) \to V \tag{13.7.1}$$

$$\overleftrightarrow{F}^E(\cdot, t) : K(t) \to P \otimes V. \tag{13.7.2}$$

The function $\vec{\boldsymbol{\phi}}^E$ is called the spatial density of the stuff; $\overleftrightarrow{\mathbf{F}}^E$ is the spatial current density or spatial flux density of the stuff. And $\vec{\gamma}^E$, the creation rate of the stuff, is defined to be

$$\vec{\gamma}^E = \partial_t \vec{\boldsymbol{\phi}}^E + \vec{\partial} \cdot \overleftrightarrow{\mathbf{F}}^E . \tag{13.7.3}$$

Definition 13.7.53 Suppose $t \in \mathcal{R}$, $\vec{r} \in K(t)$, and $dV(\vec{r})$ is a small open subset of $K(t)$ containing $\vec{r}$; $dV(\vec{r})$ will also denote the volume of that subset. Suppose $dA(\vec{r}, t)$ is a small nearly plane oriented surface containing $\vec{r}$, and $\hat{n}$ is the unit normal on the front side of $dA(\vec{r}, t)$. We also use $dA(\vec{r}, t)$ to denote the area of the small surface. Suppose $dA(\vec{r}, t)$ moves with speed W normal to itself, with $W > 0$ when the motion is in the direction of $\hat{n}$. Then

i) $dV(\vec{r}) \vec{\boldsymbol{\phi}}^E(\vec{r}, t)$ is called the amount of stuff in $dV(\vec{r})$ at time t;

ii) $dA(\vec{r}, t)[\hat{n} \cdot \overleftrightarrow{\mathbf{F}}^E(\vec{r}, t) - W \vec{\boldsymbol{\phi}}^E(\vec{r}, t)]$ is called the rate at which the stuff flows across $[dA(\vec{r}, t), \hat{n}]$ from *back* to *front*.

iii) $dV(\vec{r})\vec{\gamma}^E(\vec{r}, t)$ is called the rate at which stuff is created in $dV(\vec{r})$ at time t.

Definition 13.7.54 Let $K'(t)$ be any open subset of $K(t)$, with piecewise continuous boundary $\partial K'(t)$ and unit outward normal $\hat{n}(\vec{r})$ at $\vec{r} \in \partial K'(t)$. Suppose $K'(t)$ moves so that the outward velocity of $\partial K'(t)$ normal to itself at $\vec{r} \in \partial K'(t)$ is $W(\vec{r}, t)$. Then

i) $\int_{K'(t)} dV(\vec{r}) \vec{\boldsymbol{\phi}}^E(\vec{r}, t)$ is called the amount of stuff in $K'(t)$ at time t. Denote if by $\vec{\Phi}[K'(t)]$

ii) $\int_{\partial K'(t)} dA(\vec{r}) \left[\hat{n}(\vec{r}) \cdot \overleftrightarrow{\mathbf{F}}^E(\vec{r}, t) - W(\vec{r}, t) \vec{\boldsymbol{\phi}}^E(\vec{r}, t)\right]$ is called the rate at which stuff flows out of $K'(t)$ across $\partial K'(t)$ at time t. Denote it by $\vec{\mathcal{F}}[K'(t)]$.

iii) $\int_{K'(t)} dV(\vec{r})\vec{\gamma}^E(\vec{r},t)$ is called the rate at which the stuff is created in $K'(t)$ at time t. Denote it by $\vec{\Gamma}[K'(t)]$.

Remark 13.7.68 $d/dt\,\vec{\Phi}[K'(t)] = \vec{\Gamma}[K'(t)] - \vec{\mathcal{F}}[K'(t)]$ for any open subset $K'(t)$ of $K(t)$, moving in any way whatever.

Proof:

$$\frac{d}{dt}\vec{\Phi}\,[K'(t)] = \int_{K'(t)} dV(\vec{r})\partial_t \vec{\boldsymbol{\phi}}^{\,E}(\vec{r},t) + \int_{\partial K'(t)} dA(\vec{r})W\,\vec{\boldsymbol{\phi}}^{\,E}(\vec{r},t)$$

$$\vec{\mathcal{F}}\,[K'(t)] = \int_{K'(t)} dV(\vec{r})\vec{\partial}\cdot \overleftrightarrow{\mathbf{F}}^{E}(\vec{r},t) - \int_{\partial K'(t)} dA(\vec{r})W\,\vec{\boldsymbol{\phi}}^{\,E}(\vec{r},t).$$

Therefore,

$$\begin{aligned}\frac{d}{dt}\vec{\Phi}\,[K'(t)] + \vec{\mathcal{F}}\,[K'(t)] &= \int_{K'(t)} dV(\vec{r})\left(\partial_t \vec{\boldsymbol{\phi}}^{\,E} + \vec{\partial}\cdot \overleftrightarrow{\mathbf{F}}^{E}\right) \\ &= \int_{K'(t)} dV(\vec{r})\vec{\gamma}^E(\vec{r},t) = \vec{\Gamma}\,[K'(t)]\,.\end{aligned}$$

Definition 13.7.55 Suppose the $K(t)$ in definition 13.7.52 is the set of positions occupied by the particles of a certain continuum at time t. Suppose ρ and $\vec{v}$ are the density of mass and the particle velocity in the contiuum. Then define

i) $\vec{\phi}^E(\vec{r},t) = \vec{\boldsymbol{\phi}}^{\,E}(\vec{r},t)/\rho^E(\vec{r},t)$

ii) $\overleftrightarrow{F}^{E}(\vec{r},t) = \overleftrightarrow{\mathbf{F}}^{E}(\vec{r},t) - \vec{v}^E(\vec{r},t)\,\vec{\boldsymbol{\phi}}^{\,E}(\vec{r},t)$.

Let $\vec{\phi}$, $\vec{\boldsymbol{\phi}}$, $\overleftrightarrow{F}$, $\overleftrightarrow{\mathbf{F}}$ be, respectively the physical quantities whose Eulerian descriptions are $\vec{\phi}^E$, $\vec{\boldsymbol{\phi}}^{\,E}$, $\overleftrightarrow{F}^{E}$, $\overleftrightarrow{\mathbf{F}}^{E}$. Then the definitions of $\vec{\phi}^E$ and $\overleftrightarrow{F}^{E}$ imply

$$\vec{\phi} = \vec{\boldsymbol{\phi}}/\rho \text{ or } \vec{\boldsymbol{\phi}} = \rho\vec{\phi} \tag{13.7.4}$$

$$\vec{\boldsymbol{\phi}} = \overleftrightarrow{\mathbf{F}} - \vec{v}\,\vec{\boldsymbol{\phi}} \text{ or } \overleftrightarrow{\mathbf{F}} = \overleftrightarrow{F} + \rho\vec{v}\vec{\phi}. \tag{13.7.5}$$

The physical quantity $\vec{\phi}$ is called the density of stuff per unit mass of continuum material ("material density of the stuff," for short), while $\vec{\boldsymbol{\phi}}$ is the density of stuff per unit volume of physical space. The physical quantity $\overleftrightarrow{F}$ is the flux density of the stuff relative to the material in the continuum, or the material flux density, while $\overleftrightarrow{\mathbf{F}}$ is the flux density of the stuff in space.

Remark 13.7.69 Suppose the two pairs of functions $(\vec{\boldsymbol{\phi}}, \overleftrightarrow{\mathbf{F}})$ and $(\phi, \vec{F})$ are related by (13.7.4). Then

$$\partial_t \vec{\boldsymbol{\phi}} + \vec{\partial}\cdot \overleftrightarrow{\mathbf{F}} = \rho D_t \vec{\phi} + \vec{\partial}\cdot \overleftrightarrow{F} \tag{13.7.6}$$

so the creation rate of the stuff $(\vec{\boldsymbol{\phi}}, \overleftrightarrow{\mathbf{F}})$ is

$$\vec{\gamma} = \rho D_t \vec{\phi} + \vec{\partial}\cdot \overleftrightarrow{F}\ . \tag{13.7.7}$$

Proof:

$$\begin{aligned}
\partial_t \vec{\boldsymbol{\phi}} &= (\partial_t \rho)\,\vec{\phi} + \rho\left(\partial_t \vec{\phi}\right) \\
\vec{\partial}\cdot \vec{\mathbf{F}} &= \vec{\partial}\cdot \overleftrightarrow{F} + \left[\vec{\partial}\cdot(\rho\vec{v})\right]\vec{\phi} + (\rho\vec{v})\cdot\vec{\partial}\vec{\phi} \\
\partial_t \boldsymbol{\phi} + \vec{\partial}\cdot\vec{\mathbf{F}} &= \left[\partial_t\rho + \vec{\partial}\cdot(\rho\vec{v})\right]\vec{\phi} + \rho\left[\partial_t\vec{\phi} + \vec{v}\cdot\left(\vec{\partial}\vec{\phi}\right)\right] + \vec{\partial}\cdot \overleftrightarrow{F} \\
&= 0\vec{\phi} + \rho D_t\vec{\phi} + \vec{\partial}\cdot \overleftrightarrow{F}\ .
\end{aligned}$$

QED.

Stuffs can be added if their densities and fluxes have the same domain and same range. We have

Definition 13.7.56 Suppose $(\vec{\boldsymbol{\phi}}_\nu, \overleftrightarrow{\mathbf{F}}_\nu)$ for each $v = 1, \cdots, N$ is a stuff, and that all these stuffs have the same $K(t)$ and V in definition (13.7.1). Suppose $c_1, \cdots, c_N \in \mathcal{R}$. Then $\sum_{\nu=1}^N c_\nu(\vec{\boldsymbol{\phi}}_\nu, \overleftrightarrow{\mathbf{F}}_\nu)$ stands for the stuff $(\sum_{\nu=1}^N c_\nu \vec{\boldsymbol{\phi}}_\nu), (\sum_{\nu=1}^N c_\nu \overleftrightarrow{\mathbf{F}}_\nu)$.

Corollary 13.7.48 The stuff $\sum_{\nu=1}^N c_\nu(\vec{\boldsymbol{\phi}}_\nu, \overleftrightarrow{\mathbf{F}}_\nu)$ has creation rate $\sum_{\nu=1}^N c_\nu \gamma_\nu$.

Proof:

Obvious from (13.7.3).

Corollary 13.7.49 In a continuum with particle velocity $\vec{v}$ and density ρ, the material density and material flux density of the stuff $\sum_{\nu=1}^N c_\nu(\boldsymbol{\phi}_\nu, \vec{\mathbf{F}}_\nu)$ are, respectively $\sum_{\nu=1}^N c_\nu \phi_\nu$ and $\sum_{\nu=1}^N c_\nu \vec{F} v$.

Proof:

Obvious from (13.7.4) and (13.7.5).

Now we consider a number of stuffs which arise out of the conservation laws. The name of each stuff is capitalized and underlined, and then its spatial density $\vec{\boldsymbol{\phi}}$, spatial flux $\overleftrightarrow{\mathbf{F}}$, material density $\vec{\phi}$, material flux $\overleftrightarrow{F}$, and creation rate $\vec{\gamma}$ are given. The stuff is not defined by $\vec{\boldsymbol{\phi}}$ or $\vec{\phi}$ alone. Both $\vec{\boldsymbol{\phi}}$ and $\overleftrightarrow{\mathbf{F}}$, or both $\vec{\phi}$ and $\overleftrightarrow{F}$ must be given. Then γ is calculated from (13.7.3). The choice of $\overleftrightarrow{F}$ is dictated by convenience in applications.

Mass: $\boldsymbol{\phi} = \rho$, $\vec{\mathbf{F}} = \rho\vec{v}$, $\phi = 1$, $\vec{\mathbf{F}} = \vec{0}$, $\gamma = \partial_t \boldsymbol{\phi} + \vec{\partial} \cdot \vec{\mathbf{F}} = \partial_t\rho + \vec{\partial} \cdot (\rho\vec{v}) = 0$. Mass is not created, and its material flux is zero.

Momentum: $\vec{\boldsymbol{\phi}} = \rho\vec{v}$, $\overleftrightarrow{\mathbf{F}} = \rho\vec{v}\vec{v} - \overleftrightarrow{S}$, $\vec{\phi} = \vec{v}$, $\overleftrightarrow{F} = -\overleftrightarrow{S}$, $\vec{\gamma} = \rho D_t\phi + \vec{\partial}\cdot\vec{F} = \rho D_t\vec{v} - \vec{\partial}\cdot \overleftrightarrow{S} = \vec{f}$.

Intrinsic Angular Momentum: $\vec{\boldsymbol{\phi}} = \rho\vec{l}$, $\overleftrightarrow{\mathbf{F}} = \rho\vec{v}\vec{l} - \overleftrightarrow{M}$, $\vec{\phi} = \vec{l}$, $\overleftrightarrow{F} = -\overleftrightarrow{M}$, so $\gamma = \rho D_t\vec{l} - \vec{\partial}\cdot \overleftrightarrow{M} = \vec{m} + A_P\langle 2\rangle \overleftrightarrow{S}$.

Kinetic Angular Momentum:

$$
\begin{aligned}
\vec{\boldsymbol{\phi}} &= \rho(\vec{\mathbf{r}} \times \vec{v}), \quad \overleftrightarrow{\mathbf{F}} = \rho\vec{v}(\vec{\mathbf{r}} + \vec{v}) + \overleftrightarrow{S} \times \vec{\mathbf{r}}, \\
\vec{\phi} &= \vec{\mathbf{r}} \times \vec{v}, \quad \overleftrightarrow{F} = \overleftrightarrow{S} \times \vec{\mathbf{r}} \\
\vec{\gamma} &= \rho D_t (\vec{\mathbf{r}} \times \vec{v}) + \vec{\partial} \cdot \left(\overleftrightarrow{S} \times \vec{\mathbf{r}}\right) = \rho\vec{\mathbf{r}} \times \vec{a} - \vec{\mathbf{r}} \times \left(\vec{\partial}\cdot \overleftrightarrow{S}\right) - A_P\langle 2\rangle \overleftrightarrow{S} \\
\vec{\gamma} &= \vec{\mathbf{r}} \times \vec{f} - A_P\langle 2\rangle \overleftrightarrow{S} .
\end{aligned}
$$

Total Angular Momentum:

$$
\begin{aligned}
\vec{\boldsymbol{\phi}} &= \rho\left(\vec{l} + \vec{\mathbf{r}} \times \vec{v}\right), \quad \overleftrightarrow{\mathbf{F}} = \rho\vec{v}\left(\vec{l} + \vec{\mathbf{r}} \times \vec{v}\right) - \overleftrightarrow{M} + \overleftrightarrow{S} \times \vec{\mathbf{r}} \\
\phi &= \vec{l} + \vec{\mathbf{r}} \times \vec{v}, \quad \overleftrightarrow{F} = -\overleftrightarrow{M} + \overleftrightarrow{S} \times \vec{\mathbf{r}}.
\end{aligned}
$$

Since total angular momentum is the sum of intrinsic and kinetic angular momentum in the sense of definition (13.7.4), its creation rate is the sum of their creation rates, by corollary (13.7.48). Thus

$$\vec{\gamma} = \vec{m} + \vec{\mathbf{r}} \times \vec{f}.$$

Internal Energy:

$$\begin{aligned} \boldsymbol{\phi} &= \rho U, \quad \vec{\mathbf{F}} = \rho\vec{v}U + \vec{H} \\ \phi &= U, \quad \vec{F} = \vec{H} \\ \gamma &= \rho D_t U + \vec{\partial}\cdot\vec{H} \end{aligned}$$

so

$$\gamma = h + \overleftrightarrow{S}\langle 2\rangle\vec{\partial}\vec{v}.$$

An alternative definition of internal energy suggests itself if h is entirely h_{nuc}, the radioactive heating rate. Let $\phi^L_{nuc}(\vec{x},t)$ be the nuclear energy in joules per kilogram of mass near particle $\vec{x}$. As the radioactive nuclei near $\vec{x}$ decay, they emit γ-rays and massive particles. We will assume that these lose most of their energy by collision with molecules so close to $\vec{x}$ that the macroscopic properties of the continuum are nearly constant over the region where the collisions occur. Then kinetic energy of molecular motion is added to the material near $\vec{x}$ at the rate $-D_t\phi^L_{nuc}(\vec{x},t)$ watts/kg. To convert to watts/meter^3, we must multiply by the density, so $h^L_{nuc}(\vec{x},t) = -\rho^L(\vec{x},t)D_t\phi^L_{nuc}(\vec{x},t)$ or

$$h_{nuc} = -\rho D_t\phi_{nuc}. \tag{13.7.8}$$

We can regard $U + \phi_{nuc}$ as total internal energy density, and define a stuff called TOTAL INTERNAL ENERGY with

$$\begin{aligned} \boldsymbol{\phi} &= \rho\left(U + \phi_{nuc}\right), \qquad & \vec{\mathbf{F}} &= \rho\vec{v}\left(U + \phi_{nuc}\right) + \vec{H} \\ \phi &= U + \phi_{nuc}, & \vec{F} &= \vec{H}. \end{aligned}$$

The creation rate of this energy is $\gamma = \rho D_t\phi + \vec{\partial}\cdot\vec{F} = \rho D_t U + \vec{\partial}\cdot\vec{H} + \rho D_t\phi_{nuc}$ $= h + \overleftrightarrow{S}\langle 2\rangle\vec{\partial}\vec{v} - h_{nuc}$. If $h = h_{nuc}$, then $\gamma = \overleftrightarrow{S}\langle 2\rangle\vec{\partial}\vec{v}$.

This "total internal energy" is not very useful when the value of ϕ_{nuc} is unaffected by what happens to the material. That is the situation at the low temperatures and pressures inside the earth (with one possible small exception; see Stacey, Physics of the Earth, p. 28). In such situations, h_{nuc}/ρ is a property of the material which is known at any particle $\vec{x}$ at any time t, independent of the motion of the continuum. By contrast, the molecular U is not known at $\vec{x}, t$ until we have solved the problem of finding how the continuum moves. In the deep interiors of stars, pressures and temperatures are high enough to affect nuclear reaction rates, so it is useful to include ϕ_{nuc} with U.

Kinetic Energy:

$$\begin{aligned} \boldsymbol{\phi} &= \frac{1}{2}\rho v^2, \qquad \vec{\mathbf{F}} = \rho\vec{v}\left(\frac{1}{2}v^2\right) - \overleftrightarrow{S} \cdot \vec{v} \\ \phi &= \frac{1}{2}v^2, \qquad \vec{F} = -\overleftrightarrow{S} \cdot \vec{v} \end{aligned}$$

$$\begin{aligned} \gamma &= \rho D_t\phi + \vec{\partial}\cdot\vec{F} = \frac{1}{2}\rho D_t v^2 - \vec{\partial}\cdot\left(\overleftrightarrow{S}\cdot\vec{v}\right) \\ &= \rho\vec{v}\cdot D_t\vec{v} - \partial_i\left(S_{ij}v_j\right) \text{ relative to an orthonormal basis for } P \\ &= \rho\vec{v}\cdot\vec{a} - \left(\partial_i S_{ij}\right)v_j - S_{ij}\partial_i v_j \\ &= \left(\rho\vec{a} - \vec{\partial}\cdot\overleftrightarrow{S}\right)\cdot\vec{v} - \overleftrightarrow{S}\langle 2\rangle\vec{\partial}\vec{v} \\ \gamma &= \vec{f}\cdot\vec{v} - \overleftrightarrow{S}\langle 2\rangle\vec{\partial}\vec{v}. \end{aligned}$$

IK Energy: (Internal plus Kinetic Energy).

This is the sum of internal energy and kinetic energy in the sense of definition 13.7.56, so it has

$$\begin{aligned} &\boldsymbol{\phi} = \rho\left(\tfrac{1}{2}v^2 + U\right), \qquad \vec{\mathbf{F}} = \rho\vec{v}\left(\tfrac{1}{2}v^2 + U\right) + \vec{H} - \overleftrightarrow{S}\cdot\vec{v} \\ &\phi = \tfrac{1}{2}v^2 + U, \qquad \vec{F} = \vec{H} - \overleftrightarrow{S}\cdot\vec{v} \\ &\gamma = h + \vec{f}\cdot\vec{v}. \end{aligned}$$

Potential Energy: Suppose the body force density $\vec{f}$ is given by $\vec{f} = -\rho\vec{\partial}\psi$ where the potential function $\psi^E(\vec{r}, t)$ is independent of time, i.e., $\partial_t\psi = 0$. Then $D_t\psi = \partial_t\psi + \vec{v}\cdot\vec{\partial}\psi = \vec{v}\cdot\vec{\partial}\psi$ so $\rho D_t\psi = -\vec{f}\cdot\vec{v}$.

Then the stuff "potential energy" is defined by

$$\begin{aligned} \phi &= \psi, \quad \vec{F} = \vec{0} \\ \boldsymbol{\phi} &= \rho\psi, \quad \vec{\mathbf{F}} = \rho\vec{v}\psi \end{aligned}$$

so its creation rate is $\gamma = \rho D_t \psi = -\vec{f} \cdot \vec{v} \cdot \gamma$.

IKP Energy: (Internal plus kinetic plus potential energy).

If $\vec{f} = -\rho\vec{\partial}\psi$ with $\partial_t \psi = 0$, then we define the stuff "IKP energy" as the sum of internal, kinetic and potential energies in the sense of definition (13.7.56). Then

$$\begin{aligned} \phi &= \frac{1}{2}v^2 + U + \psi, \quad \vec{F} = \vec{H} - \overleftrightarrow{S} \cdot \vec{v} \\ \boldsymbol{\phi} &= \rho\left(\frac{1}{2}v^2 + U + \psi\right), \quad \mathbf{F} = \rho\vec{v}\left(\frac{1}{2}v^2 + U + \psi\right) + \vec{H} - \overleftrightarrow{S} \cdot \vec{v}. \end{aligned}$$

By corollary 13.7.48, γ is the sum of the three γ's for kinetic, internal and potential energy, so $\gamma = h$. If we include nuclear potential energy in U and neglect other sources of h (e.g. solar heating), then $\gamma = 0$. Usually, this is not done, and IKP energy is called "total energy."

13.7.1 Boundary conditions

So far we have dealt only with physical quantities which were continuously differentiable to the extent needed to justify using Gauss's theorem and the vanishing integral theorem. Now we analyze what happens when there is a surface S in the continuum, across which physical quantities may have jump discontinuities. That is, $\lim \phi^E(\vec{r}, t)$ exists as $\vec{r}$ approaches $\vec{r}_0 \in S$ from either side of S, but the two limits are different.

Examples of such surfaces are the interface between the ocean and atmosphere, the ocean bottom, the core-mantle boundary, and a shock wave in air, water or rock. (Shock waves in rock are excited only by nuclear blasts and meteor impacts. Earthquakes are too weak to excite them.)

We will assume that it is possible to choose a front side of S (Möbius strips are excluded) and that we have done so. We will denote the unit normal to S on its front, or positive, side by $\hat{\nu}$, and we will call its back side the negative side. S is a large surface, not nearly plane, but piecewise smooth, so $\hat{\nu}$ can vary with position $\vec{r} \in S$. Moreover, we will permit S to move, so that $S(t)$ may vary with time t. Then so will $\hat{\nu}(\vec{r}, t)$, the normal to $S(t)$ at $\vec{r} \in S(t)$. We will write $W(\vec{r}, t)$ for the velocity with which $S(t)$ moves

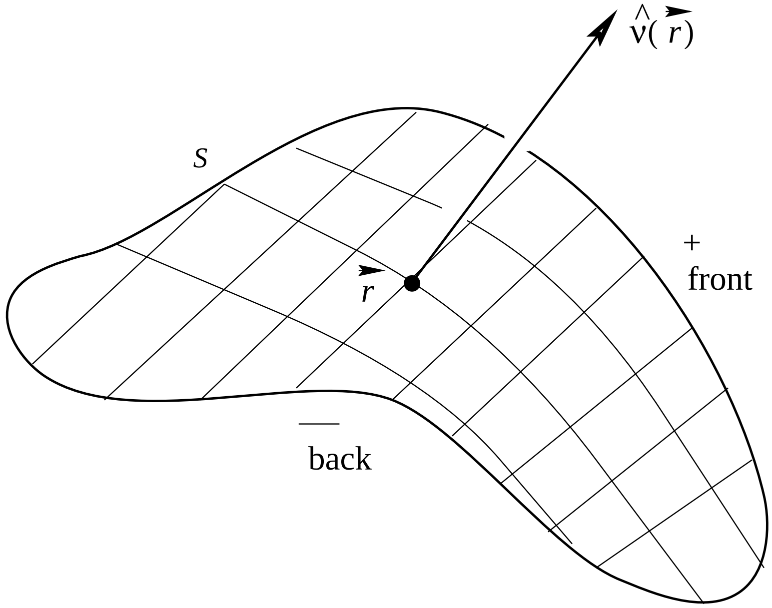


Figure 13.17:

normal to itself at point $\vec{r} \in S(t)$, and we take $W > 0$ if the motion is in the direction of $\hat{\nu}$, $W < 0$ if opposite.

If the surface is given in physical space P, we write it as $S^E(t)$. If it is given in label space L, we write it as $S^L(t)$. If ϕ is any physical quantity with a jump discontinuity across $S(t)$, we write $\phi^E(\vec{r}_0, t)^+$ for the limiting value of $\phi^E(\vec{r}, t)$ as $\vec{r} \to \vec{r}_o \in S^E(t)$ from in front of $S^E(t)$. We write $\phi^E(\vec{r}_0, t)^-$ for the limit of $\phi^E(\vec{r}, t)$ as $\vec{r} \to \vec{r}_0$ from behind $S^E(t)$. And we define

$$\left[\phi^E\left(\vec{r}_0, t\right)\right]_-^+ = \phi^E(\vec{r}_0, t)^+ - \phi^E(\vec{r}_0, t)^-. \tag{13.7.9}$$

This quantity is called the jump in ϕ^E across $S^E(t)$. We introduce a similar notation for label space, with

$$\left[\phi^L\left(\vec{x}_0, t\right)\right]_-^+ = \phi^L\left(\vec{x}_0, t\right)^+ - \phi^L\left(\vec{x}_0, t\right)^-.$$

We will need the surface version of the vanishing integral theorem. This says that if V is a Euclidean vector space, S is a surface in P or L, and $\vec{f} : S \to V$ is continuous, and $\int_{S'} dA(\vec{r})\vec{f}(\vec{r}) = \vec{O}_V$ for every open patch S' on S, then $\vec{f}(\vec{r}) = \vec{O}_V$ for all $\vec{r} \in S$. Here an "open patch" on S is any set $S' = K' \cap S$, where K' is an open subset of P or L. The proof of the surface version of the vanishing integral theorem is essentially the same as the proof of the original theorem 11.5.26, so we omit it.

We will consider only the Eulerian forms of the conservation laws, so we will omit the superscript ^{E}. The reader can easily work out the Lagrangian forms of the law.

We use the notation introduced on page 163, but now we assume that the surface $S(t)$ passes through $K'(t)$. Therefore, we must augment the notation. We use $K'(t)^+$ for the part (excluding $S(t)$) of $K'(t)$ which is in front of $S(t)$, $\partial^+ K'(t)$ for the part of $\partial K'(t)$ in front of $S(t)$, and $\partial K'(t)^+$ for the whole boundary of $K'(t)^+$, including $K'(t) \cap S(t)$. If ϕ is a function defined and continuous on the closed set $\overline{K'(t)} = K'(t) \cup \partial K'(t)$, except for a jump discontinuity on $S(t)$, we write ϕ^+ for the continuous function on $\overline{K'(t)^+}$ defined as ϕ except on $K'(t) \cap S(t)$, where its value is the ϕ^+ of equation (13.7.9). We label objects behind $S(t)$ in the same way, except that $^+$ is replaced by $^-$. The unit outward normal to $\partial K'(t)^+$ will be written $\hat{n}^+$. The unit outward normal to $\partial K'(t)^-$ will be written $\hat{n}^-$. The unit outward normal to $\partial K'(t)$ will be written $\hat{n}$. Then $\hat{n}^+ = \hat{n}$ on $\partial^+ K'(t)$ and $\hat{n}^+ = -\hat{\nu}$ on $K'(t) \cap S(t)$. And $\hat{n}^- = \hat{n}$ on $\partial^- K'(t)$, while $\hat{n}^- = \hat{\nu}$ on $K'(t) \cap S(t)$. As on page 163, we choose $K'(t)$ so that it always consists of the same particles. Then the outward normal velocity of $\partial K'(t)^+$ relative to itself is $\hat{n} \cdot \vec{v}$ on $\partial^+ K'(t)$ and $-W$ on $K'(t) \cap S(t)$. Here $\vec{v}$ is the particle velocity in the continuum. The outward normal velocity of $\partial K'(t)^-$ relative to itself is $\hat{n} \cdot \vec{v}$ on $\partial^- K'(t)$ and W on $K'(t) \cap S(t)$.

13.7.2 Mass conservation

Let $M[K'(t)]$ denote the mass of the material in $K'(t)$. Since $K'(t)$ always consists of the same particles, the physical law of mass conservation is

$$\frac{d}{dt} M\left[K'(t)\right] = 0. \tag{13.7.10}$$

We need a mathematical identity like (13.1.9). We begin with

$$M\left[K'(t)\right] = \int_{K'(t)} dV(\vec{r}) \rho\left(\vec{r}, t\right). \tag{13.7.11}$$

This looks obvious until we recognize that there are physical situations which we might want to model by placing a mass per unit area on $S(t)$, which would have to be added

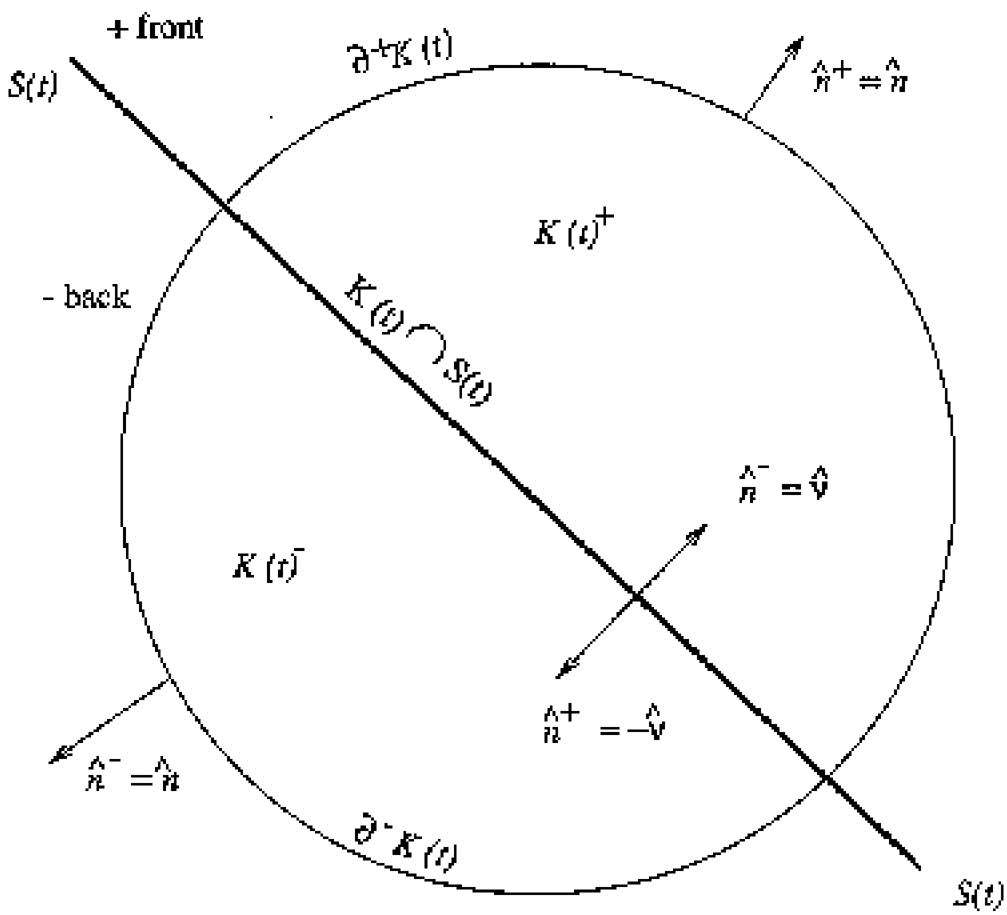

+ front
$\partial^+K(t)$
$S(t)$
$\hat{n}^+ = \hat{n}$
$K(t)^+$
- back
$K(t) \cap S(t)$
$\hat{n}^- = \hat{\nu}$
$K(t)^-$
$\hat{n}^+ = -\hat{\nu}$
$\hat{n}^- = \hat{n}$
$\partial^- K(t)$
$S(t)$

Figure 13.18:

to (13.7.11). Examples are $S(t) =$ the soap film in a soap bubble in air, or $S(t) =$ the ice pack on a polar sea. We ignore these possibilities and assume

$$\text{mass on } S(t) = 0. \tag{13.7.12}$$

Next, we try to differentiate (13.7.10) using (11.6.1). The jump discontinuity on $S(t) \cap K'(t)$ prevents this, so we break (13.7.11) into

$$M\left[K'(t)\right] = M\left[K'(t)^+\right] + M\left[K'(t)^-\right] \text{ with} \tag{13.7.13}$$

$$M\left[K'(t)^+\right] = \int_{K'(t)^+} dV(\vec{r})\rho^+(\vec{r},t) \tag{13.7.14}$$

$$M\left[K'(t)^-\right] = \int_{K'(t)^-} dV(\vec{r})\rho^-(\vec{r},t). \tag{13.7.15}$$

The two integrands in (13.7.14) and (13.7.15) are continuous, so we can apply (11.6.1) to obtain

$$\frac{d}{dt}M\left[K'(t)^+\right] = \int_{K'(t)^+} dV(\vec{r})\,\partial_t\rho^+ + \int_{\partial^+K'(t)} dA\,\hat{n}\cdot\vec{v}^+\rho^+ - \int_{K'(t)\cap S(t)} dA\,W\rho^+. \tag{13.7.16}$$

We would like to use Gauss's theorem on the right in (13.7.16), but $\partial^+K'(t)$ is not a closed surface.

To close it we must add $K'(t) \cap S(t)$. Then

$$\begin{aligned}
\int_{\partial^+K'(t)} dA\hat{n}\cdot(\rho\vec{v})^+ &= \int_{\partial K'(t)^+} dA\,\hat{n}^+\cdot(\rho\vec{v})^+ - \int_{K'(t)\cap S(t)} dA\,\hat{n}^+\cdot(\rho\vec{v})^+ \\
&= \int_{K'(t)^+} dV\,\vec{\partial}\cdot(\rho\vec{v})^+ + \int_{K'(t)\cap S(t)} dA\,\hat{\nu}\cdot(\rho\vec{v})^+
\end{aligned}$$

Using this in (13.7.16) gives

$$\frac{d}{dt}M\left[K'(t)^+\right] = \int_{K'(t)^+} dV\left[\partial_t\rho + \vec{\partial}\cdot(\rho\vec{v})\right]^+ + \int_{K'(t)\cap S(t)} dA\left[\rho\left(\hat{\nu}\cdot\hat{v} - W\right)\right]^+.$$

But $[\partial_t\rho + \vec{\partial}\cdot(\rho\vec{v})]^+$ in $K'(t)^+$ is just the value of $\partial_t\rho + \vec{\partial}\cdot(\rho\vec{v})$ there, so it vanishes, and

$$\frac{d}{dt}M\left[K'(t)^+\right] = \int_{K'(t)\cap S(t)} dA\left[\rho(\hat{\nu}\cdot\vec{v} - W)\right]^+. \tag{13.7.17}$$

By an exactly similar argument,

$$\frac{d}{dt}M\left[K'(t)^-\right] = \int_{K'(t)\cap S(t)} dA\left[\rho(\hat{\nu}\cdot\vec{v} - W)\right]^-. \tag{13.7.18}$$

Therefore,

$$\frac{d}{dt}M\left[K'(t)\right] = \int_{K'(t)\cap S(t)} dA\left[\rho(\hat{\nu}\cdot\vec{v} - W)\right]_{-}^{+}. \tag{13.7.19}$$

Combining (13.7.10) and (13.7.19) gives for $K' = K'(t)$

$$\int_{K'(t)\cap S(t)} dA\left[\rho(\hat{\nu}\cdot\vec{v} - W)\right]_{-}^{+} = 0. \tag{13.7.20}$$

Since K' can be any open subset of $K(t)$, $K'\cap S(t)$ can be any open patch on $S(t)$. Therefore, by the surface version of the vanishing integral theorem

$$\left[\rho\left(\hat{\nu}\cdot\vec{v} - W\right)\right]_{-}^{+} = 0 \qquad \text{on } S(t). \tag{13.7.21}$$

By analogy with definition 13.7.53 ii), we call either $[\rho(\hat{\nu}\cdot\vec{v}-W)]^{+}$ or $[\rho(\hat{\nu}\cdot\vec{v}-W)]^{-}$ the flux of mass across $S(t)$, and write it $m(\vec{r},t)$ since it depends on $\vec{r}$ and t. Thus

$$m = \left[\rho(\hat{\nu}\vec{v} - W)\right]^{+} = \left[\rho(\hat{\nu}\vec{v} - W)\right]^{-}. \tag{13.7.22}$$

If $m = 0$, $S(t)$ is a boundary between materials. It is then called a contact surface. It is not crossed by particles. Then

$$\hat{\nu}\cdot\vec{v}^{\,+} = \hat{\nu}\cdot\hat{\nu}^{-} = W. \tag{13.7.23}$$

If $m \neq 0$, $S(t)$ is a shock wave or shock front.

13.7.3 Momentum conservation

Again refer to figure 13.18. Let $\vec{P}[K'(t)]$ be the momentum in $K'(t)$, while $\vec{\mathcal{F}}[K'(t)]$ is the force on that material. Then mometum conservation requires

$$\frac{d}{dt}\vec{P}\left[K'(t)\right] = \vec{\mathcal{F}}\left[K'(t)\right]. \tag{13.7.24}$$

We must find mathematical identities for the two sides of (13.7.24). Because of (13.7.12), we have

$$\vec{P}\left[K'(t)\right] = \int_{K'(t)} dV\rho\vec{v}. \tag{13.7.25}$$

Again we cannot appeal directly to (11.6.1) to differentiate (13.7.25), because of the jump discontinuity on $S(t)$. Therefore we imitate the procedure beginning on page 224. We have

$$\vec{P}\left[K'(t)\right] = \vec{P}\left[K'(t)^{+}\right] + \vec{P}\left[K'(t)^{-}\right] \tag{13.7.26}$$

where

$$\vec{P}\left[K'(t)^{+}\right] = \int_{K'(t)^{+}} dV (\rho\vec{v})^{+} \tag{13.7.27}$$

$$\vec{P}\left[K'(t)^{-}\right] = \int_{K'(t)^{-}} dV (\rho\vec{v})^{-}. \tag{13.7.28}$$

Then, appealing to (11.6.1), we have

$$\begin{aligned} \frac{d}{dt}\vec{P}\left[K'(t)^{+}\right] &= \int_{K'(t)^{+}} dV\, \partial_t(\rho\vec{v})^{+} \\ &\quad + \int_{\partial^{+}K'(t)} dA\, \hat{n}\cdot(\vec{v}\rho\vec{v})^{+} \\ &\quad - \int_{K'(t)\cap S(t)} dA\, W\, (\rho\vec{v})^{+} \end{aligned}$$

and

$$\begin{aligned} \int_{\partial^{+}K'(t)} dA\, \hat{n}\cdot(\vec{v}\rho\vec{v})^{+} &= \int_{\partial K'(t)^{+}} dA\, \hat{n}^{+}\cdot(\rho\vec{v}\vec{v})^{+} \\ &\quad - \int_{K'(t)\cap S(t)} dA\, \hat{n}^{+}\cdot(\rho\vec{v}\vec{v})^{+} \\ &= \int_{K'(t)^{+}} dV\, \vec{\partial}\cdot(\rho\vec{v}\vec{v})^{+} \\ &\quad + \int_{K'(T)\cap S(t)} dA\, \hat{\nu}\cdot(\rho\vec{v}\vec{v})^{+}. \end{aligned}$$

Thus

$$\begin{aligned} \frac{d}{dt}\vec{P}\left[K'(t)^{+}\right] &= \int_{K'(t)^{+}} dV \left[\partial_t\left(\rho\vec{v}\right) + \vec{\partial}\cdot\left(\rho\vec{v}\vec{v}\right)\right]^{+} \\ &\quad + \int_{K'(t)\cap S(t)} dA \left[\left(\hat{\nu}\cdot\vec{v} - W\right)\rho\vec{v}\right]^{+}. \end{aligned}$$

Now

$$\begin{aligned} \partial_t(\rho\vec{v}) + \vec{\partial}\cdot(\rho\vec{v}\vec{v}) &= \left[\partial_t\rho + \vec{\partial}\cdot(\rho\vec{v})\right]\vec{v} + \rho\left[\partial_t\vec{v} + \vec{v}\cdot\vec{\partial}\vec{v}\right] \\ &= \rho D_t\vec{v} = \rho\vec{a} = \vec{\partial}\cdot\overleftrightarrow{S} + \vec{f} \end{aligned}$$

so, using Gauss's theorem

$$\begin{aligned}
\frac{d}{dt}\vec{P}\left[K'(t)^{+}\right] &= \int_{\partial K'(t)^{+}} dA\, \hat{n}^{+}\cdot \overset{\leftrightarrow}{S}^{+} \\
&\quad + \int_{K'(t)^{+}} dV\, \vec{f} + \int_{K'(t)\cap S(t)} dA\left[\left(\hat{\nu}\cdot\vec{v} - W\right)\rho\vec{v}\right]^{+} \\
&= \int_{K'(t)^{+}} dV\, \vec{f} + \int_{\partial^{+}K'(t)} dA\, \hat{n}\cdot \overset{\leftrightarrow}{S} \\
&\quad + \int_{K'(t)\cap S(t)} dA\left[-\hat{\nu}\cdot \overset{\leftrightarrow}{S} + \left(\left(\hat{\nu}\cdot\vec{v}\right) - W\right)\rho\vec{v}\right]^{+}.
\end{aligned}$$

Using 13.7.22, we finally get

$$\begin{aligned}
\frac{d}{dt}\vec{P}\left[K'(t)^{+}\right] &= \int_{K'(t)^{+}} dV\vec{f} + + \int_{\partial^{+}K'(t)} dA\, \hat{n}\cdot \overset{\leftrightarrow}{S} \\
&\quad + \int_{K'(t)\cap S(t)} dA\left[-\hat{\nu}\cdot \overset{\leftrightarrow}{S} + m\vec{v}\right]^{+}.
\end{aligned} \tag{13.7.29}$$

An exactly similar argument gives

$$\begin{aligned}
\frac{d}{dt}\vec{P}\left[K'(t)^{-}\right] &= \int_{K'(t)^{-}} dV\vec{f} + \int_{\partial^{-}K'(t)} dA\, \hat{n}\cdot \overset{\leftrightarrow}{S} \\
&\quad - \int_{K'(t)\cap S(t)} dA\left[m\vec{v} - \hat{\nu}\cdot \overset{\leftrightarrow}{S}\right]^{-}.
\end{aligned} \tag{13.7.30}$$

Therefore, adding (13.7.29) and (13.7.30),

$$\begin{aligned}
\frac{d}{dt}\vec{P}\left[K'(t)\right] &= \int_{K'(t)} dV\vec{f} + \int_{\partial K'(t)} dA\, \hat{n}\cdot \overset{\leftrightarrow}{S} \\
&\quad + \int_{K'(t)\cap S(t)} dA\left[m\vec{v} - \hat{\nu}\cdot \overset{\leftrightarrow}{S}\right]_{-}^{+}.
\end{aligned} \tag{13.7.31}$$

We still need a mathematical identity for $\vec{\mathcal{F}}[K'(t)]$ in (13.7.24). This seems easy:

$$\vec{\mathcal{F}}\left[K'(t)\right] = \int_{K'(t)} dV\vec{f} + \int_{\partial K'(t)} dA\, \hat{n}\cdot \overset{\leftrightarrow}{S}. \tag{13.7.32}$$

But (13.7.32) does involve assumptions. If $S(t)$ is an air-water interface holding an electric charge per unit area, there will be an electrostatic force per unit area on $S(t)$ which must be added to (13.7.32). In addition, the surface integral on the right in (13.7.32) assumes that the stress on the material just inside $\partial K'(t)$ by the material just outside $\partial K'(t)$ is $\hat{n}\cdot \overset{\leftrightarrow}{S}^{-}$ behind $S(t)$ and $\hat{n}\cdot \overset{\leftrightarrow}{S}^{+}$ in front of $S(t)$, all the way to $S(t)$. In fact, if $S(t)$ is an air-water interface, for example, the water molecules just behind $S(t)$ and the air molecules just in front will be arranged differently from those in the bulk of the air and water. Therefore, along the curve $\partial K'(t) \cap S(t)$ there will be an additional force per unit length exerted by the molecules just outside $\partial K'(t)$ on the molecules just inside $\partial K'(t)$. This force per unit length is called the surface stress in $S(t)$. For an air-water interface it is tangent to $S(t)$ and normal to $\partial K'(t) \cap S(t)$, and its magnitude is the surface tension. We assume

$$\text{there is no surface stress or surface force per unit area on } S(t). \tag{13.7.33}$$

Then (13.7.32) is correct, and substituting (13.7.31) and (13.7.32) in (13.7.24) gives

$$\int_{K' \cap S(t)} dA \left[m\vec{v} - \hat{\nu} \cdot \overleftrightarrow{S} \right]_-^+ = 0 \tag{13.7.34}$$

for any open subset K' of $K(t)$. By the surface version of the vanishing integral theorem, it follows that

$$\left[m\vec{v} - \hat{\nu} \cdot \overleftrightarrow{S} \right]_-^+ = 0 \text{ on } S(t). \tag{13.7.35}$$

If $S(t)$ is a boundary between two materials rather than a shock, then $m = 0$ so (13.7.35) reduces to

$$\left[\hat{\nu} \cdot \overleftrightarrow{S} \right]_-^+ = 0 \text{ on } S(t). \tag{13.7.36}$$

At a boundary between two materials (a contact surface), the stress on the boundary (the "normal stress") is continuous across that boundary.

Equation (13.7.35) has a simple physical interpretation. Consider a small patch dA on $S(t)$. Mass $m\delta t\, dA$ crosses dA from back to front in time δt, and gains momentum $m\delta t\, [\vec{v}]_-^+ \, dA = [m\vec{v}]_-^+ \, \delta t dA$. Therefore, the patch requires momentum to be supplied to it at the rate $[m\vec{v}]_-^+ dA$ per second. Where does this come from? In the material the flux density of momentum is $-\overleftrightarrow{S}$, so momentum arrives at the back of dA at the rate $-\hat{\nu} \cdot \overleftrightarrow{S}^+ \, dA$ and leaves the front at the rate $-\hat{\nu} \cdot \overleftrightarrow{S}^- \, dA$. The net momentum accumulation rate in dA, available to supply the required $[m\vec{v}]_-^+ dA$, is $-\hat{\nu} \cdot \overleftrightarrow{S}^- \, dA + \hat{\nu} \cdot \overleftrightarrow{S}^+ \, +dA$ $= \hat{\nu} \cdot [\overleftrightarrow{S}]_-^+ dA$. Thus (13.7.35) simply says that the difference in momentum flux into the two sides of dA supplies the momentum needed to accelerate the material crossing dA. If $m = 0$, this reduces to action = reaction, i.e. (13.7.36).

13.7.4 Energy conservation

The physical law of conservation of energy is

$$\frac{d}{dt}\mathcal{E}\left[K'(t)\right] = \mathcal{W}\left[K'(t)\right] \tag{13.7.37}$$

where $\mathcal{E}[K'(t)]$ is the total kinetic plus internal energy in $K'(t)$ and $\mathcal{W}[K'(t)]$ is the rate at which energy is being added. We refer again to figure 13.18 and follow the by now well-trodden path to find mathematical identities for both sides of (13.7.37). We have

$$\mathcal{E}\left[K'(t)\right] = \mathcal{E}\left[K'(t)^+\right] + \mathcal{E}\left[K'(t)^-\right]$$

where

$$\mathcal{E}\left[K'(t)^+\right] = \int_{K'(t)^+} dV \left[\rho\left(\frac{1}{2}v^2 + U\right)\right]^+ \tag{13.7.38}$$

with a similar formula for $\mathcal{E}[K'(t)^-]$. Then

$$\begin{aligned}\frac{d}{dt}\mathcal{E}\left[K'(t)^+\right] \;=\; & \int_{K'(t)^+} dV\,\partial_t\left[\rho\left(\frac{1}{2}v^2 + U\right)\right]^+ \\ & + \int_{\partial^+ K'(t)} dA\,\hat{n}\cdot\left[\vec{v}\rho\left(\frac{1}{2}v^2 + U\right)\right]^+ - \int_{K'(t)\cap S(t)} dA\,W\left[\rho\left(\frac{1}{2}v^2 + U\right)\right]^+.\end{aligned}$$

Also

$$\begin{aligned}\textstyle\int_{\partial^+ K'(t)} dA\,\hat{n}\cdot\left[\rho\vec{v}\left(\frac{1}{2}v^2 + U\right)\right]^+ &= \textstyle\int_{\partial' K'(t)^+} dA\,\hat{n}^+\cdot\left[\rho\vec{v}\left(\frac{1}{2}v^2 + U\right)\right]^+ \\ & \quad \textstyle - \int_{K'(t)\cap S(t)} dA\,\hat{n}^+\cdot\left[\rho\vec{v}\left(\frac{1}{2}v^2 + U\right)\right]^+ \\ &= \textstyle\int_{K'(t)^+} dV\,\vec{\partial}\cdot\left[\rho\vec{v}\left(\frac{1}{2}v^2 + U\right)\right]^+ + \int_{K'(t)\cap S(t)} dA\,\hat{\nu}\cdot\left[\rho\vec{v}\left(\frac{1}{2}v^2 + U\right)\right]^+.\end{aligned}$$

Thus,

$$\begin{aligned}\frac{d}{dt}\mathcal{E}\left[K'(t)^+\right] \;=\; & \int_{K'(t)^+} dV\left\{\partial_t\left[\rho\left(\frac{1}{2}v^2 + U\right)\right] + \vec{\partial}\cdot\left[\rho\vec{v}\left(\frac{1}{2}v^2 + U\right)\right]\right\}^+ \\ & + \int_{K'(t)\cap S(t)} dA\left[\left(\hat{\nu}\cdot\vec{v} - W\right)\rho\left(\frac{1}{2}v^2 + U\right)\right]^+.\end{aligned}$$

Now on page 221 we see that IK energy is a stuff with

$$\Phi = \rho\left(\frac{1}{2}v^2 + U\right), \quad \vec{\mathbf{F}} = \rho\vec{v}\left(\frac{1}{2}v^2 + U\right) + \vec{H} - \overleftrightarrow{S}\cdot\vec{v}.$$

Therefore $\partial_t\Phi + \vec{\partial}\cdot\mathbf{F} = \gamma = h + \vec{f}\cdot\vec{v}$ for this stuff. Thus $\partial_t\left[\rho(\frac{1}{2}v^2+U)\right] + \vec{\partial}\cdot\left[\rho\vec{v}\left(\frac{1}{2}v^2+U\right)\right] + \vec{\partial}\cdot\vec{H} - \vec{\partial}\cdot\left(\overset{\leftrightarrow}{S}\cdot\vec{v}\right) = h + \vec{f}\cdot\vec{v}$, and

$$\begin{aligned}
\frac{d}{dt}\mathcal{E}\left[K'(t)^+\right] &= \int_{K'(t)^+} dV\left[-\vec{\partial}\cdot\vec{H} + \vec{\partial}\cdot\left(\overset{\leftrightarrow}{S}\cdot\vec{v}\right) - h + \vec{f}\cdot\vec{v}\right]^+ \\
&\quad + \int_{K'(t)\cap S(t)} dA\ \ m\left(\frac{1}{2}v^2+U\right)^+ \\
&= \int_{K'(t)^+} dV\left(h+\vec{f}\cdot\vec{v}\right)^+ + \int_{\partial K'(t)^+} dA\,\hat{n}^+\cdot\left(-\vec{H}+\overset{\leftrightarrow}{S}\cdot\vec{v}\right)^+ \\
&\quad + \int_{K'(t)\cap S(t)} dA\,m\left(\frac{1}{2}v^2+U\right)^+
\end{aligned}$$

$$\begin{aligned}
\frac{d}{dt}\mathcal{E}\left[K'(t)^+\right] &= \int_{K'(t)^+} dV\left(h+\vec{f}\cdot\vec{v}\right)^+ + \int_{\partial^+K'(t)} dA\,\hat{n}\cdot\left(-\vec{H}+\overset{\leftrightarrow}{S}\cdot\vec{v}\right) \\
&\quad + \int_{K'(t)\cap S(t)} dA\left[m\left(\frac{1}{2}v^2+U\right) + \hat{\nu}\cdot\left(\vec{H}-\overset{\leftrightarrow}{S}\cdot\vec{v}\right)\right]^+ .
\end{aligned} \tag{13.7.39}$$

Similarly,

$$\begin{aligned}
\frac{d}{dt}\mathcal{E}\left[K'(t)^-\right] &= \int_{K'(t)^-} dV\left(h+\vec{f}\cdot\vec{v}\right)^- + \int_{\partial^+K'(t)} dA\,\hat{n}\cdot\left(-\vec{H}+\overset{\leftrightarrow}{S}\cdot\vec{v}\right) \\
&\quad - \int_{K'(t)\cap S(t)} dA\left[m\left(\frac{1}{2}v^2+U\right) + \hat{\nu}\cdot\left(\vec{H}-\overset{\leftrightarrow}{S}\cdot\vec{v}\right)\right]^- .
\end{aligned} \tag{13.7.40}$$

Therefore

$$\begin{aligned}
\frac{d}{dt}\mathcal{E}\left[K'(t)\right] &= \int_{K'(t)} dV\left(h+\vec{f}\cdot\vec{v}\right) + \int_{\partial^+K'(t)} dA\,\hat{n}\cdot\left(-\vec{H}+\overset{\leftrightarrow}{S}\cdot\vec{v}\right) \\
&\quad + \int_{K'(t)\cap S(t)} dA\left[m\left(\frac{1}{2}v^2+U\right) + \hat{\nu}\cdot\left(\vec{H}-\overset{\leftrightarrow}{S}\cdot\vec{v}\right)\right]_-^+ .
\end{aligned} \tag{13.7.41}$$

Also,

$$\mathcal{W}\left[K'(t)\right] = \int_{K'(t)} dV\left(h+\vec{f}\cdot\vec{v}\right) + \int_{\partial K'(t)} dA\,\hat{n}\cdot\left(-\vec{H}+\vec{S}\cdot\vec{v}\right).$$

Therefore, (13.7.37) implies

$$\int_{K'\cap S(t)} dA\left[m\left(\frac{1}{2}v^2+U\right) + \hat{\nu}\cdot\left(\vec{H}-\overset{\leftrightarrow}{S}\cdot\vec{v}\right)\right]_-^+ = 0 \tag{13.7.42}$$

for every open subset K' of $K(t)$. By the surface version of the vanishing integral theorem,

$$\left[m\left(\frac{1}{2}v^2+U\right) + \hat{\nu}\cdot\vec{H} - \hat{\nu}\cdot\overset{\leftrightarrow}{S}\cdot\vec{v}\right]_-^+ = 0 \text{ on } S(t). \tag{13.7.43}$$

If the material is stationary, $m = 0$ and $\vec{v} = 0$, so (13.7.43) reduces to the thermal boundary condition

$$\left[\hat{\nu} \cdot \vec{H}\right]_{-}^{+} = 0. \tag{13.7.44}$$

At a material boundary which is not a shock, $m = 0$, so $[\hat{\nu} \cdot \overleftrightarrow{S}]^{+} = [\hat{\nu} \cdot \overleftrightarrow{S}]^{-}$, and $[\hat{\nu} \cdot \overleftrightarrow{S} \cdot \vec{v}]_{-}^{+} = [\hat{\nu} \cdot \overleftrightarrow{S}]^{+} \cdot \vec{v}^{+} - (\hat{\nu} \cdot \overleftrightarrow{S})^{-} \cdot \vec{v}^{-} = (\hat{\nu} \cdot \overleftrightarrow{S}) \cdot [\vec{v}]_{-}^{+}$ where $\hat{\nu} \cdot \overleftrightarrow{S} = [\hat{\nu} \cdot \overleftrightarrow{S}]^{+} = [\hat{\nu} \cdot \overleftrightarrow{S}]^{-}$. Also, from (13.7.23), if $\vec{v} = v_{\nu}\hat{\nu} + \vec{v}_S$ where $v_v = \hat{\nu} \cdot \vec{v}$, then $v_{\nu}^{+} = v_{\nu}^{-}$, so $[\vec{v}]_{-}^{+} = [\vec{v}_S]^{+}$. Therefore (13.7.43) becomes

$$\left[\hat{\nu} \cdot \vec{H}\right]_{-}^{+} = \left(\hat{\nu} \cdot \overleftrightarrow{S}\right) \cdot [\vec{v}_S]_{-}^{+}. \tag{13.7.45}$$

The quantity $[\hat{\nu} \cdot \vec{H}]_{-}^{+}$ is the net heat flow per unit area per second from the boundary into the surrounding medium. Equation (13.7.45) says that heat is generated mechanically when the two materials slip past one another at the relative velocity $[\vec{v}_S]_{-}^{+}$ and when the surface stress $(\hat{\nu} \cdot \overleftrightarrow{S})$ on the boundary has a component in the direction of the tangential relative slip velocity. Brune, Henyey and Roy (J. Geophys. Res. 74, 3821) used the absence of an observed heat flow anomaly along the San Andreas fault, together with an estimate of $[\vec{v}_S]_{-}^{+}$ on the fault, to put an upper bound on the component of $\hat{\nu} \cdot \overleftrightarrow{S}$ in the direction of slip. They used (13.7.45).

Even when $m \neq 0$, (13.7.43) can be made to resemble 13.7.45. We have

$$\begin{aligned} \left[\frac{1}{2}mv^2\right]_{-}^{+} &= \frac{1}{2}m\left(\vec{v}^{+} - \vec{v}^{-}\right) \cdot \left(\vec{v}^{+} + \vec{v}^{-}\right) \\ &= \frac{1}{2}\hat{\nu} \cdot \left(\overleftrightarrow{S}^{+} - \overleftrightarrow{S}^{-}\right) \cdot \left(\vec{v}^{+} + \vec{v}^{-}\right) \text{ from 13.7.35} \end{aligned}$$

so

$$\begin{aligned} \left[\frac{1}{2}mv^2 - \hat{\nu} \cdot \overleftrightarrow{S} \cdot \vec{v}\right]_{-}^{+} &= \frac{1}{2}\hat{\nu} \cdot \left[- \overleftrightarrow{S}^{+} \cdot \vec{v}^{+} - \overleftrightarrow{S}^{-} \cdot \vec{v}^{+} + \overleftrightarrow{S}^{+} \cdot \vec{v}^{-} + \overleftrightarrow{S}^{-} \cdot \vec{v}^{-}\right] \\ &= -\hat{\nu} \cdot \frac{1}{2}\left(\overleftrightarrow{S}^{+} + \overleftrightarrow{S}^{-}\right) \cdot [\vec{v}]_{-}^{+}. \end{aligned}$$

Define

$$\langle \overleftrightarrow{S} \rangle_{-}^{+} = \frac{1}{2}\left(\overleftrightarrow{S}^{+} + \overleftrightarrow{S}^{-}\right) = \text{ average } \overleftrightarrow{S} \text{ on } S. \tag{13.7.46}$$

Then (13.7.43) can be written in general as

$$\left[\hat{\nu} \cdot \vec{H} + m\,U\right]_{-}^{+} = \hat{\nu} \cdot \langle \overleftrightarrow{S} \rangle_{-}^{+} \cdot [\vec{v}]_{-}^{+}. \tag{13.7.47}$$

Chapter 14

Strain and Deformation

14.1 Finite deformation

Consider a continuum which at time t occupies open subset $K(t)$ of physical space P. Suppose $K'(t_0)$ is a very small open subset of $K(t_0)$. We would like to be able to visualize the particles in $K'(t_0)$, and to see their spatial positions at a particular later time t. The time difference $t - t_0$ is not small, but $K'(t_0)$ is a very small set, in a sense to be made precise shortly.

We use t_0-position labelling. Then label space L is the same as physical space P, and $\vec{\mathbf{r}}^L(\vec{x}, t) =$ position at time t of the particle which was at position $\vec{x}$ at time t_0. The particles in which we are interested have their labels in the set $H' = K'(t_0)$. First choose a particular particle $\vec{x}_A \in K'(t_0)$. Let $\vec{r}_A$ be its position at time t. Then $\vec{r}_A = \vec{\mathbf{r}}^L(\vec{x}_A, t)$. Consider any other particle $\vec{x} \in K'(t_0)$, and let $\vec{r}$ be its position at time t. Then $\vec{r} = \vec{\mathbf{r}}^L(\vec{x}, t)$. Define $\vec{\rho} = \vec{r} - \vec{r}_A$, $\vec{\xi} = \vec{x} - \vec{x}_A$. Then

$$\vec{\rho} = \vec{\mathbf{r}}^L(\vec{x}, t) - \vec{\mathbf{r}}^L(\vec{x}_A, t) = \vec{\mathbf{r}}^L(\vec{x}_A + \vec{\xi}, t) - \vec{\mathbf{r}}^L(\vec{x}_A, t)$$

so

$$\vec{\rho} = \vec{\xi} \cdot \vec{D}\,\vec{\mathbf{r}}^L(\vec{x}_A, t) + \|\vec{\xi}\|\vec{R}\left(\vec{\xi}\right) \tag{14.1.1}$$

where $\vec{R}(\vec{\xi})$ is the remainder function for $\vec{\mathbf{r}}^L(\cdot, t)$ at $\vec{x}_A$.

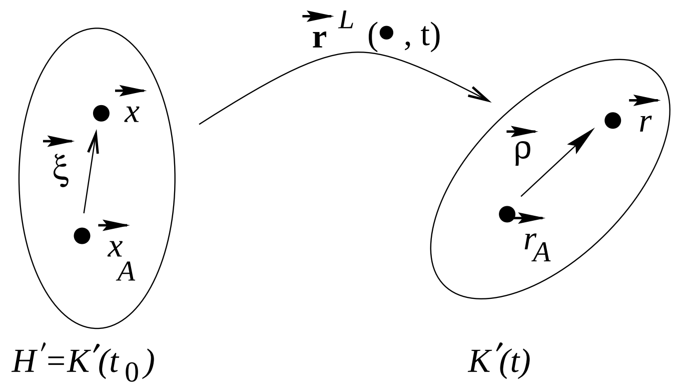


Figure 14.1:

Now we assume that $K'(t_0)$ has been chosen so small that for every $\vec{x} \in K'(t_0)$,

$$\|\vec{R}\left(\vec{\xi}\right)\| << \|\vec{D}\,\vec{\mathbf{r}}^L\,(\vec{x}_A, t)\,\|. \tag{14.1.2}$$

This is the sense in which $K'(t_0)$ must be a small set. If (14.1.2) is satisfied, then we can neglect the remainder term in (14.1.1) and write

$$\vec{\rho} = \vec{\xi} \cdot \vec{D}\,\vec{\mathbf{r}}^L\,(\vec{x}_A, t)\,. \tag{14.1.3}$$

If we measure initial positions in $K'(t_0)$ as displacement vectors from an origin at $\vec{x}_A$, and final positions in $K'(t)$ as displacement vectors from an origin $\vec{r}_A$, the mapping from initial to final positions $\vec{\xi} \longmapsto \vec{\rho}$ is linear, and is (14.1.3). Thus the effect of the motion on the particles in $K'(t_0)$ is clear. The particle at $\vec{x}_A$ is displaced to $\vec{r}_A$. The relative positions of nearby particles, relative to particle $\vec{x}_A$'s position, suffer the linear mapping (14.1.3).

As we have noted (see 12.2.7), $|\det \vec{D}\,\vec{\mathbf{r}}^L(\vec{x}_A, t')|$ is never 0 for any t'. But $\vec{\mathbf{r}}^L(x, t_0) = \vec{x}$, so $\vec{D}\,\vec{\mathbf{r}}^L(\vec{x}, t_0) = \overleftrightarrow{I}_P$ and $\det \vec{D}\,\vec{\mathbf{r}}^L(\vec{x}_A, t_0) = 1$. Then, since $\det \vec{D}\,\vec{\mathbf{r}}^L(\vec{x}_A, t')$ depends continuously on t' and is never 0, it cannot change sign, so it is always positive. Therefore

$$\det \vec{D}\,\vec{\mathbf{r}}^L\,(\vec{x}_A, t) > 0. \tag{14.1.4}$$

It follows from the polar decomposition theorem (page D-27) that there are unique tensors $\overleftrightarrow{C}$ and $\overleftrightarrow{R} \in P \otimes P$ with these properties:

i) $\overleftrightarrow{C}$ is symmetric and positive definite,

ii) $\overleftrightarrow{R}$ is proper orthogonal (i.e., a rotation through some angle θ between 0 and π about some axis $\hat{w}$).

iii)
$$\vec{D}\,\vec{\mathbf{r}}^{\,L}(\vec{x}_A,t) = \overleftrightarrow{C}\cdot\overleftrightarrow{R}\,. \tag{14.1.5}$$

Thus (14.1.3) is
$$\vec{\rho} = \vec{\xi}\cdot\overleftrightarrow{C}\cdot\overleftrightarrow{R}\,. \tag{14.1.6}$$

Thus the effect of the motion on $K'(t_0)$ is to displace $\vec{x}_A$ to $\vec{r}_A$; then to subject the relative position vectors (relative to particle $\vec{x}_A$) to the symmetric operator C; and then to rotate everything through angle θ about an axis passing through $\vec{r}_A$ in the direction of $\hat{w}$.

The displacement $\vec{x}_A \longmapsto \vec{r}_A$ and the rotation $\overleftrightarrow{R}$ are easy to visualize. The mapping $\overleftrightarrow{C}$ may be worth discussing. Since $\overleftrightarrow{C}^T = \overleftrightarrow{C}$ there is an orthonormal basis $\hat{y}_1$, $\hat{y}_2$, $\hat{y}_3$ for P which consists of eigenvectors of C. That is, there are scalars c_1, c_2, c_3 such that

$$C(\hat{y}_i) = c_{(i)}\hat{y}_{(i)}, \quad i = 1,2,3. \tag{14.1.7}$$

Then $\hat{y}_i\cdot\overleftrightarrow{C} = c_{(i)}\hat{y}_{(i)}$ so $\overleftrightarrow{C} = \overleftrightarrow{I}_P\cdot\overleftrightarrow{C} = (\hat{y}_i\,\hat{y}_i)\cdot\overleftrightarrow{C} = \hat{y}_i\left(\hat{y}_i\cdot\overleftrightarrow{C}\right) = \sum_{i=1}^{3} c_{(i)}\,\hat{y}_{(i)}\,\hat{y}_{(i)}$. Thus

$$\overleftrightarrow{C} = c_1\hat{y}_1\hat{y}_1 + c_2\hat{y}_2\hat{y}_2 + c_3\hat{y}_3\hat{y}_3. \tag{14.1.8}$$

Since $\overleftrightarrow{C}$ is positive definite, c_1 and c_2 and c_3 must all be positive. Now $\vec{\xi} = \xi_1\hat{y}_1 + \xi_2\hat{y}_2 + \xi_3\hat{y}_3$ so $\vec{\xi}\cdot\overleftrightarrow{C} = c_1\xi_1\hat{y}_1 + c_2\xi_2\hat{y}_2 + c_3\xi_3\hat{y}_3$. If we imagine a cube with edges of unit length parallel to the axes $\hat{y}_1$, $\hat{y}_2$, $\hat{y}_3$, the effect of $\overleftrightarrow{C}$ is to multiply the i'th edge of the cube by c_i, transforming that cube to the rectangular parallelipiped in figure 14.2.

Suppose that in $K'(t_0)$ we imagine a small cube, with edges of length ε parallel to the axes $\hat{y}_1$, $\hat{y}_2$, $\hat{y}_3$ which are eigenvectors for $\overleftrightarrow{C}$. We can use the cube on the left in 14.2 by shrinking its edges to ε and placing the hidden vertex at $\vec{x}_A$. Then the effect of $\vec{\mathbf{r}}^{\,L}(\cdot,t)$ on this cube is

i) to translate it so the hidden vertex is at $\vec{r}_A$;

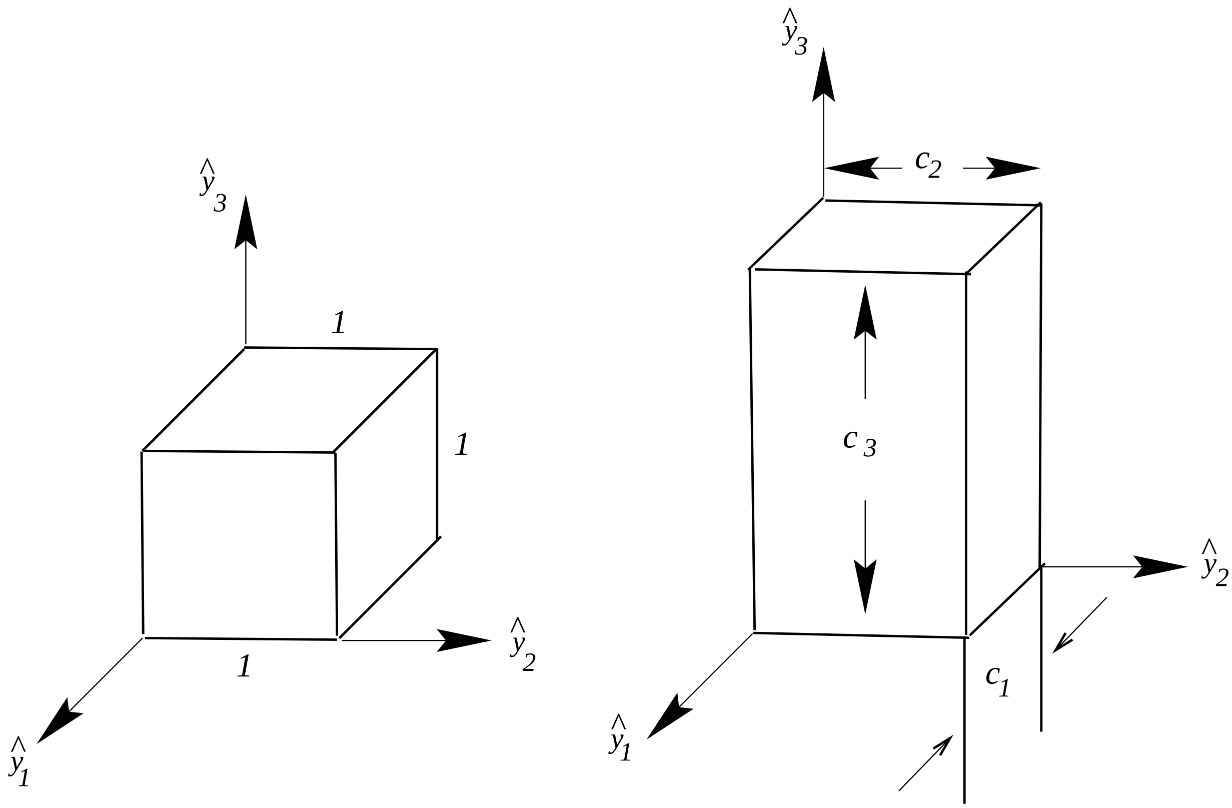


Figure 14.2:

ii) to stretch or compress it by the factor c_i in the direction $\hat{y}_i$, as shown in Figure 14.2;

iii) to rotate it through angle θ, with $0 \leq \theta \leq \pi$, about an axis passing through $\vec{r}_A$ in the direction $\hat{w}$.

Now (14.1.5) implies

$$\vec{D}\,\vec{\mathbf{r}}^{\,L}(\vec{x}_A, t) = \overleftrightarrow{R} \cdot \overleftrightarrow{C}' \tag{14.1.9}$$

where

$$\overleftrightarrow{C}' = \overleftrightarrow{R}^{-1} \cdot \overleftrightarrow{C} \cdot \overleftrightarrow{R} = \sum_{i=1}^{3} c_i \left(\hat{y}_i \cdot \overleftrightarrow{R}\right)\left(\hat{y}_i \cdot \overleftrightarrow{R}\right). \tag{14.1.10}$$

Therefore we can perform the rotation $\overleftrightarrow{R}$ as step ii), and the stretching or compression as step iii). But then, from (14.1.10), the orthogonal axes along which stretching or compression occurs are the rotated axes $\hat{y}_1 \cdot \overleftrightarrow{R}$, $\hat{y}_2 \cdot \overleftrightarrow{R}$, $\hat{y}_3 \cdot \overleftrightarrow{R}$.

All this motivates the following definition:

Definition 14.1.57 Suppose a continuum is labelled with t_0-positions, and $\vec{\mathbf{r}}^{\,L} : K(t_0) \times \mathcal{R} \to P$ is the resulting Lagrangian description of the motion. Suppose $t \in \mathcal{R}$. Then $\vec{\mathbf{r}}^{\,L}(\cdot, t) : K(t_0) \to K(t)$ is called the "deformation" of the continuum from t_0 to t. For any $\vec{x}_A \in K(t_0)$, $\vec{D}\,\vec{\mathbf{r}}^{\,L}(\vec{x}_A, t)$ is the "deformation gradient" at particle $\vec{x}_A$. The rotation $\overleftrightarrow{R}$ in (14.1.5) and (14.1.9) is called the "local rotation at particle $\vec{x}_A$", and $\overleftrightarrow{C}$ in (14.1.5) is the "prerotation stretch tensor at $\vec{x}_A$", while $\overleftrightarrow{C}'$ in (14.1.9,14.1.10) is the "postrotation stretch tensor at $\vec{x}_A$." We will write $\vec{D}\,\vec{\mathbf{r}}^{\,L}(\vec{x}_A, t)$ as $\overleftrightarrow{G}_{t_0}{}^{L}(\vec{x}_A, t)$. For any t_0, $\overleftrightarrow{G}_{t_0}$, is a physical quantity.

Various tensors obtained from $\overleftrightarrow{C}$ or $\overleftrightarrow{C}'$ are called finite strain tensors. The commonest is probably $\overleftrightarrow{C} - \overleftrightarrow{I}_P$, and the most useful is probably

$$\sum_{i=1}^{3} (\ln c_i)\, \hat{y}_i \hat{y}_i.$$

All these strain tensors reduce to the infinitesimal strain tensor when the deformation is nearly the identity mapping (see next section), and to $\overleftrightarrow{0}$ when $\vec{\mathbf{r}}^{\,L}(\cdot, t) = \overleftrightarrow{I}_P$.

In the foregoing discussion, the displacement of particle $\vec{x}_A$ to position $\vec{r}_A$ was so easy to visualize that we said very little about it. It is worth considering. If we fix t_0, the displacement vector of any particle $\vec{x}$ at time t is

$$\vec{s}^L(\vec{x},t) = \vec{\mathbf{r}}^L(\vec{x},t) - \vec{\mathbf{r}}^L(\vec{x},t_0). \tag{14.1.11}$$

As the notation indicates, we will consider the physical quantity $\vec{s}$, called displacement, whose Lagrangian description is the $\vec{s}^L$ defined by (14.1.11). The definition requires that we fix t_0. If we change it, we change the definition of $\vec{s}$.

For t_0-position labelling, $\vec{\mathbf{r}}^L(\vec{x},t_0) = \vec{x} = \vec{\mathbf{x}}^L(\vec{x},t)$ so $\vec{s}^L(\vec{x},t) = \left(\vec{\mathbf{r}}^L - \vec{\mathbf{x}}^L\right)(\vec{x},t)$. Therefore

$$\vec{s} = \vec{\mathbf{r}} - \vec{\mathbf{x}} \text{ for } t_0\text{-position labelling.} \tag{14.1.12}$$

Equation (14.1.12) makes it convenient to use t_0-position labelling when studying displacements. In particular, for such labelling, $(\vec{D}\vec{\mathbf{x}})^L(\vec{x},t) = \vec{D}\vec{x} = \overleftrightarrow{I}_P$ so $\vec{D}\vec{\mathbf{x}} = \overleftrightarrow{I}_P$. Thus, from (14.1.12)

$$\vec{D}\vec{\mathbf{r}} = \overleftrightarrow{I}_P + \vec{D}\vec{s} \text{ for } t_0\text{-position labelling.} \tag{14.1.13}$$

Also, $(\vec{\partial}\vec{\mathbf{r}})^E(\vec{r},t) = \vec{\partial}\vec{r} = \overleftrightarrow{I}_P$ so $\vec{\partial}\vec{\mathbf{r}} = \overleftrightarrow{I}_P$. Then from (14.1.12)

$$\vec{\partial}\vec{\mathbf{x}} = \overleftrightarrow{I}_P - \vec{\partial}\vec{s} \text{ for } t_0\text{-position labelling.} \tag{14.1.14}$$

But $(\vec{\partial}\vec{\mathbf{x}}) \cdot (\vec{D}\vec{\mathbf{r}}) = \overleftrightarrow{I}_P$, so from (14.1.13) and (14.1.14), $(\overleftrightarrow{I}_P - \overleftrightarrow{\partial}\,\vec{s}) \cdot (\overleftrightarrow{I}_P + \vec{D}\vec{s}) = \overleftrightarrow{I}_P$, or $\overleftrightarrow{I}_P + \vec{D}\vec{s} - \vec{\partial}\vec{s} - \vec{\partial}\vec{s} \cdot \vec{D}\vec{s} = \overleftrightarrow{I}_P$, or

$$\vec{D}\vec{s} = \vec{\partial}\vec{s} + \vec{\partial}\vec{s} \cdot \vec{D}\vec{s} = \vec{\partial}\vec{s} \cdot \left(\overleftrightarrow{I}_P + \vec{D}\vec{s}\right). \tag{14.1.15}$$

14.2 Infinitesimal displacements

We use t_0-position labelling throughout this section, and we discuss the displacement $\vec{s}$ defined by (14.1.11) or (14.1.12).

Definition 14.2.58 A tensor $\overleftrightarrow{T} \in P \otimes P$ is "small" if $\| \overleftrightarrow{T} \| << 1$. That is $\| \overleftrightarrow{T} \langle 2 \rangle \overleftrightarrow{T} \|^{\frac{1}{2}} << 1$. A tensor $\overleftrightarrow{Q} \in P \otimes P$ is "close to the identity" if $\overleftrightarrow{Q} - \overleftrightarrow{I}_P$ is small.

Remark 14.2.70 Suppose $\overleftrightarrow{T} \in P \otimes P$ and $\| \overleftrightarrow{T} \| < 1$. Define $(\overleftrightarrow{T})^n = \overleftrightarrow{T}^n$, so $\overleftrightarrow{T}^0 = \overleftrightarrow{I}_P$, $\overleftrightarrow{T}^1 = \overleftrightarrow{T}$, $\overleftrightarrow{T}^2 = \overleftrightarrow{T} \cdot \overleftrightarrow{T}$, $\overleftrightarrow{T}^3 = \overleftrightarrow{T} \cdot \overleftrightarrow{T} \cdot \overleftrightarrow{T}$, etc. Then $(\overleftrightarrow{I}_P + \overleftrightarrow{T})^{-1}$ and $\lim_{N\to\infty} \sum_{n=0}^{N} (-1)^n \overleftrightarrow{T}^n$ both exist, and they are equal.

Proof:

$$\| \sum_{n=M}^{N} (-1)^n \overleftrightarrow{T}^n \| \le \sum_{n=M}^{N} \| \overleftrightarrow{T} \|^n \le \frac{\| \overleftrightarrow{T} \|^M}{1 - \| \overleftrightarrow{T} \|}.$$

As M, $N \to \infty$, this $\to 0$, so the Cauchy convergence criterion assures the existence of $\lim_{N\to\infty} \sum_{n=0}^{N} (-1)^n \overleftrightarrow{T}^n$. Then

$$\begin{aligned} \left(\overleftrightarrow{I}_P + \overleftrightarrow{T}\right) \cdot \sum_{n=0}^{N} (-1)^n \overleftrightarrow{T}^n &= \sum_{n=0}^{N} (-1)^n \overleftrightarrow{T}^n + \sum_{n=0}^{N} (-1)^n \overleftrightarrow{T}^{n+1} \\ &= \sum_{n=0}^{N} (-1)^n \overleftrightarrow{T}^n - \sum_{n=1}^{N+1} (-1)^n \overleftrightarrow{T}^n = \overleftrightarrow{I}_P + (-1)^N \overleftrightarrow{T}^{N+1} . \end{aligned}$$

But $\| \overleftrightarrow{T}^{N+1} \| \le \| \overleftrightarrow{T} \|^{N+1}$, so letting $N \to \infty$ on both sides gives

$$\left(\overleftrightarrow{I}_P + \overleftrightarrow{T}\right) \cdot \sum_{n=0}^{\infty} (-1)^n \overleftrightarrow{T}^n = \overleftrightarrow{I}_P .$$

QED

Corollary 14.2.50 If $\|\vec{D}\vec{s}\| << 1$, then $\|\vec{\partial}\vec{s}\| << 1$, and to first order in $\vec{D}\vec{s}$,

$$\vec{D}\vec{s} = \vec{\partial}\vec{s}. \tag{14.2.1}$$

Proof:

By remark 14.1.14, $(\overleftrightarrow{I}_P + \vec{D}\vec{s})^{-1}$ exists and equals $\sum_{n=0}^{\infty} (-1)^n \left(\vec{D}\vec{s}\right)^n$. Then $\|(\overleftrightarrow{I}_P + \vec{D}\vec{s})^{-1}\| \le \sum_{n=0}^{\infty} \|\vec{D}\vec{s}\|^n = 1/(1 - \|D\vec{s}\|)$. By (14.1.15) , $\vec{\partial}\vec{s} = \vec{D}\vec{s} \cdot (\overleftrightarrow{I}_P + \vec{D}\vec{s})^{-1}$, so

$$\|\partial\vec{s}\| \le \|\vec{D}\vec{s}\| \| \left(I_P + \vec{D}\vec{s}\right)^{-1} \| \le \|\vec{D}\vec{s}\| / \left(1 - \|\vec{D}\vec{s}\|\right).$$

Remark 14.2.71 If $\overleftrightarrow{T} \in P \otimes P$ and $\overleftrightarrow{T}$ is symmetric and $\| \overleftrightarrow{T} \| < 1$, then $\overleftrightarrow{I}_P + \overleftrightarrow{T}$ is positive definite and its positive definite square root is

$$\left(\overleftrightarrow{I}_P + \overleftrightarrow{T} \right)^{1/2} = \sum_{n=0}^{\infty} \begin{pmatrix} \frac{1}{2} \\ n \end{pmatrix} \overleftrightarrow{T}^n \tag{14.2.2}$$

where

$$\begin{pmatrix} N \\ n \end{pmatrix} = N(N-1)\cdots(N-n+1)/n! \quad (= 1 \text{ if } n = 0).$$

Proof:

If $\| \overleftrightarrow{T} \| < 1$, all T's eigenvalues lie between -1 and 1. Therefore $\overleftrightarrow{I}_P + \overleftrightarrow{T}$ has all its eigenvalues between 0 and 2. Therefore, it is positive definite. That $\sum_{n=0}^{\infty} \binom{1/2}{n} \overleftrightarrow{T}^n$ exists, is positive definite, and when squared gives $\overleftrightarrow{I}_P + \overleftrightarrow{T}$ can be proved in the same way as remark (14.2.70). Alternatively, the effect of $\sum_{n=0}^{N} \binom{1/2}{n} \overleftrightarrow{T}^n$ can be studied on each of the vectors in an orthonormal basis for P consisting of eigenvectors of $\overleftrightarrow{T}$. Then one simply uses the known validity of (14.2.2) when $\overleftrightarrow{I}_P$ is replaced by 1 and $\overleftrightarrow{T}$ by any of its eigenvalues.

Corollary 14.2.51 Suppose $\overleftrightarrow{T}^T = \overleftrightarrow{T}$ and $\| \overleftrightarrow{T} \| < 1$. Then correct to first order in $\| \overleftrightarrow{T} \|$,

$$\left(\overleftrightarrow{I}_P + \overleftrightarrow{T} \right)^{1/2} = \overleftrightarrow{I}_P + \frac{1}{2} \overleftrightarrow{T}.$$

Theorem 14.2.29 Use the notation of definition (14.1.9), and define the displacement $\vec{s}$ by (14.1.12). Suppose that $\|\vec{D}\vec{s}\| << 1$. Then to first order in $\vec{D}\vec{s}$ we have

$$\vec{\partial}\vec{s} = \vec{D}\vec{s} \tag{14.2.3}$$

$$\overleftrightarrow{C} = \overleftrightarrow{C}' = \overleftrightarrow{I}_P + \frac{1}{2} \left[\vec{D}\vec{s}\,(\vec{x}_A, t) + \vec{D}\vec{s}\,(\vec{x}_A, t)^T \right] \tag{14.2.4}$$

$$\overleftrightarrow{R} = \overleftrightarrow{I}_P + \frac{1}{2} \left[\vec{D}\vec{s}\,(\vec{x}_A, t) - \vec{D}\vec{s}\,(\vec{x}_A, t)^T \right]. \tag{14.2.5}$$

Proof:

We have already proved (14.2.3). For the rest, abbreviate $\vec{D}\vec{s}(\vec{x}_A,t)$ as $\overleftrightarrow{T}$. Since $\vec{D}\,\vec{\mathbf{r}}\ =\overleftrightarrow{C}\cdot\overleftrightarrow{R}$ we have $(\vec{D}\,\vec{\mathbf{r}})\cdot(\vec{D}\,\vec{\mathbf{r}})^T=\overleftrightarrow{C}\cdot\overleftrightarrow{R}\cdot\overleftrightarrow{R}^T\cdot\overleftrightarrow{C}^T=\overleftrightarrow{C}\cdot\overleftrightarrow{R}\cdot\overleftrightarrow{R}^{-1}\cdot\overleftrightarrow{C}=\overleftrightarrow{C}\cdot\overleftrightarrow{I}_P\cdot\overleftrightarrow{C}=\overleftrightarrow{C}^2$, so

$$\begin{aligned}
\overleftrightarrow{C} &= \left[\left(\vec{D}\,\vec{\mathbf{r}}\right)\cdot\left(\vec{D}\,\vec{\mathbf{r}}\right)^T\right]^{1/2}=\left[\left(\overleftrightarrow{I}_P+\overleftrightarrow{T}\right)\cdot\left(\overleftrightarrow{I}_P+\overleftrightarrow{T}^T\right)\right]^{1/2}\\
&= \left[\overleftrightarrow{I}_P+\overleftrightarrow{T}+\overleftrightarrow{T}^T+\overleftrightarrow{T}\cdot\overleftrightarrow{T}^T\right]^{1/2}\\
&= \overleftrightarrow{I}_P+\frac{1}{2}\left(\overleftrightarrow{T}+\overleftrightarrow{T}^T+\overleftrightarrow{T}\cdot\overleftrightarrow{T}^T\right)+O\left\{\|\overleftrightarrow{T}+\overleftrightarrow{T}^T+\overleftrightarrow{T}\cdot\overleftrightarrow{T}^T\|^2\right\}\\
&= \overleftrightarrow{I}_P+\frac{1}{2}\left(\overleftrightarrow{T}+\overleftrightarrow{T}^T\right)+O\left\{\|\overleftrightarrow{T}\|^2\right\}.
\end{aligned}$$

Also, $\vec{D}\,\vec{\mathbf{r}}\ =\overleftrightarrow{R}\cdot\overleftrightarrow{C}'$, so $\left(\vec{D}\,\vec{\mathbf{r}}\right)^T\cdot\left(\vec{D}\,\vec{\mathbf{r}}\right)=\overleftrightarrow{C}'^T\cdot\overleftrightarrow{R}^T\cdot\overleftrightarrow{R}\cdot\overleftrightarrow{C}'=\overleftrightarrow{C}'\cdot\overleftrightarrow{R}^{-1}\cdot\overleftrightarrow{R}\cdot\overleftrightarrow{C}'=\overleftrightarrow{C}'\cdot\overleftrightarrow{I}_P\cdot\overleftrightarrow{C}'=\left(\overleftrightarrow{C}'\right)^2$ so

$$\begin{aligned}
\vec{C}' &= \left[\left(\vec{D}\,\vec{\mathbf{r}}\right)^T\cdot\left(\vec{D}\,\vec{\mathbf{r}}\right)\right]^{1/2}=\left[\left(\overleftrightarrow{I}_P+\overleftrightarrow{T}^T\right)\cdot\left(\overleftrightarrow{I}_P+\overleftrightarrow{T}\right)\right]^{1/2}\\
&= \left[\overleftrightarrow{I}_P+\overleftrightarrow{T}^T+\overleftrightarrow{T}+\overleftrightarrow{T}^T\cdot\overleftrightarrow{T}\right]\\
&= \overleftrightarrow{I}_P+\frac{1}{2}\left(\overleftrightarrow{T}^T+\overleftrightarrow{T}+\overleftrightarrow{T}^T\cdot\overleftrightarrow{T}\right)+O\left\{\|\overleftrightarrow{T}^T+\overleftrightarrow{T}+\overleftrightarrow{T}^T+\cdot\overleftrightarrow{T}\|^2\right\}\\
&= \overleftrightarrow{I}_P+\frac{1}{2}\left(\overleftrightarrow{T}+\overleftrightarrow{T}^T\right)+O\left\{\|\overleftrightarrow{T}\|^2\right\}.
\end{aligned}$$

Finally, $\overleftrightarrow{R}=\overleftrightarrow{C}^{-1}\cdot\vec{D}\,\vec{\mathbf{r}}\ =\overleftrightarrow{C}^{-1}\cdot\left(\overleftrightarrow{I}_P+\overleftrightarrow{T}\right)$ and $\overleftrightarrow{C}^{-1}=\overleftrightarrow{I}_P-\frac{1}{2}\left(\overleftrightarrow{T}+\overleftrightarrow{T}^T\right)+O\left\{\|\vec{T}\|^2\right\}$ so

$$\begin{aligned}
\overleftrightarrow{R} &= \left(\overleftrightarrow{I}_P-\frac{1}{2}\left(\overleftrightarrow{T}+\overleftrightarrow{T}^T\right)+O\left\{\|\vec{T}\|^2\right\}\right)\cdot\left(\overleftrightarrow{I}_P+\overleftrightarrow{T}\right)\\
&= \overleftrightarrow{I}_P+\frac{1}{2}\left(\overleftrightarrow{T}-\overleftrightarrow{T}^T\right)+O\left\{\|\vec{T}\|^2\right\}.
\end{aligned}$$

QED.

Definition 14.2.59 Use t_0-position labelling so the displacement $\vec{s}$ is defined by (14.1.12). Suppose $\|\vec{D}\vec{s}^L(\vec{x},t)\|<<1$. Then the tensor

$$\overleftrightarrow{\sigma}^L(\vec{x},t)=\frac{1}{2}\left[\vec{D}\vec{s}^L(\vec{x},t)+\vec{D}\vec{s}^L(\vec{x},t)^T\right] \tag{14.2.6}$$

is called the infinitesimal strain tensor at particle $\vec{x}$ at time t, and

$$\overleftrightarrow{w}^L(\vec{x},t) = \frac{1}{2}\left[\vec{D}\vec{s}^L(\vec{x},t) - \vec{D}\vec{s}^L(\vec{x},t)^T\right] \tag{14.2.7}$$

is the infinitesimal rotation tensor. The vector

$$\vec{w}^L(\vec{x},t) = \frac{1}{2}A_P\langle 2\rangle \overleftrightarrow{w}^L(\vec{x},t) \tag{14.2.8}$$

is called the infinitesimal rotation vector at particle $\vec{x}$ at time t.

Corollary 14.2.52 $\vec{w}^L(\vec{x},t)$ can be written $1/2\vec{D}\times\vec{s}^L(\vec{x},t)$ or $1/2\left(\vec{\partial}\times\vec{s}\right)^L(\vec{x},t)$ because

$$\vec{w}^L(\vec{x},t) = \frac{1}{2}A_P\langle 2\rangle\vec{D}\vec{s}^L(\vec{x},t) \tag{14.2.9}$$

and, to first order in $\vec{D}\vec{s}$,

$$\vec{w}^L(\vec{x},t) = \frac{1}{2}A_P\langle 2\rangle\left(\vec{\partial}\vec{s}\right)^L(\vec{x},t). \tag{14.2.10}$$

Proof:

(14.2.10) follows from (14.2.9) and (14.2.3). To prove (14.2.9) note that for any tensor $\overleftrightarrow{T}$ we have

$$A_P\langle 2\rangle \overleftrightarrow{T}^T = -A_P\langle 2\rangle \overleftrightarrow{T}. \tag{14.2.11}$$

To see this, look at components relative to a pooob in P. Then (14.2.11) is equivalent to $\varepsilon_{ijk}T_{kj} = -\varepsilon_{ijk}T_{jk}$. But by changing index names we get $\varepsilon_{ijk}T_{kj} = \varepsilon_{ijk}T_{jk} = -\varepsilon_{ijk}T_{jk}$. Applying (14.2.11) to (14.2.7, 14.2.8) with $\overleftrightarrow{T} = \vec{D}\vec{s}^L(\vec{x},t)$ gives (14.2.9). QED

Corollary 14.2.53 Let $\hat{y}_1$, $\hat{y}_2$, $\hat{y}_3$ be an orthonormal basis for P consisting of eigenvectors of $\overleftrightarrow{\sigma}(\vec{x}_A,t)$, where $\vec{x}_A$ is as in definition 14.1.9. Then to first order in $\vec{D}\vec{s}^L(\vec{x}_A,t)$,

$$\overleftrightarrow{C} = \overleftrightarrow{C}' = \overleftrightarrow{I}_P + \overleftrightarrow{\sigma}(\vec{x}_A,t) = \sum_{i=1}^{3}(1+\sigma_i)\,\hat{y}_i\hat{y}_i \tag{14.2.12}$$

where σ_1, σ_2, σ_3 are the eigenvalues of $\overleftrightarrow{\sigma}(\vec{x}_A,t)$ belonging to the eigenvectors $\hat{y}_1$, $\hat{y}_2$, $\hat{y}_3$. And $\overleftrightarrow{R}$ is a rotation in the right-handed-screw sense through angle $\|\vec{w}^L(\vec{x}_A,t)\|$ about the axis through $\vec{O}_P$ in the direction $\vec{w}^L(\vec{x}_A,t)$.

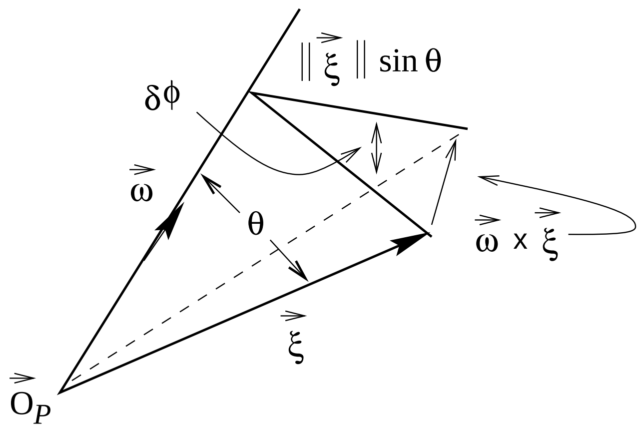


Figure 14.3:

Proof:

We show $\overleftrightarrow{\sigma}(\vec{x}_A, t) = \sum_{i=1}^{3} \sigma_i \hat{y}_i \hat{y}_i$ in the same way we showed (14.1.8). The rest of (14.2.12) follows from (14.2.4). To prove the results about $\overleftrightarrow{R}$, we note from exercise 9c that

$$\overleftrightarrow{w}^L(\vec{x}_A, t) = A_P \cdot \vec{w}^L(\vec{x}_A, t).$$

Then from (14.2.5). to first order in $\vec{D}\vec{s}$

$$\overleftrightarrow{R} = \overleftrightarrow{I}_P + A_P \cdot \vec{w}^L(\vec{x}_A, t).$$

If we apply $\overleftrightarrow{R}$ to a vector $\vec{\xi}$, the result is

$$\vec{\xi} \cdot \overleftrightarrow{R} = \vec{\xi} + \vec{\xi} \cdot A_P \cdot \vec{w}^L(\vec{x}_A, t).$$

We claim $\vec{\xi} \cdot A_P \cdot \vec{v} = \vec{v} \times \vec{\xi}$. Looking at components relative to a pooob, $(\vec{\xi} \cdot A_p \cdot \vec{v})_i = \xi_j \varepsilon_{jik} v_k = \varepsilon_{ikj} v_k \xi_j \left(\vec{v} \times \vec{\xi}\right)_i$. Therefore

$$\vec{\xi} \cdot \overleftrightarrow{R} = \vec{\xi} + \vec{w}^L(\vec{x}_A, t) \times \vec{\xi}. \tag{14.2.13}$$

Abbreviate $\vec{w}^L(\vec{x}_A, t)$ by $\vec{w}$, and examine figure 14.3 where $\vec{\xi} \cdot \vec{R} = \vec{\xi} + \vec{w} \times \vec{\xi}$. Since $\|\vec{w} \times \vec{\xi}\| = \|\vec{w}\| \|\vec{\xi}\| \sin\theta$, clearly $\delta\phi = \|\vec{w}\|$. QED.

Now suppose

$$t = t_0 + \delta t \tag{14.2.14}$$

where δt is small enough that it is a good approximation to write

$$\vec{\mathbf{r}}^{\,L}(\vec{x}, t) = \vec{\mathbf{r}}^{\,L}(\vec{x}, t_0) + \delta t D_t \vec{\mathbf{r}}^{\,L}(\vec{x}, t_0) + O(\delta t)^2 . \tag{14.2.15}$$

Then, since $D_t \vec{\mathbf{r}} = \vec{v}$ and $\vec{s}^L(\vec{x}, t) = \vec{\mathbf{r}}^{\,L}(\vec{x}, t) - \vec{\mathbf{r}}^{\,L}(\vec{x}, t_0)$, we have

$$\vec{s}^L(\vec{x}, t) = \delta t \vec{v}^L(\vec{x}, t_0) + O(\delta t)^2 \tag{14.2.16}$$

Then, to first order in δt,

$$\overleftrightarrow{\sigma}^L(\vec{x}, t) = \overleftrightarrow{\sigma}^L(\vec{x}, t_0 + \delta t) = \delta t \frac{1}{2}\left[\vec{D}\vec{v}^L(\vec{x}, t_0) + \vec{D}\vec{v}^L(\vec{x}, t_0)^T\right] \tag{14.2.17}$$

$$\overleftrightarrow{w}^L(\vec{x}, t) = \overleftrightarrow{w}^L(\vec{x}, t_0 + \delta t) = \delta t \frac{1}{2}\left[\vec{D}\vec{v}^L(\vec{x}, t_0) - \vec{D}\vec{v}^L(\vec{x}, t_0)^T\right] \tag{14.2.18}$$

$$\begin{aligned} \overleftrightarrow{w}^L(\vec{x}, t) = \overleftrightarrow{w}^L(\vec{x}, t_0 + \delta t) &= \delta t \frac{1}{2} A_p\langle 2\rangle \left[\vec{D}\vec{v}^L(\vec{x}, t_0)\right] \\ &= \delta t \frac{1}{2}\vec{D} \times \vec{v}^L(\vec{x}, t_0) . \end{aligned} \tag{14.2.19}$$

Consider the material near particle $\vec{x}_A$ at various times $t_0 + \delta t$, with small δt. Relative to particle $\vec{x}_A$, the material stretches as in (14.2.12), the fractional stretching σ_i being proportional to δt, $\sigma_i = \dot{\sigma}_i \delta t$, where $\dot{\sigma}_i$ is one of the eigenvalues of $1/2\left[\vec{D}\vec{v}^L(\vec{x}_A, t_0) + \vec{D}\vec{v}^L(\vec{x}_A, t_0)^T\right]$, and $\hat{y}_i$ in (14.2.12) is the corresponding eigenvector. And relative to particle $\vec{x}_A$, the material also rotates through angle $\delta t \|\frac{1}{2}\vec{D}\times\vec{v}^L(\vec{x}_A, t)\|$ about an axis through $\vec{x}_A$ and in direction $\frac{1}{2}\vec{D} \times \vec{v}^L(\vec{x}_0, t)$. In other words, it rotates with angular velocity $1/2\vec{D} \times \vec{v}^L(\vec{x}_A, t)$. To first order in δt $\overleftrightarrow{C}' = \overleftrightarrow{C}$, so the rotation can occur before or after the stretching. We prefer to think of them as occurring simultaneously. The foregoing analysis began by choosing a time t_0 at which to introduce t_0-position labels. We can choose t_0 to be any time we like; the foregoing results are true for all $t_0 \in \mathcal{R}$. Therefore, we can define physical quantities $\dot{\overleftrightarrow{\sigma}}$, $\dot{\overleftrightarrow{w}}$, $\dot{\vec{w}}$ by requiring for any $\vec{x} \in P$ and any $t_0 \in \mathbf{r}$ that

$$\dot{\overleftrightarrow{\sigma}}^E(\vec{x}, t_0) = \frac{1}{2}\left[\vec{D}\vec{v}^L(\vec{x}, t_0) + \vec{D}\vec{v}^L(\vec{x}, t_0)^T\right] \tag{14.2.20}$$

$$\overleftrightarrow{w}^E(\vec{x}, t_0) = \frac{1}{2}\left[\vec{D}\vec{v}^L(\vec{x}, t_0) - \vec{D}\vec{v}^L(\vec{x}, t_0)^T\right] \tag{14.2.21}$$

$$\overset{\leftrightarrow}{w}^E(\vec{x}, t_0) = \frac{1}{2}\vec{D} \times \vec{v}^L(\vec{x}, t_0) \qquad . \tag{14.2.22}$$

Here it is understood that to evaluate the left side of (14.2.20, 14.2.21, 14.2.22) at any particular t_0, we use that t_0 to establish t_0-position labels in establishing the Lagrangian descriptions needed on the right. Each t_0 on the left calls for a new Lagrangian description on the right, the one using t_0-position labels.

But with t_0-position labels, $\vec{D}\,\vec{\mathbf{r}}^L(\vec{x}, t_0) = \overleftrightarrow{I}_P$, so $\vec{\partial}^E f^E(\vec{x}, t_0) = \vec{D}^L f^L(\vec{x}, t_0)$ for any physical quantity f. Therefore $\vec{D}\vec{v}^L$ can be replaced by $\vec{\partial}\vec{v}^E$ on the right in 14.2.20. Since the Eulerian description of physical quantities are independent of the way particles are labelled, this replacement leads to the conclusions

$$\overleftrightarrow{\sigma} = \frac{1}{2}\left[\vec{\partial}\vec{v} + \left(\vec{\partial}\vec{v}\right)^T\right] \tag{14.2.23}$$

$$\overleftrightarrow{w} = \frac{1}{2}\left[\vec{\partial}\vec{v} - \left(\vec{\partial}\vec{v}\right)^T\right] \tag{14.2.24}$$

$$\dot{\vec{w}} = \frac{1}{2}\vec{\partial} \times \vec{v}. \tag{14.2.25}$$

The quantity $2\dot{\vec{w}}$ is called the vorticity, and $\overleftrightarrow{\sigma}$ is the strain-rate tensor.

Chapter 15

Constitutive Relations

15.1 Introduction

Let us examine the conservation laws as a system. They are

	Eulerian Form	Lagrangian Form	
Mass	$D_t\rho + \rho\vec{\partial}\cdot\vec{v} = 0$	$D_t\tilde{\rho} = 0$	(15.1.1)
Momentum	$\rho D_t\vec{v} = \vec{\partial}\cdot \overleftrightarrow{S} + \vec{f}$	$\tilde{\rho} D_t\vec{v} = \vec{D}\cdot \overleftrightarrow{\tilde{S}} + \vec{\tilde{f}}$	(15.1.2)
Internal Energy	$\rho D_t U + \vec{\partial}\cdot\vec{H} = h + \overleftrightarrow{S}\langle 2\rangle\partial\vec{v}$	$\tilde{\rho} D_t U + \vec{D}\cdot\vec{\tilde{H}} = \tilde{h} + \overleftrightarrow{\tilde{S}}\langle 2\rangle\vec{D}\vec{v}.$	(15.1.3)

If we knew the initial values of ρ, $\vec{v}$ and U, or $\tilde{\rho}$, $\vec{v}$ and U, at some instant t_0, we might hope to integrate (15.1.3) forward in time to find later values of these quantities. To do this, clearly we must know $\vec{f}$ and h, or $\vec{\tilde{f}}$ and $\tilde{h}$, and we will assume henceforth that these are given to us. But there still remains the question of finding $\vec{H}$ and $\overleftrightarrow{S}$, or $\vec{\tilde{H}}$ and $\overleftrightarrow{\tilde{S}}$. The equations which give their values when the state of the material is known are called the "constitutive equations" of the material. Ideally, we would like to be able to deduce these from the molecular pictures of matter, but theory and computational techniques are not yet adequate to this task for most materials. We must rely on experimental measurements. The outcomes of these experiments suggest idealized mathematical models for various classes of real materials.

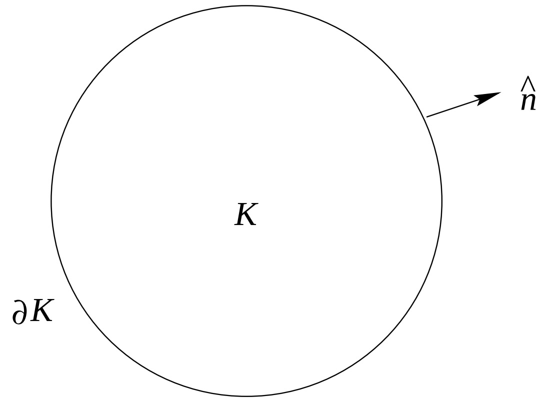


Figure 15.1:

In modelling a real material, one usually begins by considering its homogeneous thermostatic equilibrium states (HTES). Then one studies small deviations from thermostatic equilibrium, by linearizing in those deviations. Large deviations from thermodynamic equilibrium are not fully understood, and we will not consider them. They make plasma physics difficult, for example.

A homogeneous thermostatic equilibrium state (HTES) is said to exist in a material when no macroscopically measurable property of that material changes with time, and when the macroscopic environments of any two particles are the same. The second law of thermodynamics then implies that there can be no heat flow in the material. The material can have a constant velocity, but if it does we study it from a laboratory moving with the same velocity, so we may assume $\vec{v} = \vec{0}$ in an HTES. Then $D_t U = \partial_t U$. But $\partial_t U = 0$ so, from (15.1.3), $h = 0$. The Cauchy stress tensor $\overleftrightarrow{S}$ must be the same everywhere at all times in an HTES, so $\vec{\partial} \cdot \overleftrightarrow{S} = \vec{0}$, and (15.1.3 [momentum]) implies $\vec{f} = \vec{0}$. Therefore, in an HTES we have

$$\begin{aligned}
&\rho = \text{constant}, \quad \vec{H} = \vec{0} \quad \vec{v} = \vec{0}. \\
&U = \text{constant}, \quad h = 0 \\
&\overleftrightarrow{S} = \text{constant}, \quad \vec{f} = \vec{0}
\end{aligned}$$

If the material occupies a volume K with boundary ∂K and unit outward normal $\hat{n}$, then in order to maintain the material in an HTES we must apply to the surface ∂K a stress $\hat{n}\cdot \overleftrightarrow{S}$, and we must prevent surface heat flow and volume heating. That is we need

$$\begin{cases} \text{surface stress applied to } \partial K \text{ at } \vec{r} \text{ is } \hat{n}(\vec{r})\cdot \overleftrightarrow{S} \\ \hat{n}(\vec{r}) \cdot \vec{H} = 0 \text{ at every } \vec{r} \in \partial K \\ h = 0 \text{ at every } \vec{r} \in K. \end{cases} \tag{15.1.4}$$

Two HTES's are different if they differ in any macroscopically measurable quantity, including chemical composition. Most pure substances require only the values of a few of their physical properties in order to specify their HTES's. For example, two isotropically pure samples of H_2O (no deuterium, only O^{16}) liquid which have the same density and the same internal energy per unit mass will have the same values of pressure, temperature, entropy per gram, electrical resistivity, index of refraction for yellow light, neutron mean absorption length at 15 kev, etc. For water, the set of all possible HTES's is a two-parameter family. (We have ignored, and will ignore, electrical properties. It is possible to alter U for water by polarizing it with an electric field.)

Changes from one HTES to another are produced by violating some of the conditions (15.1.4). Changes produced by altering the surface stress applied to ∂K are said to "do work," while changes produced by letting $h \neq 0$ are said to involve "internal heating or cooling," and $\hat{n} \cdot \vec{H} \neq 0$ on ∂K involves "heat flow at ∂K." For the materials in which we are interested, these changes can be described as follows. We pick one HTES as a reference state, and call it HTES$_0$ We use t_0-position labelling, where t_0 is any time when the material is in HTES$_0$. We carry out a change by violating (15.1.4), and wait until time t_1 when the material has settled into a steady state, HTES$_1$. The Lagrangian description of the motion is $\vec{\mathbf{r}}^{\,L}(\vec{x}, t)$, and it is independent of t for $t \leq t_0$ and for $t \geq t_1$. We write

$$\vec{\mathbf{r}}_{01}\left(\vec{x}\right) = \vec{\mathbf{r}}^{\,L}\left(\vec{x}, t_1\right). \tag{15.1.5}$$

$$\overleftrightarrow{G}_{01}\left(\vec{x}\right) = \vec{D}\,\vec{\mathbf{r}}_{01}\left(\vec{x}\right). \tag{15.1.6}$$

Since the environments of all particles must be alike in HTES$_1$, they must have been subjected to the same stretch and the same rotation so

$$\overleftrightarrow{G}_{01} \text{ is constant, independent of } \vec{x}. \tag{15.1.7}$$

($\overleftrightarrow{G}_{01}$ is called the deformation gradient of HTES_1 relative to HTES_0.) Then from (15.1.6), $\vec{\mathbf{r}}_{01}(\vec{x}) = \vec{x} \cdot \overleftrightarrow{G}_{01} + \vec{\mathbf{r}}_{01}(\vec{0})$. If $\vec{\mathbf{r}}_{01}(\vec{0}) \neq \vec{0}$, we may move the material with a uniform displacement so $\vec{\mathbf{r}}_{01}(\vec{0}) = \vec{0}$. Thus we may assume

$$\vec{\mathbf{r}}_{01}(\vec{x}) = \vec{x} \cdot \overleftrightarrow{G}_{01} . \tag{15.1.8}$$

Here $\vec{x}$ is particle position in HTES_0 and $\vec{x} \cdot \overleftrightarrow{G}_{01}$ is particle position in HTES_1. If we want to study another HTES, say HTES_2, we can adopt either HTES_0 or HTES_1 as the reference state. If we use HTES_1 for refrence, then particle position in HTES_2 is $\vec{x}_1 \cdot \overleftrightarrow{G}_{12}$ where $\vec{x}_1$ is particle position in HTES_1 and $\overleftrightarrow{G}_{12}$ is calculated from (15.1.6) but with t_1 position labelling. If a particle has position $\vec{x}$ in HTES_0, then $\vec{x}_1 = \vec{x} \cdot \overleftrightarrow{G}_{01}$ and so $\vec{x}_2 =$ position in $\mathrm{HTES}_2 = \vec{x}_1 \cdot \overleftrightarrow{G}_{12} = \vec{x} \cdot \overleftrightarrow{G}_{01} \cdot \overleftrightarrow{G}_{12}$. But also $\vec{x}_2 = \vec{x} \cdot \overleftrightarrow{G}_{02}$, so

$$\overleftrightarrow{G}_{02} = \overleftrightarrow{G}_{01} \cdot \overleftrightarrow{G}_{12} \tag{15.1.9}$$

for any HTES_0, HTES_1, HTES_2.

15.2 Elastic materials

The observed HTES's of many material are approximated well by the following idealization:

Definition 15.2.60 An elastic material is one whose HTES's have these properties:

i) If HTES_0 is any HTES, there is an $M_0 > 0$ such that for any $\overleftrightarrow{G} \in P \otimes P$ with $\| \overleftrightarrow{G} - \overleftrightarrow{I}_P \| < M_0$, there is an HTES_1 whose deformation gradient relative to HTES_0 is $\overleftrightarrow{G}_{10} = \overleftrightarrow{G}$;

ii) Once a reference state HTES_0 is chosen, any other equilibrium state HTES_1 is completely determined by its deformation gradient $\overleftrightarrow{G}_{01}$ relative to HTES_0, and by its entropy per unit mass N_1.

iii) U, the internal energy per unit mass, is unchanged by rigid rotations. That is, if $\overleftrightarrow{G} \in P \otimes P$ and $\overleftrightarrow{R} \in \Theta^+(P)$ then

$$U\left(\overleftrightarrow{G}, N\right) = U\left(\overleftrightarrow{G} \cdot \overleftrightarrow{R}, N\right) . \tag{15.2.1}$$

Now consider two different HTES's of an elastic material, say HTES_1 and HTES_2, neither of them the reference state HTES_0. Suppose $\vec{\mathbf{r}}^{\,L} : K_0 \times \mathcal{R} \to P$ is the Lagrangian description of the motion which carries the material from state 1 to state 2. Here K_0 is the region occupied by the material in HTES_0, so the total mass of the material is

$$M = |K_0|\tilde{\rho} \tag{15.2.2}$$

where $|K_0|$ = volume of K_0. The total internal energy of the material is

$$\mathbf{U}_i = MU_i = |K_0|\tilde{\rho}U_i \text{ in } \mathrm{HTES}_i. \tag{15.2.3}$$

Let $\triangle_{12}W$ be the work done on the material and $\triangle_{12}Q$ the heat added to it in going from HTES_1 to HTES_2. Energy conservation requires

$$\triangle_{12}W + \triangle_{12}Q = \mathbf{U}_2 - \mathbf{U}_1. \tag{15.2.4}$$

Now suppose HTES_1 and HTES_2 are not very different. Then we will write $df = f_2 - f_1$ for any property f of the two states. Thus (15.2.4) is $d\,\mathbf{U} = \triangle_{12}W + \triangle_{12}Q$. According to the second law of thermodynamics, correct to first order in small changes,

$$\triangle_{12}Q \le \theta d\,\mathbf{N} = \theta(\mathbf{N}_2 - \mathbf{N}_1) \tag{15.2.5}$$

where $\mathbf{N}_i = |K_0|\tilde{\rho}N_i$ = total entropy of state i, and θ is the smaller of θ_1 and θ_2. It is possible to find processes $\vec{\mathbf{r}}^{\,L}$ in which equality is approximated arbitrarily closely in (15.2.5). These are called "reversible processes." In such processes

$$\triangle_{12}Q = \theta|K_0|\tilde{\rho}dN. \tag{15.2.6}$$

In any transition from state 1 to state 2, reversible or not,

$$\tilde{\rho}D_tU = -\vec{D}\cdot\tilde{\vec{H}} + \tilde{h} + \overset{\leftrightarrow}{\tilde{S}}\langle 2\rangle\vec{D}\vec{v}$$

so, integrating over K_0 and using Gauss's theorem,

$$\tilde{\rho}\int_{K_0} dV_L(\vec{x})D_tU = \int_{K_0} dV_L\tilde{h} - \int_{\partial K_0} dA_L\,\hat{n}_L\cdot\tilde{\vec{H}} + \int_{K_0} dV_L\,\overset{\leftrightarrow}{\tilde{S}}\langle 2\rangle\vec{D}\vec{v}$$

or

$$\frac{d}{dt}\mathbf{U} = \dot{Q} + \int_{K_0} dV_L \overset{\leftrightarrow}{\tilde{S}}\langle 2\rangle \vec{D}\vec{v}$$

where $\dot{Q}$ is the rate at which heat is added to the material. Thus, integrating from t_1 to t_2

$$\mathbf{U}_2 - \mathbf{U}_1 = \triangle_{12}Q + \int_{t_1}^{t_2} dt \int_{K_0} dV_L \overset{\leftrightarrow}{\tilde{S}} \langle 2\rangle \vec{D}\vec{v}.$$

Comparing with (15.2.4) we see that

$$\triangle_{12}W = \int_{t_1}^{t_2} dt \int_{K_0} dV_L \overset{\leftrightarrow}{\tilde{S}} \langle 2\rangle \vec{D}\vec{v}.$$

Now if the transformation is done slowly, $\overset{\leftrightarrow}{\tilde{S}}$ will be nearly constant, about equal to its value in HTES_1 or HTES_2 so

$$\triangle_{12}W = \overset{\leftrightarrow}{\tilde{S}} \langle 2\rangle \int_{K_0} dV_L \int_{t_1}^{t_2} dt \vec{D}\vec{v}.$$

But $\vec{v} = d_t\,\vec{\mathbf{r}}$ so $\vec{D}\vec{v} = \vec{D}D_t\,\vec{\mathbf{r}}$, $(\vec{D}\vec{v})^L(\vec{x},t) = \vec{D}D_t\,\vec{\mathbf{r}}^{\,L}(\vec{x},t) = D_t\vec{D}\,\vec{\mathbf{r}}^{\,L}(\vec{x},t)$
$= (D_t \overset{\leftrightarrow}{G})^L(\vec{x},t)$. Thus, $\vec{D}\vec{v} = D_t \overset{\leftrightarrow}{G}$, and

$$\triangle_{12}W = \overset{\leftrightarrow}{\tilde{S}} \langle 2\rangle \int_{K_0} dV_L(\vec{x}) \left[\overset{\leftrightarrow}{G} (\vec{x},t_2) - \overset{\leftrightarrow}{G} (\vec{x},t_1)\right].$$

But $\overset{\leftrightarrow}{G} (\vec{x},t_1)$ and $\overset{\leftrightarrow}{G} (\vec{x},t_2)$ refer to HTES's. They are independent of $\vec{x}$. Thus

$$\triangle_{12}W = |K_0| \overset{\leftrightarrow}{\tilde{S}} \langle 2\rangle d \overset{\leftrightarrow}{G}. \tag{15.2.7}$$

From (15.2.4), (15.2.6) and (15.2.7), $d\,\mathbf{U} = |K_0| \overset{\leftrightarrow}{\tilde{S}} \langle 2\rangle d \overset{\leftrightarrow}{G} + |K_0|\, \tilde{\rho}\theta dN$. Since $d\,\mathbf{U} = |K_0|\tilde{\rho}dU$, dividing by $|K_0|\tilde{\rho}$ gives

$$dU = d \overset{\leftrightarrow}{G} \langle 2\rangle \frac{\overset{\leftrightarrow}{\tilde{S}}}{\tilde{\rho}} + \theta dN \tag{15.2.8}$$

for small reversible changes between HTES's. Here θ and $\overset{\leftrightarrow}{\tilde{S}}$ refer to either the initial or the final state, since they are close.

But U is a function of $\overset{\leftrightarrow}{G}$ and N. Equation (15.2.8) says this function is differentiable, and that

$$\theta\left(\overset{\leftrightarrow}{G}, N\right) = \frac{\partial}{\partial N} U\left(\overset{\leftrightarrow}{G}, N\right) \tag{15.2.9}$$

$$\frac{\overset{\leftrightarrow}{\tilde{S}}\,(\overset{\leftrightarrow}{G}, N)}{\tilde{\rho}} = \vec{\partial}_{\overset{\leftrightarrow}{G}} U\left(\overset{\leftrightarrow}{G}, N\right), \tag{15.2.10}$$

where $\vec{\partial}_{\overset{\leftrightarrow}{G}} U(\overset{\leftrightarrow}{G}, N)$ is an abbreviation for $\left[\vec{\partial} U(\cdot, N)\right](\vec{G})$. Since $P \otimes P$ is a Euclidean space, the theory of such gradients is already worked out. In particular, if $\hat{y}_1$, $\hat{y}_2$, $\hat{y}_3$ is a pooob for P, then the dyads $\hat{y}_i \hat{y}_j$ are an orthonormal basis for $P \otimes P$, and

$$\overset{\leftrightarrow}{G} = G_{ij} \hat{y}_i \hat{y}_j. \tag{15.2.11}$$

Then

$$\partial_{\overset{\leftrightarrow}{G}} U\left(\overset{\leftrightarrow}{G}, N\right) = \hat{y}_i \hat{y}_j \frac{\partial U(\overset{\leftrightarrow}{G}, N)}{\partial G_{ij}}$$

so, from (15.2.10)

$$\frac{\tilde{S}_{ij}}{\rho} = \frac{\partial U(\overset{\leftrightarrow}{G}, N)}{\partial G_{ij}}. \tag{15.2.12}$$

To obtain the Cauchy stress tensor we use (13.3.30), namely

$$\left(\frac{\overset{\leftrightarrow}{S}}{\rho}\right) = \overset{\leftrightarrow}{G}^T \cdot \left(\frac{\overset{\leftrightarrow}{\tilde{S}}}{\tilde{\rho}}\right). \tag{15.2.13}$$

The equation

$$U = U\left(\overset{\leftrightarrow}{G}, N\right) \tag{15.2.14}$$

is called the equation of state of the elastic material. It can be solved for N as a function of $\overset{\leftrightarrow}{G}$ and U,

$$N = N\left(\overset{\leftrightarrow}{G}, U\right). \tag{15.2.15}$$

This serves as the basis for

Definition 15.2.61 A perfectly elastic material is one in which the small region near any particle $\vec{x}$ at any time t behaves as if it were in an HTES with $\overset{\leftrightarrow}{G} = \vec{D}\,\vec{\mathbf{r}}^L(\vec{x}, t)$ and $U = U^L(\vec{x}, t)$.

In a perfectly elastic material, we will have (see page 250) $h = 0$, $\vec{H} = \vec{0}$, and $\overset{\leftrightarrow}{\tilde{S}}$ will be given by (15.2.10), so the missing information needed to integrate (15.1.1, 15.1.2, 15.1.3) is supplied by (15.2.12), which is known if the equation of state is known.

15.3 Small disturbances of an HTES

In seismology, the earth is at rest before an event. This resting state is used to label particles by their positions in the resting state. Then particle displacements are given by

$$\vec{s}^L(\vec{x},t) = \vec{\mathbf{r}}^L(\vec{x},t) - \vec{x}. \tag{15.3.1}$$

Usually it is a very good assumption in seismology that

$$\|\vec{D}\vec{s}^L(\vec{x},t)\| << 1. \tag{15.3.2}$$

Expressed in terms of deformation gradients, we have

$$\overleftrightarrow{G} = \vec{D}\vec{\mathbf{r}} = \overleftrightarrow{I}_P + \vec{D}\vec{s}$$

so

$$\overleftrightarrow{G} = \overleftrightarrow{I}_P + \delta\overleftrightarrow{G} \tag{15.3.3}$$

where

$$\delta\overleftrightarrow{G} = \vec{D}\vec{s}$$

and

$$\|\delta\overleftrightarrow{G}\| << 1. \tag{15.3.4}$$

We want to study HTES's which are close to the zeroth or initial state in the sense of (15.3.4). We use subscripts 0 to refer to the initial state. Then, since particles are labelled by their positions in HTES_0,

$$\overleftrightarrow{G}_0 = \overleftrightarrow{I}_P \tag{15.3.5}$$

$$\tilde{\rho}_0 = \rho_0 \tag{15.3.6}$$

$$\overleftrightarrow{\tilde{S}} = \overleftrightarrow{S}_0 . \tag{15.3.7}$$

Furthermore, if we write $\delta f = f_1 - f_0$ for any physical quantity, f_1 being its value in HTES_1, (since f is constant throughout the material, we needn't distinguish between f^L and f^E) then from (15.2.8), to first order in δN and $\delta\overleftrightarrow{G}$ we have

$$\delta U = \theta\delta N + \frac{1}{\rho_0}\overleftrightarrow{S}_0\langle 2\rangle\delta\overleftrightarrow{G} .$$

If $\delta \overleftrightarrow{G}^T = -\delta \overleftrightarrow{G}$, then $\overleftrightarrow{G} = \overleftrightarrow{I}_P + \delta \overleftrightarrow{G}$ is a rotation to first order in $\delta \overleftrightarrow{G}$, and therefore $\delta N = 0$ implies $\delta U = 0$. Consequently,

$$\overleftrightarrow{S}_0 \langle 2 \rangle \delta \overleftrightarrow{G} = 0 \text{ if } \delta \overleftrightarrow{G}^T = -\delta \overleftrightarrow{G} .$$

But then for any $\delta \overleftrightarrow{G}$, $\overleftrightarrow{S}_0 \langle 2 \rangle \left(\delta \overleftrightarrow{G} - \delta \overleftrightarrow{G}^T \right) = 0$, so $\left(\overleftrightarrow{S}_0 - \overleftrightarrow{S}_0^T \right) \langle 2 \rangle \delta \overleftrightarrow{G} = 0$. Thus

$$\overleftrightarrow{S}_0 = \overleftrightarrow{S}_0^T . \tag{15.3.8}$$

The Cauchy stress tensor is symmetric in every HTES of an elastic material.

We can go further. Since HTES$_1$ is close to HTES$_0$, all the changes δf are small, so we can linearize (15.2.12) and (15.2.13). The changes $\delta\theta = \theta_1 - \theta_0$, $\delta \tilde{\overleftrightarrow{S}} = \tilde{\overleftrightarrow{S}}_1 - \tilde{\overleftrightarrow{S}}^0$ are given in terms of $\delta N = N_1 - N_0$ and $\delta \overleftrightarrow{G} = \overleftrightarrow{G}_1 - \overleftrightarrow{G}_0$ by

$$\begin{aligned}
\delta\theta &= \delta N \frac{\partial \theta \left(\overleftrightarrow{G}_0, N_0 \right)}{\partial N} + \delta \overleftrightarrow{G} \langle 2 \rangle \vec{\partial}_{\overleftrightarrow{G}} \theta \left(\overleftrightarrow{G}_0, N_0 \right) \\
\delta \tilde{\overleftrightarrow{S}} &= \delta N \frac{\partial \tilde{\overleftrightarrow{S}} \left(\overleftrightarrow{G}_0, N_0 \right)}{\partial N} + \delta \overleftrightarrow{G} \langle 2 \rangle \partial_{\overleftrightarrow{G}} \tilde{\overleftrightarrow{S}} \left(\overleftrightarrow{G}_0, N_0 \right) .
\end{aligned}$$

We define the following tensors, $\overleftrightarrow{\overleftrightarrow{Q}}_0 \in \otimes^4 P$, $\overleftrightarrow{W}_0 \in \otimes^2 P$, $B_0 \in \mathcal{R}$:

$$\overleftrightarrow{\overleftrightarrow{Q}}_0 = \vec{\partial}_{\overleftrightarrow{G}} \tilde{\overleftrightarrow{S}} \left(\overleftrightarrow{G}_0, N_0 \right) = \tilde{\rho}_0 \vec{\partial}_{\overleftrightarrow{G}} \vec{\partial}_{\overleftrightarrow{G}} U \left(\overleftrightarrow{G}_0, N_0 \right) \tag{15.3.9}$$

$$\begin{aligned}
\overleftrightarrow{W}_0 &= \frac{\partial}{\partial N} \tilde{\overleftrightarrow{S}} \left(\overleftrightarrow{G}_0, N_0 \right) = \tilde{\rho} \frac{\partial}{\partial N} \vec{\partial}_{\overleftrightarrow{G}} U \left(\overleftrightarrow{G}_0, N_0 \right) \\
&= \tilde{\rho} \vec{\partial}_{\overleftrightarrow{G}} \frac{\partial}{\partial N} U \left(\overleftrightarrow{G}_0, N_0 \right) = \tilde{\rho}_0 \vec{\partial}_{\overleftrightarrow{G}} \theta \left(\vec{\partial} \overleftrightarrow{G}_0, N_0 \right)
\end{aligned} \tag{15.3.10}$$

$$B_0 = \frac{\partial}{\partial N} \theta \left(\overleftrightarrow{G}_0, N_0 \right) = \frac{\partial^2}{\partial N^2} U \left(\overleftrightarrow{G}_0, N_0 \right) . \tag{15.3.11}$$

Then

$$\delta\theta = \delta N B_0 + \delta \overleftrightarrow{G} \langle 2 \rangle \left(\frac{\overleftrightarrow{W}_0}{\tilde{\rho}_0} \right) \tag{15.3.12}$$

(Note: $\tilde{\rho} = \rho$.) Also

$$\delta \tilde{\overleftrightarrow{S}} = \delta N \overleftrightarrow{W}_0 + \delta \overleftrightarrow{G} \langle 2 \rangle \overleftrightarrow{\overleftrightarrow{Q}}_0 . \tag{15.3.13}$$

Furthermore, if $\hat{y}_1$, $\hat{y}_2$, $\hat{y}_3$ is any pooob for P, and we write $\overleftrightarrow{G}$ as in (15.2.11) and write

$$\overleftrightarrow{\overleftrightarrow{Q}}_0 = Q_{ijkl}\hat{y}_i\hat{y}_j\hat{y}_k\hat{y}_l,$$

then

$$Q_{ijkl} = \partial^2 U/\partial G_{ij}\partial G_{kl} = \partial^2 U/\partial G_{kl}\partial G_{ij}$$

so

$$Q_{ijkl} = Q_{klij} \text{ or } (13)(24)\ \overleftrightarrow{\overleftrightarrow{Q}}_0 = \overleftrightarrow{\overleftrightarrow{Q}}_0\ . \tag{15.3.14}$$

We would also like to calculate $\delta\ \overleftrightarrow{S} = \overleftrightarrow{S}_1 - \overleftrightarrow{S}_0$, the change in the Cauchy stress tensor. From (15.2.13) and (13.1.3),

$$\overleftrightarrow{S}_1 = \left|\det \overleftrightarrow{G}_{01}\right|^{-1} \overleftrightarrow{G}^T{}_{01} \cdot \tilde{\overleftrightarrow{S}}_1.$$

Now $\det \overleftrightarrow{G}(t)$ depends continuously on t in going from HTES_0 to HTES_1, and can never vanish. It is 1 in HTES_0, so is always > 0. Hence $|\det \overleftrightarrow{G}| = \det \overleftrightarrow{G}$ and

$$\overleftrightarrow{S}_1 = \left(\det \overleftrightarrow{G}_{01}\right)^{-1} \overleftrightarrow{G}^T{}_{01} \cdot \tilde{\overleftrightarrow{S}}_1. \tag{15.3.15}$$

Now $\overleftrightarrow{G}_{01} = \overleftrightarrow{G} = \overleftrightarrow{I}_P + \delta\ \overleftrightarrow{G}$ and, taking components relative to $\hat{y}_1$, $\hat{y}_2$, $\hat{y}_3$

$$\begin{aligned}
\det \overleftrightarrow{G} &= \begin{vmatrix} 1+\delta G_{11} & \delta G_{12} & \delta G_{13} \\ \delta G_{21} & 1+\delta G_{22} & \delta G_{23} \\ \delta G_{31} & \delta G_{32} & 1+\delta G_{33} \end{vmatrix} \\
&= 1 + \delta G_{ii} + O\left(\|\delta\ \overleftrightarrow{G}\|^2\right) = 1 + \mathrm{tr}\,\delta\ \overleftrightarrow{G} + O\|\ \overleftrightarrow{\delta}\ G\|^2.
\end{aligned}$$

Thus

$$\left(\det \overleftrightarrow{G}\right)^{-1} = 1 - \mathrm{tr}\,\delta\ \overleftrightarrow{G} + O\left(\|\delta\ \overleftrightarrow{G}\|^2\right). \tag{15.3.16}$$

Then

$$\overleftrightarrow{S}_1 = \left(1 - \mathrm{tr}\,\delta\ \overleftrightarrow{G}\right)\left(\overleftrightarrow{I}_P + \delta\ \overleftrightarrow{G}^T\right) \cdot \left(\overleftrightarrow{S}_0 + \delta\tilde{\overleftrightarrow{S}}\right),$$

where we have used (15.3.15) and have linearized in $\delta\ \overleftrightarrow{G}$. Keeping only terms up to first order in $\delta\ \overleftrightarrow{G}$, we have

$$\overleftrightarrow{S}_1 = \overleftrightarrow{S}_0 - [\mathrm{tr}\ (\delta G)]\ \overleftrightarrow{S}_0 + \delta\ \overleftrightarrow{G}^T \cdot \overleftrightarrow{S}_0 + \delta\tilde{\overleftrightarrow{S}},$$

or

$$\delta \overleftrightarrow{S} = \delta \tilde{\overleftrightarrow{S}} + \delta \overleftrightarrow{G}^T \cdot \overleftrightarrow{S}_0 - \overleftrightarrow{S}_0 \left(\operatorname{tr} \delta \overleftrightarrow{G} \right). \tag{15.3.17}$$

If we take components relative to the pooob, $\hat{y}_1$, $\hat{y}_2$, $\hat{y}_3$ from P, and write

$$\overleftrightarrow{S}_0 = S_{ij} \hat{y}_i \hat{y}_j \tag{15.3.18}$$

then (15.3.17) becomes

$$\begin{aligned} \delta S_{kl} &= \delta \tilde{S}_{kl} + \delta G_{ik} S_{il} - \delta G_{jj} S_{kl} \\ &= (\delta G_{ij}) Q_{ijkl} + \delta N W_{kl} + \delta G_{ij} (\delta_{jk} S_{il} - \delta_{ij} S_{kl}) \\ &= (\delta G_{ij}) (Q_{ijkl} + S_{il} \delta_{kj} - \delta_{ij} S_{kl}) + \delta N W_{kl}. \end{aligned}$$

We define a tensor $\overleftrightarrow{\overleftrightarrow{F}} \in \otimes^4 P$ by

$$\overleftrightarrow{\overleftrightarrow{F}}_0 = (Q_{ijkl} + S_{il} \delta_{kj} - \delta_{ij} S_{kl}) \hat{y}_i \hat{y}_j \hat{y}_k \hat{y}_l,$$

or,

$$\overleftrightarrow{\overleftrightarrow{F}}_0 = \overleftrightarrow{\overleftrightarrow{Q}}_0 + (24) \left(\overleftrightarrow{S}_0 \overleftrightarrow{I}_P \right) - \overleftrightarrow{I}_P \overleftrightarrow{S}_0 . \tag{15.3.19}$$

Then

$$F_{ijkl} = Q_{ijkl} + S_{il} \delta_{kj} - \delta_{ij} S_k \tag{15.3.20}$$

and

$$\delta \overleftrightarrow{S} = \delta \overleftrightarrow{G} \langle 2 \rangle \overleftrightarrow{\overleftrightarrow{F}}_0 + \delta N \overleftrightarrow{W}_0 . \tag{15.3.21}$$

This equation is of exactly the same form as (15.3.13), but $\overleftrightarrow{\overleftrightarrow{Q}}_0$ is replaced by $\overleftrightarrow{\overleftrightarrow{F}}_0$.

Now $\delta \overleftrightarrow{S}$ is the difference between the Cauchy stress tensors in two HTES's, so $\delta \overleftrightarrow{S}^T = \delta \overleftrightarrow{S}$ for any $\delta \overleftrightarrow{G}$ whatever and any δN. First, take $\delta \overleftrightarrow{G} = 0$. Then we must have

$$\overleftrightarrow{W}_0^T = \overleftrightarrow{W}_0 . \tag{15.3.22}$$

Next, take $\delta N = 0$. Then

$$\delta S_{kl} = \delta G_{ij} F_{ijkl} = \delta S_{lk} = \delta G_{ij} F_{ijlk},$$

so

$$\delta G_{ij}\left(F_{ijkl}-F_{ijlk}\right)=0$$

for every $\delta \overleftrightarrow{G}$. Hence

$$F_{ijkl}=F_{ijlk} \tag{15.3.23}$$

or

$$(34)\ \overleftrightarrow{\overleftrightarrow{F}}_0=\overleftrightarrow{\overleftrightarrow{F}}_0\ . \tag{15.3.24}$$

Now in general, $\overleftrightarrow{G}$ is a joint property of two HTES's, the zeroth or initial state used to label particles, and the final state. However, $\overleftrightarrow{S}_0$, B_0 , $\overleftrightarrow{W}_0$, $\overleftrightarrow{\overleftrightarrow{Q}}_0$ and $\overleftrightarrow{\overleftrightarrow{F}}_0$ are properties of the initial state alone, HTES_0. As we have seen $\overleftrightarrow{S}_0$ and $\overleftrightarrow{W}_0$ are symmetric, and $\overleftrightarrow{\overleftrightarrow{Q}}_0$ satisfies (15.3.14), while $\overleftrightarrow{\overleftrightarrow{F}}_0$ satisfies (15.3.24). If the initial HTES is isotropic, all these tensors must be isotropic. From this follows

Remark 15.3.72 In an isotropic HTES of an elastic material, there are constants p, α, λ, μ such that (λ, μ are the Lamé parameters of the material)

$$\begin{cases} \overleftrightarrow{S}=-p\ \overleftrightarrow{I}_P \\ \overleftrightarrow{W}=-\alpha\ \overleftrightarrow{I}_P \\ \overleftrightarrow{\overleftrightarrow{F}}=\lambda\ \overleftrightarrow{I}_P\overleftrightarrow{I}_P+\mu\left[(23)+(24)\right]\ \overleftrightarrow{I}_P\overleftrightarrow{I}_P\ \ \text{or} \\ F_{ijkl}=\lambda\delta_{ij}\delta_{kl}+\mu\left(\delta_{ik}\delta_{jl}+\delta_{il}\delta_{kj}\right)\ \text{relative to any pooob for } P. \end{cases} \tag{15.3.25}$$

Proof:

In our discussion of $I[\Theta^+(P)]$ we proved all of this except that for $\overleftrightarrow{\overleftrightarrow{F}}$ we proved only the existence of three scalars λ, μ, ν such that

$$F_{ijkl}=\lambda\delta_{ij}\delta_{kl}+\mu\delta_{ik}\delta_{jl}+\nu\delta_{il}\delta_{kj}.$$

But by (15.3.23), $\mu\delta_{ik}\delta_{jl}+\nu\delta_{il}\delta_{kj}=\mu\delta_{il}\delta_{jk}+\nu\delta_{ik}\delta_{lj}$, or $(\mu-\nu)(\delta_{ik}\delta_{jl}-\delta_{il}\delta_{jk})=0$. Setting $i=k=1$, $j=l=2$ gives $\mu-\nu=0$, so $\nu=\mu$, and we have the expression for $\overleftrightarrow{\overleftrightarrow{F}}$ given in the remark 15.3.72.

Corollary 15.3.54 For HTES's close to an isotropic HTES_0,

$$\left\{\begin{array}{l} \delta\theta = (\delta N)\, B - \left(tr \overset{\leftrightarrow}{\sigma}\right)(\alpha/\rho) \\ \delta \overset{\leftrightarrow}{S} = \lambda \left(\mathrm{tr}\ \overset{\leftrightarrow}{\sigma}\right) \overset{\leftrightarrow}{I}_P + 2u \overset{\leftrightarrow}{\sigma} \end{array}\right. \tag{15.3.26}$$

where $\overset{\leftrightarrow}{\sigma} = 1/2(\delta \overset{\leftrightarrow}{G} + \delta \overset{\leftrightarrow}{G^T}) = 1/2 \left(\vec{D}\vec{s} + \left(\vec{D}\vec{s}\right)^T\right)$ = infinitesimal strain tensor.

In most cases, the HTES_0 of the elastic material is not isotropic, either because the material itself is anisotropic or because $\overset{\leftrightarrow}{S}_0$ is anisotropic. Then the stiffness tensor $\overset{\leftrightarrow}{\overset{\leftrightarrow}{F}}_0$ gives the response of $\overset{\leftrightarrow}{S}$ to small changes $\delta \overset{\leftrightarrow}{G}$ in the deformation gradient $\overset{\leftrightarrow}{G}$ away from $\overset{\leftrightarrow}{G}_0 = \overset{\leftrightarrow}{I}_P$. The stiffness tensor has two sets of symmetries (see 15.3.23):

$$F_{ijkl} = F_{ijlk}. \tag{15.3.27}$$

The second symmetry is obtained (15.3.14) and (15.3.20). Let

$$R_{ijkl} := \delta_{jk} S_{il} - \delta_{ij} S_{kl}. \tag{15.3.28}$$

Then from (15.3.20, $Q_{ijkl} = F_{ijkl} - R_{ijkl}$, so from (15.3.14)

$$F_{ijkl} = F_{klij} + R_{ijkl} - R_{klij}. \tag{15.3.29}$$

Since F_{ijkl} have to be measured experimentally, and there are 81 such components, we would like to use (15.3.27) and (15.3.29) to see how many of these components are ind-pendent. Only those need be measured and recorded.

To answer this question we need one more symmetry, deducible from (15.3.27) and (15.3.29). We have from those two equations

$$F_{ijkl} = F_{klji} + R_{ijkl} - R_{klij}.$$

Interchange i and j in (15.3.29), so

$$F_{kljk} = F_{jikl} - R_{jikl} + R_{klji}.$$

Then

$$F_{ijkl} = F_{jikl} + R_{ijkl} - R_{klij} + R_{klji} - R_{jikl}. \tag{15.3.30}$$

Now define $\overset{\leftrightarrow}{\overset{\leftrightarrow}{E}}_0 \in \otimes^4 P$ by

$$E_{ijkl} = \frac{1}{4}\left[F_{ijkl} + F_{jikl} + F_{klij} + F_{lkij}\right]. \tag{15.3.31}$$

Then it is a trivial algebraic identity that

$$\begin{aligned}
F_{ijkl} &= E_{ijkl} + \frac{1}{4}\left[F_{ijkl} - F_{ijkl}\right] + \frac{1}{4}\left[F_{ijkl} - F_{jikl}\right] \\
&\quad + \frac{1}{4}\left[F_{ijkl} - F_{klij}\right] + \frac{1}{4}\left[F_{ijkl} - F_{lkij}\right] \\
&= E_{ijkl} + \frac{1}{4}\left[F_{ijkl} - F_{jikl}\right] + \frac{1}{4}\left[F_{ijkl} - F_{klij}\right] \\
&\quad + \frac{1}{4}\left[F_{ijkl} - F_{klij} + F_{klij} - F_{lkij}\right] \\
&= E_{ijkl} + \frac{1}{4}\left[F_{ijkl} - F_{jikl}\right] + \frac{1}{2}\left[F_{ijkl} - F_{klij}\right] \\
&\quad + \frac{1}{4}\left[F_{klij} - F_{lkij}\right].
\end{aligned}$$

Using (15.3.30) once as written and once with the interchanges $i \leftrightarrow k$ and $j \leftrightarrow l$, and using (15.3.29 once as written, we get

$$\begin{aligned}
F_{ijkl} &= E_{ijkl} + \frac{1}{4}\left[R_{ijkl} - R_{klij} + R_{klji} - R_{jikl}\right] \\
&\quad + \frac{1}{4}\left[R_{klij} - R_{ijkl} + R_{ijlk} - R_{lkij}\right] \\
&\quad + \frac{1}{2}\left[R_{ijkl} - R_{klij}\right] \\
&= E_{ijkl} + \frac{1}{2}\left(R_{ijkl} - R_{klij}\right) + \frac{1}{4}\left(R_{ijlk} - R_{lkij}\right) \\
&\quad - \frac{1}{4}\left(R_{jikl} - R_{klji}\right) \\
&= E_{ijkl} + \frac{1}{2}\left(\delta_{jk}S_{il} - \delta_{ij}S_{kl} - \delta_{li}S_{kj} + \delta_{kl}S_{ij}\right) \\
&\quad + \frac{1}{4}\left(\delta_{jl}S_{ik} - \delta_{ij}S_{lk} - \delta_{ki}S_{lj} + \delta_{lk}S_{ij}\right) \\
&\quad - \frac{1}{4}\left(\delta_{ik}S_{jl} - \delta_{ji}S_{kl} - \delta_{lj}S_{ki} + \delta_{kl}S_{ji}\right).
\end{aligned}$$

Therefore

$$F_{ijkl} = E_{ijkl} + \frac{1}{2}\left[\delta_{kl}S_{ij} - \delta_{ij}S_{kl} + \delta_{jl}S_{ik} - \delta_{ik}S_{jl} + \delta_{jk}S_{il} - \delta_{il}S_{jk}\right] \tag{15.3.32}$$

where we have used $\delta_{ij} = \delta_{ji}$ and $S_{ij} = S_{ji}$. Now we must measure the six independent components of S_{ij} and the components of E_{ijkl} in order to know $\overset{\leftrightarrow}{\overset{\leftrightarrow}{F}}_0$. How many independent components does $\overset{\leftrightarrow}{\overset{\leftrightarrow}{E}}$ have? From (15.3.31)

$$E_{ijkl} = E_{jikl} = E_{ijlk} = E_{klij}. \tag{15.3.33}$$

Therefore we can think of E_{ijkl} as a 6×6 matrix whose rows are counted by $(ij) = (11)$, (12), (13), (22), (23), (33) and whose columns are counted by $(kl) = (11)$, (12), (13), (22), (23), (33). We need not consider $(ij) = (21)$, (31), or (32) because of the first of equations (15.3.33). The fact that $E_{ijkl} = E_{klij}$ means that this 6×6 matrix is symmetric. Therefore it contains only 21 independent components. These can be chosen arbitrarily but not all choices describe stable materials.

A tensor in $\otimes^4 P$ with the symmetries (15.3.33) is called an "elastic tensor." The set of all such tensors in a subspace of $\otimes^4 P$, which we call $\mathcal{E}^4(P)$. We have just proved that $\dim \mathcal{E}^4(P) = 21$. We have also shown that if $\overset{\leftrightarrow}{\overset{\leftrightarrow}{F}} \in \otimes^4 P$ has the symmetries (15.3.27, 15.3.28, 15.3.29) then it can be written in the form (15.3.32) with $\overset{\leftrightarrow}{\overset{\leftrightarrow}{E}} \in \mathcal{E}^4(P)$. Conversely, if $\overset{\leftrightarrow}{\overset{\leftrightarrow}{E}} \in \mathcal{E}^4(P)$ and $\overset{\leftrightarrow}{S}^T = \overset{\leftrightarrow}{S}$ and $\overset{\leftrightarrow}{\overset{\leftrightarrow}{F}}$ is given by (15.3.32), then $\overset{\leftrightarrow}{\overset{\leftrightarrow}{F}}$ has the symmetries (15.3.27, 15.3.28, 15.3.29), a fact which is easily verified and is left to the reader. Therefore it is not possible to reduce the problem further. In principle, to measure $\overset{\leftrightarrow}{\overset{\leftrightarrow}{F}}_0$, it may be necessary to measure the six independent components of $\overset{\leftrightarrow}{S}_0$ and the 21 independent components of $\overset{\leftrightarrow}{\overset{\leftrightarrow}{\mathcal{E}}}_0$. Actually, only five of the components of $\overset{\leftrightarrow}{S}_0$ need be measured. Suppose $\overset{\leftrightarrow}{\triangle}_0$ is the stress deviator for $\overset{\leftrightarrow}{S}_0$, so

$$\overset{\leftrightarrow}{S}_0 = -p_0 \overset{\leftrightarrow}{I}_P + \overset{\leftrightarrow}{\triangle}_0, \quad \text{with } -3p_0 = \operatorname{tr} \overset{\leftrightarrow}{S}_0 . \tag{15.3.34}$$

Then substituting (15.3.34) in (15.3.27) shows that the isotropic part of $\overset{\leftrightarrow}{S}_0$ makes no contribution to $\overset{\leftrightarrow}{\overset{\leftrightarrow}{F}}_0$, and we can rewrite (15.3.32) as

$$F_{ijkl} = E_{ijkl} + \frac{1}{2}\left[\delta_{kl}\triangle_{ij} - \delta_{ij}\triangle_{kl} + \delta_{jl}\triangle_{ik} - \delta_{ik}\triangle_{jl} + \delta_{jk}\triangle_{il} - \delta_{il}\triangle_{jk}\right]. \tag{15.3.35}$$

Therefore $\overset{\leftrightarrow}{\overset{\leftrightarrow}{F}}_0$ has 26 independent components, 21 in $\overset{\leftrightarrow}{\overset{\leftrightarrow}{E}}_0$ and 5 in $\overset{\leftrightarrow}{\overset{\leftrightarrow}{\triangle}}_0$.

Suppose a material is originally isotropic in an HTES with $\overleftrightarrow{\overleftrightarrow{S}}_0 = \overleftrightarrow{\overleftrightarrow{0}}$. Suppose the material is strongly compressed so that $\overleftrightarrow{\overleftrightarrow{S}}_0 = -p_0 \overleftrightarrow{I}_P$ where p_0 is very large (say of the order of mantle pressure, 0.1 to 1.35 megabar). The new HTES will still be isotropic, and $\overleftrightarrow{\overleftrightarrow{F}}_0 = \overleftrightarrow{\overleftrightarrow{E}}_0$ will have the form (15.3.25), which involves only two constants, the Lamé parameters λ and μ. Suppose then that a small (a few kilobars) stress deviator is added, say $\overleftrightarrow{\triangle}_0$. Then $\overleftrightarrow{\overleftrightarrow{E}}_0$ will change by a small amount $\delta \overleftrightarrow{\overleftrightarrow{E}}$ which depends linearly on $\overleftrightarrow{\overleftrightarrow{\triangle}}_0$, and $\overleftrightarrow{\overleftrightarrow{F}}_0$ will change by $\delta \overleftrightarrow{\overleftrightarrow{F}} = \delta \overleftrightarrow{\overleftrightarrow{E}} +$ terms involving $\overleftrightarrow{\overleftrightarrow{\triangle}}_0$ in (15.3.35). Now $\delta \overleftrightarrow{\overleftrightarrow{E}} = J_0\langle 2\rangle \overleftrightarrow{\triangle}_0$ where $J_0 \in \otimes^6 P$. This tensor is a property of the original highly compressed isotropic HTES. Therefore it is a member of $I^6(\theta^+(V))$, i.e. it is isotropic. By Weyl's theorem, it is a linear combination of all possible permutations of $\overleftrightarrow{I}_P \overleftrightarrow{I}_P \overleftrightarrow{I}_P$. These are as follows:

$$
\begin{array}{ccc}
\delta_{ij}\delta_{kl}\delta_{mn} & \delta_{ij}\delta_{km}\delta_{ln} & \delta_{ij}\delta_{kn}\delta_{lm} \\
\delta_{ik}\delta_{jl}\delta_{mn} & \delta_{ik}\delta_{jm}\delta_{ln} & \delta_{ik}\delta_{jn}\delta_{lm} \\
\delta_{il}\delta_{jk}\delta_{mn} & \delta_{il}\delta_{jm}\delta_{kn} & \delta_{il}\delta_{jn}\delta_{km} \\
\delta_{im}\delta_{jk}\delta_{ln} & \delta_{im}\delta_{jl}\delta_{kn} & \delta_{im}\delta_{jn}\delta_{kl} \\
\delta_{in}\delta_{jk}\delta_{lm} & \delta_{in}\delta_{jl}\delta_{km} & \delta_{in}\delta_{jm}\delta_{kl}.
\end{array}
$$

Since $\triangle_{mn} = \triangle_{nm}$, we can combine two terms which differ only in interchanging m and n. Since $\triangle_{mm} = 0$, we can discard terms with δ_{mn}. Thus there are scalars α, β, γ, ε, ζ such that $2J_{ijklmn} = \alpha(\delta_{ij}\delta_{km}\delta_{ln} + \delta_{ij}\delta_{kn}\delta_{lm}) + \beta(\delta_{ik}\delta_{jm}\delta_{ln} + \delta_{ik}\delta_{jn}\delta_{lm}) + \gamma(\delta_{il}\delta_{jm}\delta_{kn} + \delta_{il}\delta_{jn}\delta_{km}) + \delta(\delta_{im}\delta_{jk}\delta_{ln} + \delta_{in}\delta_{jk}\delta_{lm}) + \varepsilon(\delta_{im}\delta_{jl}\delta_{kn} + \delta_{in}\delta_{jl}\delta_{km}) + \zeta(\delta_{im}\delta_{jn}\delta_{kl} + \delta_{in}\delta_{jm}\delta_{kl})$. Moreover, J_{ijklmn} must satisfy (15.3.33) for any fixed m, n, so $\beta = \gamma = \delta = \varepsilon$ and $\alpha = \zeta$. Therefore J_0 involves only two independent constants, and

$$
\begin{aligned}
2J_{ijklmn} = \; & \alpha\delta_{ij}\left(\delta_{km}\delta_{ln} + \delta_{kl}\delta_{lm}\right) + \alpha\delta_{kl}\left(\delta_{im}\delta_{jn} + \delta_{in}\delta_{jm}\right) \\
& + \beta\Big\{\delta_{ik}\left(\delta_{jm}\delta_{ln} + \delta_{jn}\delta_{lm}\right) + \delta_{il}\left(\delta_{jm}\delta_{kn} + \delta_{jn}\delta_{km}\right) \\
& + \delta_{jk}\left(\delta_{im}\delta_{ln} + \delta_{in}\delta_{lm}\right) + \delta_{jl}\left(\delta_{im}\delta_{kn} + \delta_{in}\delta_{km}\right)\Big\}.
\end{aligned}
$$

Then

$$
\begin{aligned}
\delta E_{ijkl} = \; & J_{ijklmn}\triangle_{mn} = \alpha\left(\delta_{ij}\triangle_{kl} + \delta_{kl}\triangle_{ij}\right) \\
& + \beta\left(\delta_{ik}\triangle_{jl} + \delta_{il}\triangle_{jk} + \delta_{jk}\triangle_{il} + \delta_{jl}\triangle_{ik}\right).
\end{aligned}
$$

Thus, finally, correct to first order in $\triangle_{ij}$, (15.3.35) gives

$$
\begin{aligned}
F_{ijkl} &= \lambda\delta_{ij}\delta_{kl} + u\left(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}\right) && (15.3.36)\\
&+ \left(\alpha - \frac{1}{2}\right)\delta_{ij}\triangle_{kl} + \left(\alpha + \frac{1}{2}\right)\delta_{kl}\triangle_{ij} && (15.3.37)\\
&+ \left(\beta - \frac{1}{2}\right)\left(\delta_{ik}\triangle_{jl} + \delta_{il}\triangle_{jk}\right) + \left(\beta + \frac{1}{2}\right)\left(\delta_{jk}\triangle_{il} + \delta_{jl}\triangle_{ik}\right). && (15.3.38)
\end{aligned}
$$

If the anisotropy in $\overset{\leftrightarrow}{\overset{\leftrightarrow}{F}}_0$ arises solely because an isotropic material has been subjected to a small (anisotropic) deviatoric stress then $\overset{\leftrightarrow}{\overset{\leftrightarrow}{F}}_0$ involves only 9 independent constants, the five independent components of $\overset{\leftrightarrow}{\overset{\leftrightarrow}{\triangle}}_0$ and the four constants λ, u, α, β. If the anisotropy was already present in the crystal structure of the material, either because it was a single crystal, or because its micro-crystals were not randomly oriented, then $\overset{\leftrightarrow}{E}_0$ can involve 21 and $\overset{\leftrightarrow}{\overset{\leftrightarrow}{F}}_0$ 26 independent parameters.

15.4 Small disturbances in a perfectly elastic earth

Neglecting dissipation, we will treat the earth as a perfectly elastic body in studying its response to earthquakes. Before an earthquake, we assume the earth is at rest in static equilibrium. We label the particles by their equilibrium positions, so we use t_0-position labelling where t_0 is any time before the earthquake. When the earth is at rest, we have

$$
\begin{cases}
\vec{\mathbf{r}}^{\,L}\left(\vec{x},t\right) = \vec{x}\\
\overleftrightarrow{G}^{L}\left(\vec{x},t\right) = \left(\mathbf{D}\ \vec{\mathbf{r}}\right)^{L}\left(\vec{x},t\right) = \overleftrightarrow{I}_P
\end{cases}
\qquad (15.4.1)
$$

so the Eulerian and Lagrangian descriptions of any physical quantity are the same, and the two kinds of density and two kinds of stress tensor are the same. Using subscript $_0$ to denote the equilibrium state,

$$
\begin{aligned}
\rho_0\left(\vec{x}\right) &= \tilde{\rho}_0\left(\vec{x}\right)\\
\overleftrightarrow{S}_0\left(\vec{x}\right) &= \overleftrightarrow{\tilde{S}}_0\left(\vec{x}\right).
\end{aligned}
$$

The body force is $\vec{f}\left(\vec{x}\right) = \tilde{\vec{f}}_0(\vec{x}) = \rho_0(\vec{x})\vec{g}_0(\vec{x})$ where $\vec{g}$ is the acceleration of gravity, and $\vec{g}_0$ is its equilibrium value. The conservation equations reduce to

$$
\vec{0} = \vec{\partial}\cdot \overleftrightarrow{S}_0 - \rho_0\vec{g}_0 = \vec{D}\cdot \overleftrightarrow{\tilde{S}}_0 - \tilde{\rho}_0\vec{g}_0. \qquad (15.4.2)
$$

Now suppose an earthquake occurs, after which

$$\vec{\mathbf{r}}^{L}(\vec{x},t)=\vec{x}+\vec{s}^{L}(\vec{x},t), \tag{15.4.3}$$

with

$$\|\vec{D}\vec{s}^{L}(\vec{x},t)\| << 1 \text{ for all } \vec{x},t. \tag{15.4.4}$$

Then

$$\overleftrightarrow{G}^{L}(\vec{x},t)=\vec{D}\,\vec{\mathbf{r}}^{L}(\vec{x},t)=\overleftrightarrow{I}_{P}+\vec{D}\vec{s}^{L}(\vec{x},t). \tag{15.4.5}$$

For earthquakes, $\vec{s}$ as well as $\vec{D}\vec{s}$ is quite small, and it is almost always possible to replace the exact conservation laws and constitutive relations by linearized versions in which all powers of $\vec{s}$ or $\vec{D}\vec{s}$ higher than the first are omitted. In these problems it is useful to be able to convert easily between Eulerian and Lagrangian representations. For example, gravity is easy to calculate in the Eulerian representation, and the constitutive relations are simple in the Lagrangian representation. We discuss this conversion now.

For any physical quantity f, we have

$$f^{L}(\vec{x},t)=f^{E}(\vec{r},t) \text{ where } \vec{r}=\vec{\mathbf{r}}^{L}(\vec{x},t).$$

If $\vec{\mathbf{r}}^{L}(\vec{x},t)=\vec{x}+\vec{s}^{L}(\vec{x},t)$ then to first order in $\vec{s}$ we have

$$\begin{aligned} f^{L}(\vec{x},t) &= f^{E}\left(\vec{x}+\vec{s}^{L}(\vec{x},t),t\right)=f^{E}(\vec{x},t) \\ &\quad +\vec{s}^{L}(\vec{x},t)\cdot\vec{\partial}f^{E}(\vec{x},t). \end{aligned} \tag{15.4.6}$$

If f is small of first order in $\vec{s}$, we have

$$f^{L}(\vec{x},t)=f^{E}(\vec{x},t) \text{ or } f^{L}=f^{E} \text{ if } f=O\|\vec{s}\|. \tag{15.4.7}$$

If f is small of order $\vec{s}$, there is no need to distinguish between its Eulerian and Lagrangian representations, so we can write both $f^{E}(\vec{x},t)$ and $f^{L}(\vec{x},t)$ simply as $f(\vec{x},t)$. That is

$$f(\vec{x},t):=f^{E}(\vec{x},t)=f^{L}(\vec{x},t) \text{ if } f=O\|\vec{s}\|. \tag{15.4.8}$$

This is true in particular for $f=\vec{s}$. If f is of order $\vec{s}$, we can use f not merely for the pair of functions (f^{L},f^{E}) but for either function alone, since to order $\vec{s}$ they are equal. Even when f is not small of order $\vec{s}$, $f^E(\vec{x},t) - f_0(\vec{x})$ is, so to first order in $\vec{s}$, $\vec{s}^L(\vec{x},t)\cdot\vec{\partial}f^E(\vec{x},t)$ $= \vec{s}(\vec{x},t)\cdot\vec{\partial}f_0(\vec{x})$ and (15.4.6) for any f is

$$f^L(\vec{x},t) = f^E(\vec{x},t) + \vec{s}(\vec{x},t)\cdot\vec{\partial}f_0(\vec{x}) \tag{15.4.9}$$

or, as a relationship between functions,

$$f^L = f^E + \vec{s}\cdot\vec{\partial}f_0, \tag{15.4.10}$$

for any physical quantity f.

Relations between derivatives are also needed. For any physical quantity f, $D_t f = \partial_t f + \vec{v}\cdot\vec{\partial}f$. But $\vec{v} = D_t\vec{s} = O\|\vec{s}\|$, and $f - f_0$ is of order $\vec{s}$, so correct to first order in $\vec{s}$, $\vec{v}\cdot\vec{\partial}f = \vec{v}\cdot\vec{\partial}f_0$, and

$$D_t f = \partial_t + \vec{v}\cdot\vec{\partial}f_0. \tag{15.4.11}$$

If f is $O\|\vec{s}\|$ then $\vec{v}\cdot\vec{\partial}f_0$ is $O\|\vec{s}\|^2$, so correct to $O\|\vec{s}\|$

$$D_t f = \partial_t f \text{ if } f = O\|\vec{s}\|. \tag{15.4.12}$$

In particular, correct to $O\|\vec{s}\|$

$$\vec{v} = D_t\vec{s} = \partial_t\vec{s}, \quad \vec{a} = D_t\vec{s} = \partial_t^2\vec{s}. \tag{15.4.13}$$

If f is any physical quantity, $\vec{D}f = (\vec{D}\,\mathbf{\vec{r}}\,)\cdot\vec{\partial}f = (\overset{\leftrightarrow}{I}_P + \vec{D}\vec{s})\cdot\vec{\partial}f = \vec{\partial}f + \vec{D}\vec{s}\cdot\vec{\partial}f$. If $f = O\|\vec{s}\,\|$ then correct to $O\|\vec{s}\,\|$

$$\vec{D}f = \vec{\partial}f \text{ if } f = O\|\vec{s}\|. \tag{15.4.14}$$

In particular, correct to $O\|\vec{s}\|$, $\vec{D}\vec{s} = \vec{\partial}\vec{s}$. Therefore, correct to $O\|\vec{s}\|$, $\vec{D}\vec{s}\cdot\vec{\partial}f = \vec{\partial}\vec{s}\cdot\vec{\partial}f$ $= \vec{\partial}\vec{s}\cdot\vec{\partial}f_0$. Thus for any f

$$\vec{D}f = \vec{\partial}f + \vec{\partial}\vec{s}\cdot\vec{\partial}f_0 \text{ correct to } O\|\vec{s}\|. \tag{15.4.15}$$

For any physical quantity f we define $\delta^L f$ and $\delta^E f$, the Lagrangian and Eulerian derivatives of f from its equilibrium value. Both are physical quantities, and they are defined by

$$\left(\delta^L f\right)^L(\vec{x},t) = f^L(\vec{x},t) - f_0(\vec{x}) \tag{15.4.16}$$

$$\left(\delta^E f\right)^E (\vec{r},t) = f^E (\vec{r},t) - f_0 (\vec{r})\,. \tag{15.4.17}$$

Both $\delta^L f$ and $\delta^E f$ are small of order $\|\vec{s}\|$, so we can apply to them the convention (15.4.8) even when f is not small of order $\vec{s}$. Thus, for the functions,

$$\delta^L f = f^L - f_0 \tag{15.4.18}$$

$$\delta^E f = f^E - f_0. \tag{15.4.19}$$

Then $\delta^L f - \delta^E f = f^L - f^E = \vec{s}\cdot\vec{\partial} f_0$ from (15.4.10), so

$$\delta^L f = \delta^E f + \vec{s}\cdot\vec{\partial} f_0 \text{ to first order in } \vec{s}. \tag{15.4.20}$$

From (15.4.1), $\overleftrightarrow{G}_0 = \overleftrightarrow{I}_P$ so $\vec{s}\cdot\vec{\partial}\,\overleftrightarrow{G}_0 = \overleftrightarrow{0}$. Therefore

$$\delta^L \overleftrightarrow{G} = \delta^E \overleftrightarrow{G} = \vec{\partial}\vec{s} \text{ to first order in } \vec{s}. \tag{15.4.21}$$

With the help of this machinery, the Eulerian and Lagrangian conservation equations and the constitutive relations can be linearized. They produce either Eulerian or Lagrangian equations of motion, and it is conventional to work with the former, even in seismology. We will give the derivation of these equations, based on the Lagrangian constitutive relations and Eulerian conservation equations.

The Eulerian mass conservation equation is

$$\partial_t \rho^E (\vec{r},t) + \vec{\partial}\cdot\rho^E (\vec{r},t)\,\vec{v}^E (\vec{r},t) = 0.$$

We have $\rho^E = \rho_0 + \delta^E\rho$ and $\partial_t\rho_0 = 0$, so, correct to first order in $\vec{s}$,

$$\partial_t\left(\delta^E\rho\right) + \vec{\partial}\cdot(\rho_0\vec{v}) = 0.$$

But $\vec{v} = D_t\vec{s} = \partial_t\vec{s}$ so, correct to first order in $\vec{s}$,

$$\partial_t\left[\delta^E\rho + \vec{\partial}\cdot(\rho_0\vec{s})\right] = 0.$$

For $t = t_0$ (before the earthquake) $\delta^E\rho = 0$ and $\vec{s} = 0$. Therefore

$$\delta^E\rho = -\vec{\partial}\cdot(\rho_0\vec{s})\,. \tag{15.4.22}$$

The Eulerian momentum conservation equation is

$$\left(\rho D_t^2 \vec{s}\right)^E (\vec{r}, t) = \vec{\partial} \cdot \overleftrightarrow{S}^E (\vec{r}, t) - \rho^E (\vec{r}, t)\, \vec{g}^E (\vec{r}, t) .$$

Here $\overleftrightarrow{S}^E = \overleftrightarrow{S}_0 + \delta^E \overleftrightarrow{S}$, $\rho^E = \rho_0 + \delta^E \rho$, $\vec{g}^E = \vec{g}_0 + \delta^E \vec{g}$, so correct to first order in $\vec{s}$,

$$\rho_0 \partial_t^2 \vec{s} = \vec{\partial} \cdot \overleftrightarrow{S}_0 + \vec{\partial} \cdot \left(\delta^E \overleftrightarrow{S}\right) - \rho_0 \vec{g}_0 - \rho_0 \left(\delta^E \vec{g}\right) - \left(\delta^E \rho\right) \vec{g}_0 .$$

But then we can apply (15.4.2) to obtain

$$\rho_0 \partial_t^2 \vec{s} = \vec{\partial} \cdot \left(\delta^E \overleftrightarrow{S}\right) - \rho_0 \left(\delta^E \vec{g}\right) - \left(\delta^E \rho\right) \vec{g}_0 . \tag{15.4.23}$$

We find $\delta^E \rho$ from (15.4.22). We find $\delta^E \vec{g}$ as $\delta^E(-\vec{\partial}\psi) = -\vec{\partial}\psi^E + \vec{\partial}\psi_0 = -\vec{\partial}(\psi^E - \psi^0)$, so

$$\delta^E \vec{g} = -\vec{\partial} \left(\delta^E \psi\right) \tag{15.4.24}$$

where ψ is the gravitational potential. It is obtained from $\partial^2 \psi = 4\pi\rho\Gamma$ where $\Gamma =$ Newton's gravitational constant. Subtracting $\partial^2 \psi^0 = 4\pi\rho^0\Gamma$ gives

$$\partial^2 \delta^E \psi = 4\pi\Gamma\delta^E \rho . \tag{15.4.25}$$

In order to solve (15.4.23), it remains to express $\delta^E \overleftrightarrow{S}$ in terms of $\vec{s}$. We approximate the neighborhood of each particle in the earth as a HTES, and we assume that $\overleftrightarrow{S}$ can be calculated everywhere at all times as if the local material were in such a state. Thus at each particle $\vec{x}$ we can use (15.4.20):

$$\delta^L \overleftrightarrow{S} = \delta^L \overleftrightarrow{G} \langle 2 \rangle \overleftrightarrow{\overleftrightarrow{F}}_0 (\vec{x}) + \left(\delta^L N\right) \overleftrightarrow{W}_0 (\vec{x}) . \tag{15.4.26}$$

We must use the Lagrangian representation here, because (15.3.21) applies to a particular lump of matter, the matter close to particle $\vec{x}$. The initial HTES_0 at particle $\vec{x}$ is the HTES appropriate for the region near that particle before the earthquake. Thus this HTES_0 will vary from particle to particle, and the tensors $\overleftrightarrow{\overleftrightarrow{F}}_0$ and $\overleftrightarrow{W}_0$ really will depend on $\vec{x}$. The heat conductivity of the earth is so low that during the passage of earthquake waves almost no heat has time to flow, so there is no change in entropy per gram at each particular particle. That is

$$\delta^L N = 0 \text{ everywhere at all times.} \tag{15.4.27}$$

Thus equation (15.4.26) gives $\delta^L \overleftrightarrow{S}$ in terms of $\vec{s}$, since $\delta^L \overleftrightarrow{G} = \vec{\partial}\vec{s}$. And then (15.4.20) gives $\delta^E \overleftrightarrow{S}$ as

$$\begin{aligned}\delta^E \overleftrightarrow{S} &= \delta^L \overleftrightarrow{S} - \vec{s} \cdot \vec{\partial}\, \overleftrightarrow{S}_0 \\ \delta^E \overleftrightarrow{S} &= \left(\vec{\partial}\vec{s} \langle 2 \rangle \overleftrightarrow{\overleftrightarrow{F}}_0 \right) - \vec{s} \cdot \vec{\partial}\, \overleftrightarrow{S}_0 \,.\end{aligned}$$

Thus (15.4.23) finally becomes, to first order in $\vec{s}$,

$$\begin{aligned}\rho_0 \partial_t^2 \vec{s} &= \vec{\partial} \cdot \left[\left(\vec{\partial}\vec{s} \right) \langle 2 \rangle \overleftrightarrow{\overleftrightarrow{F}}_0 (\vec{x}) \right] - \vec{\partial} \cdot \left[\vec{s} \cdot \vec{\partial}\, \overleftrightarrow{S}_0 \right] \\ &\quad - \rho_0 \left(\delta^E \vec{g} \right) + \left[\vec{\partial} \cdot (\rho_0 \vec{s}) \right] \vec{g}_0 \end{aligned} \tag{15.4.28}$$

where $\delta^E \vec{g}$ is obtained from $\vec{s}$ via (15.4.22), (15.4.24) and (15.4.25).

For waves short compared to the radius of the earth, it is safe to neglect $\vec{g}$ and $\vec{\partial}\, \overleftrightarrow{S}_0$. If we assume that the undisturbed state is the same everywhere and is isotropic, then

$$F_{ijkl} = \lambda \delta_{ij} \delta_{kl} + \mu \left(\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk} \right)$$

so

$$\begin{aligned}\left[\left(\vec{\partial} s \right) \langle 2 \rangle \overleftrightarrow{\overleftrightarrow{F}}_0 \right]_{kl} &= \lambda \left(\vec{\partial} \cdot \vec{s} \right) \delta_{kl} + \mu \left(\partial_k s_l + \partial_l s_k \right) \\ \left(\vec{\partial}\vec{s} \right) \langle 2 \rangle \overleftrightarrow{\overleftrightarrow{F}}_0 &= \lambda \left(\vec{\partial} \cdot \vec{s} \right) \overleftrightarrow{I}_P + \mu \left[\vec{\partial}\, \vec{s} + (\partial \vec{s})^T \right].\end{aligned}$$

If we assume that there is no $\vec{x}$ dependence in the equilibrium state (homogeneous earth) then λ and μ are independent of $\vec{x}$, and (15.4.28) becomes

$$\rho_0 \partial_t^2 s_l = \partial_k \left[\left(\vec{\partial}\vec{s} \right) \langle 2 \rangle \overleftrightarrow{\overleftrightarrow{F}}_0 \right]_{kl} = \lambda \partial_l \left(\vec{\partial} \cdot \vec{s} \right) + \mu \Big(\partial_k \partial_k s_l + \partial_k \partial_l s_k \Big)$$

$$\begin{aligned}\rho_0 \partial_t^2 \vec{s} &= (\lambda + \mu)\, \vec{\partial} \left(\vec{\partial} \cdot \vec{s} \right) + \mu \partial^2 \vec{s} \\ &= (\lambda + 2\mu)\, \vec{\partial} \left(\vec{\partial} \cdot \vec{s} \right) + \mu \left[\partial^2 \vec{s} - \vec{\partial} \left(\vec{\partial} \cdot \vec{s} \right) \right] \\ &= (\lambda + 2\mu)\, \vec{\partial} \left(\vec{\partial} \cdot \vec{s} \right) - \mu \partial \times \left(\vec{\partial} \times \vec{s} \right).\end{aligned}$$

Divide this equation by ρ_0 and define

$$c_p^2 = \frac{\lambda + 2\mu}{\rho_0}, \quad c_s^2 = \frac{\mu}{\rho_0}. \tag{15.4.29}$$

Then

$$\partial_t^2 \vec{s} = c_p^2 \vec{\partial} \left(\vec{\partial} \cdot \vec{s} \right) - c_s^2 \vec{\partial} \times \left(\vec{\partial} \times \vec{s} \right). \tag{15.4.30}$$

Since we are assuming ρ_0, λ, μ independent of $\vec{x}$, the same is true of c_p^2 and c_s^2. Taking the divergence of (15.4.30) gives

$$\partial_t^2 \left(\vec{\partial} \cdot \vec{s} \right) = c_p^2 \partial^2 \left(\vec{\partial} \cdot \vec{s} \right). \tag{15.4.31}$$

The scalar $\vec{\partial} \cdot \vec{s}$ propagates in waves with velocity c_p. These are called P-waves (primary waves) or compression waves. Taking the curl of (15.4.30) and using $\vec{\partial} \times \vec{\partial} \times \vec{v} = \vec{\partial} \left(\vec{\partial} \cdot \vec{v} \right) - \partial^2 \vec{v}$ gives

$$\partial_t^2 \left(\vec{\partial} \times \vec{s} \right) = c_S^2 \partial^2 \left(\vec{\partial} \times \vec{s} \right). \tag{15.4.32}$$

The vector $\vec{\partial} \times \vec{s}$ propagates in wave with velocity c_s. These are called S-waves (secondary waves) or shear waves.

There is an interesting inequality involving c_s and c_p which we can derive by studying (15.3.21) when $\delta N = 0$ and $\overset{\leftrightarrow}{\overset{\leftrightarrow}{F}}_0$ is isotropic. In that case

$$\delta \overset{\leftrightarrow}{S} = \lambda \overset{\leftrightarrow}{I}_P \left(\operatorname{tr} \delta \overset{\leftrightarrow}{G} \right) + \mu \left[\delta \overset{\leftrightarrow}{G} + \left(\delta \overset{\leftrightarrow}{G} \right)^T \right].$$

For a pure shear, $\vec{s}(\vec{r}) = (\vec{r} \cdot \hat{y}_3)\hat{y}_1$, we have $\delta \overset{\leftrightarrow}{G} = \hat{y}_3\hat{y}_1$ so $\delta \overset{\leftrightarrow}{S} = \mu(\hat{y}_1\hat{y}_3 + \hat{y}_3\hat{y}_1)$ and $\hat{y}_3 \cdot \delta \overset{\leftrightarrow}{S} = \mu\hat{y}_1$. If $\mu \leq 0$, the material either does not resist shear or aids it once it has begun, so we must have

$$\mu > 0. \tag{15.4.33}$$

For a pure dilatation $\vec{s}(\vec{r}) = \varepsilon\vec{r}$, we have $\delta \overset{\leftrightarrow}{G} = \varepsilon \overset{\leftrightarrow}{I}_P$, so $\delta \overset{\leftrightarrow}{S} = \varepsilon(3\lambda + 2\mu) \overset{\leftrightarrow}{I}_P$. If $\varepsilon > 0$, the material expands slightly. If this does not produce a drop in pressure, the material will explode. Therefore

$$3\lambda + 2\mu > 0. \tag{15.4.34}$$

Combining these two relations gives $3(\lambda + 2\mu) > 4\mu > 0$, so

$$0 < \frac{c_s^2}{c_p^2} < \frac{3}{4}, \quad 0 < \frac{c_s}{c_p} < 0.866. \tag{15.4.35}$$

Therefore the P wave always travels faster than the S wave, and arrives first. Hence it is the "primary" wave.

15.5 Viscous fluids

A fluid is a material whose HTES's support no shear, so the Cauchy stress tensor is always isotropic. We need now a more precise definition. A fluid supports no shear in an HTES, and each HTES is uniquely determined by the density and entropy per unit mass, or by volume per unit mass, τ, and entropy per unit mass, N. Clearly $\tau = 1/\rho$, and by our assumption

$$U = U(\tau, N). \tag{15.5.1}$$

From elementary thermodynamics, it is well known that

$$dU = -p d\tau + \theta dN \text{ so} \tag{15.5.2}$$

$$p(\tau, N) = -\frac{\partial U(\tau, N)}{\partial \tau}, \tag{15.5.3}$$

$$\theta(\tau, N) = \frac{\partial U(\tau, N)}{\partial N}, \tag{15.5.4}$$

In an HTES, the Cauchy stress tensor is

$$\overleftrightarrow{S} = -p \overleftrightarrow{I}_P . \tag{15.5.5}$$

Now suppose the fluid deviates slightly from an HTES. At each point in space, $\vec{r}$, and each instant t, there will be a well defined density $\rho^E(\vec{r}, t)$ and a well defined energy per unit mass, $U^E(\vec{r}, t)$. We set $\tau = 1/\rho^E(\vec{r}, t)$, $U = U^E(\vec{r}, t)$ and use (15.5.1) to calculate $N^E(\vec{r}, t)$ as if the fluid were in an HTES at $(\vec{r}, t)$. Then with $N = N^E(\vec{r}, t)$ and $\tau = 1/\rho^E(\vec{r}, t)$ we use (15.5.3) to calculate $p^E(\vec{r}, t)$ as if the fluid were in an HTES at $(\vec{r}, t)$. We call this pressure the thermostatic equilibrium pressure at $\vec{r}, t$ and write it $p^E_{\mathrm{HTES}}(\vec{r}, t)$. The fluid at $(\vec{r}, t)$ is not in fact in the HTES defined by $U^E(\vec{r}, t)$ and $\rho^E(\vec{r}, t)$, so the actual Cauchy stress tensor $\overleftrightarrow{S}^E(\vec{r}, t)$ is not given by (15.5.5) but by

$$\overleftrightarrow{S}^E(\vec{r}, t) = -p^E_{\mathrm{HTES}}(\vec{r}, t) \overleftrightarrow{I}_P + \overleftrightarrow{V}^E(\vec{r}, t) \tag{15.5.6}$$

where (15.5.6) defines $\overleftrightarrow{V}^E$. The tensor $\overleftrightarrow{V}^E(\vec{r}, t)$ is called the viscous stress tensor at $(\vec{r}, t)$. It vanishes in HTES's.

The deviation from an HTES can be measured by the failure of θ to be constant and by the failure of $\vec{v}$ to vanish. Thus $\overleftrightarrow{V}^E(\vec{r},t)$ can be expected to depend on the functions θ^E and $\vec{v}^E$ in the whole fluid. We expect $\overleftrightarrow{V}^E$ to vanish when θ^E is constant for all $\vec{r},t$. In fact, $\vec{v}^E =$ constant for all $\vec{r},t$ is simply an HTES moving with constant velocity, so we expect

$$\begin{aligned}\overleftrightarrow{V}^E(\vec{r},t) &= \overleftrightarrow{0} \text{ if } \theta^E \text{ and } \vec{v^E} \text{ are constant} \\ &\quad \text{for all } \vec{r} \text{ and } t.\end{aligned}$$

For ordinary fluids, we would also expect $\overleftrightarrow{V}^E(\vec{r},t)$ to be affected only by the behavior of the functions θ^E and $\vec{v}^E$ near $\vec{r},t$. This behavior is determined by the derivatives, so we would expect $\overleftrightarrow{V}^E(\vec{r},t)$ to be determined by the values at $(\vec{r},t)$ of

$$\left\{\begin{array}{l} \vec{\partial}\theta^E, \quad \partial_t\theta^E, \quad \vec{\partial}\vec{\partial}\theta^E, \quad \vec{\partial}_t\vec{\partial}\theta^E, \quad \partial_t^2\theta^E, \cdots \\ \\ \vec{\partial}\vec{v^E}, \quad \partial_t\vec{v^E}, \quad \vec{\partial}\vec{\partial}\vec{v^E}, \quad \vec{\partial}_t\vec{\partial}\vec{v^E}, \quad \partial_t^2\vec{v^E}, \cdots \end{array}\right. \tag{15.5.7}$$

We will assume that the deviation from an HTES is small. In that case, $\vec{V}^E(\vec{r},t)$ can be expanded in Taylor series in the quantities (15.5.7), and we can drop all but the linear terms. We will also assume that the influence of distant fluid on $(\vec{r},t)$ drops off very rapidly with distance in either space or time, so that only first derivatives need be considered in (15.5.7). In fact, physical arguments suggest that a term $\vec{\partial}^m\partial_t^n\theta^E$ or $\vec{\partial}^m\partial_t^n\vec{v^E}$ in (15.5.7) contributes to $\overleftrightarrow{V}(\vec{r},t)$ in the ratio $(\lambda/L)^m(t/T)^n$ where λ and t are mean free path and mean time between collisions for a molecule and L and T are the macroscopic wavelength and time period of the disturbance being studied.

On the basis of the foregoing approximations, we expect that $\overleftrightarrow{V}^E(\vec{r},t)$ will depend linearly on $\vec{\partial}\theta^E(\vec{r},t)$, $\partial_t\theta^E(\vec{r},t)$, $\vec{\partial}\vec{v}^E(\vec{r},t)$ and $\partial_t\vec{v}^E(\vec{r},t)$.

We are thus led to a model in which at any point $(\vec{r},t)$ there are tensors $\overset{\leftrightarrow}{\overleftrightarrow{F}}{}^E(\vec{r},t) \in \otimes^4 P$, $\overset{\leftrightarrow}{\overleftrightarrow{J}}{}^E(\vec{r},t)$ and $\overset{\leftrightarrow}{\overleftrightarrow{K}}{}^E(\vec{r},t) \in \otimes^3 P$, and $\overleftrightarrow{H}^E(\vec{r},t) \in \otimes^2 P$ such that

$$\overleftrightarrow{V} = \left(\vec{\partial}\vec{v}\right)\langle 2\rangle \overset{\leftrightarrow}{\overleftrightarrow{F}} + \left(\vec{\partial}\theta\right)\cdot \overset{\rightarrow}{\overleftrightarrow{J}} + (\partial_t\vec{v})\cdot \overset{\rightarrow}{\overleftrightarrow{K}} + (\partial_t\theta)\overleftrightarrow{H}\,. \tag{15.5.8}$$

It seems reasonable to assume that the fluid is skew isotropic. Then the same must be true of $\overset{\leftrightarrow}{\overset{\leftrightarrow}{F}}$, $\overset{\rightarrow}{\overset{\leftrightarrow}{J}}$, $\overset{\rightarrow}{\overset{\leftrightarrow}{K}}$, and $\overset{\leftrightarrow}{H}$. Therefore at each $(\vec{r}, t)$ there are scalars λ, μ, ν, α, β, γ such that relative to any pooob for P

$$\begin{aligned} F_{ijkl} &= \lambda\delta_{ij}\delta_{kl} + \mu\delta_{ik}\delta_{jl} + \nu\delta_{il}\delta_{jk} \\ J_{ijk} &= \alpha\varepsilon_{ijk} \\ K_{ijk} &= \beta\varepsilon_{ijk} \\ H_{ij} &= \gamma\delta ij. \end{aligned}$$

It is observed experimentally that $\vec{\partial}\vec{v}$, $\vec{\partial}\theta$, $\partial_t\vec{v}$ and $\partial_t\theta$ can be adjusted independently. Therefore we can take all but $\vec{\partial}\theta$ to be 0. Then $V_{ij} = \alpha\varepsilon_{ijk}\partial_k\theta$. This $\overset{\leftrightarrow}{V}$ is antisymmetric. If we assume that the intrinsic angular momentum, body torque and torque stress in the fluid are negligible, then $S_{ij} = S_{ji}$, so $V_{ij} = V_{ji}$. Therefore we must have $\alpha = 0$. By the same argument, $\beta = 0$ and $\mu = \nu$. Thus we conclude that

$$V_{ij} = \lambda\delta_{ij}\left(\vec{\partial}\cdot\vec{v}\right) + \mu\left(\partial_i v_j + \partial_j v_i\right) + \gamma\delta_{ij}\partial_t\theta. \tag{15.5.9}$$

If $\vec{v}^E = -\varepsilon\vec{r}$, the fluid is being compressed at a uniform rate. If $\partial_t\theta = 0$ then (15.5.9) gives $V_{ij} = -\delta_{ij}(3\lambda + 2\mu)\varepsilon$. An extra pressure, over and above p_{HTES}, is required to conpress the fluid at a finite rate. The quantity $\chi = \lambda + 2\mu/3$ is therefore called the bulk viscosity:

$$\chi = \lambda + 2/3\mu. \tag{15.5.10}$$

Then

$$V_{ij} = \chi\delta_{ij}\left(\vec{\partial}\cdot\vec{v}\right) + \mu\left[\partial_i v_j + \partial_j v_i - \frac{2}{3}\delta_{ij}\left(\vec{\partial}\cdot\vec{v}\right)\right] + \gamma\delta_{ij}\partial_t\theta. \tag{15.5.11}$$

The quantity μ is called the shear viscosity because in a pure shear flow, $\vec{v}^E(\vec{r}) = w\vec{r}\cdot\hat{y}_3\hat{y}_1$, $\overset{\leftrightarrow}{V} = w\mu(\hat{y}_3\hat{y}_1 + \hat{y}_1\hat{y}_3)$ and $\hat{y}_3\cdot\overset{\leftrightarrow}{V} = \mu\hat{y}_1 w$. The tangential stress $\hat{y}_3\cdot\overset{\leftrightarrow}{V}$ is μ times the rate of shear, w. In the ocean, $\chi/\mu \approx 100$. However, except in sound waves, $\vec{\partial}\cdot\vec{v} \approx 0$, so the large value of χ in the ocean is not noticed except in the damping of short sound waves.

The second law of thermodynamics gives more information about (15.5.11), namely

$$\gamma = 0, \quad \chi \geq 0, \quad \mu \geq 0. \tag{15.5.12}$$

The argument is as follows. At a particle $\vec{x}$ at time t, U and τ are measurable physical quantities even though the material is not in an HTES. If the local values of U and τ are used in (15.5.1), a local value of N can be found for particle $\vec{x}$ at time t. Then (15.5.3) and (15.5.4) give local values of p and θ at particle $\vec{x}$ at time t. These local values of N, p, θ are not measured. They are calculated from the measured τ and U as if the material near $\vec{x}$ at time t were in an HTES. Therefore we denote them by $N^L_{\mathrm{HTES}}(\vec{x},t)$, $p^L_{\mathrm{HTES}}(\vec{x},t)$, $\theta^L_{\mathrm{HTES}}(\vec{x},t)$. Note that p_{HTES} is not $-1/3\,\mathrm{tr}\,\overleftrightarrow{S}$; the latter is a measurable quantity. If we follow particle $\vec{x}$, (15.5.2) implies

$$D_tU = -p_{\mathrm{HTES}}D_t\tau + \theta_{\mathrm{HTES}}D_tN_{\mathrm{HTES}}. \tag{15.5.13}$$

There are no other candidates for θ and N, So we drop the subscripts $_{HTES}$ on θ and N. We do not do so on p_{HTES}, because $-1/3$ tr $\overleftrightarrow{S}$ is another candidate for the pressure and we do not want to confuse the two. If we multiply (15.5.9) by ρ and use (13.6.15) we obtain

$$-\vec{\partial}\cdot\vec{H} + h + \overleftrightarrow{S}\langle 2\rangle\vec{\partial}\vec{v} = -p_{\mathrm{HTES}}\rho D_t\tau + \theta\rho D_tN. \tag{15.5.14}$$

But $\rho D_t\tau = \tau^{-1}D_t\tau = D_t\ln\tau = -D_t\ln\rho = \vec{\partial}\cdot\vec{v}$. Therefore (15.5.14) is

$$\theta\rho D_tN + \partial\cdot H = h + S\langle 2\rangle\partial\vec{v} + p_{\mathrm{HTES}}\left(\vec{\partial}\cdot\vec{v}\right). \tag{15.5.15}$$

But $p_{\mathrm{HTES}}(\vec{\partial}\cdot\vec{v}) = p_{\mathrm{HTES}}\overleftrightarrow{I}_P\langle 2\rangle\vec{\partial}\vec{v}$ so $\overleftrightarrow{S}\langle 2\rangle\vec{\partial}\vec{v} + p_{\mathrm{HTES}}(\vec{\partial}\cdot\vec{v}) = (\overleftrightarrow{S} + p_{\mathrm{HTES}}\overleftrightarrow{I}_P)\langle 2\rangle(\vec{\partial}\vec{v})$ $=\overleftrightarrow{V}\langle 2\rangle\vec{\partial}\vec{v}$ from (15.5.6). If we divide (15.5.15) by θ and use the identity $\vec{\partial}\cdot(\vec{H}/\theta) = (\vec{\partial}\cdot\vec{H})/\theta + H\cdot\vec{\partial}(1/\theta)$, (15.5.15) becomes

$$\rho D_tN + \vec{\partial}\cdot\left(\frac{\vec{H}}{\theta}\right) = \frac{h}{\theta} + \frac{1}{\theta}\overleftrightarrow{V}\langle 2\rangle\vec{\partial}\vec{v} + \vec{H}\cdot\vec{\partial}\left(\frac{1}{\theta}\right). \tag{15.5.16}$$

If we define the stuff "entropy" as the stuff whose density per unit mass is N and whose material flux density is $\vec{H}/\theta$, then (15.5.16) shows that its creation rate η is

$$\eta = \frac{h}{\theta} + \vec{H}\cdot\vec{\partial}\left(\frac{1}{\theta}\right) + \frac{1}{\theta}\overleftrightarrow{V}\langle 2\rangle\vec{\partial}\vec{v}. \tag{15.5.17}$$

Now h can be positive (e.g. ohmic heating) or negative (e.g. heat radiation by a transparent liquid into empty space if the liquid is a blob in interstellar space). However,

$\eta - h/\theta$ is the creation rate of entropy by local events in the fluid. The second law of thermodynamics is generalized to motions near HTES by requiring that everywhere at all times

$$\eta - \frac{h}{\theta} \geq 0. \tag{15.5.18}$$

To use (15.5.18), we must calculate $\overleftrightarrow{V}\langle 2\rangle\vec{\partial}\vec{v}$ in (15.5.17). Since $V_{ij} = V_{ij}$, we have

$$\begin{aligned}
\overleftrightarrow{V}\langle 2\rangle\vec{\partial}\vec{v} &= V_{ij}\partial_i v_j = \frac{1}{2}\left(V_{ij} + V_{ji}\right)\partial_i v_j = \frac{1}{2}V_{ij}\partial_i v_j + \frac{1}{2}V_{ji}\partial v_j \\
&= \frac{1}{2}V_{ij}\partial_i v_j + \frac{1}{2}V_{ij}\partial_j v_i = V_{ij}\frac{1}{2}\left(\partial_i v_j + \partial_j v_i\right) \\
&= \overleftrightarrow{V}\langle 2\rangle\frac{1}{2}\left[\vec{\partial}\vec{v} + \left(\vec{\partial}\vec{v}\right)^T\right].
\end{aligned}$$

Now suppose we write (15.5.10) as

$$V_{ij} = \delta_{ij}\left(\chi\vec{\partial}\cdot\vec{v} + \gamma\partial_t\theta\right) + \mu\left(\partial_i v_j + \partial_j v_i - \frac{2}{3}\delta_{ij}\vec{\partial}\cdot\vec{v}\right) \tag{15.5.19}$$

and write

$$\frac{1}{2}\left(\partial_i v_j + \partial_j v_i\right) = \delta_{ij}\left(\frac{1}{3}\vec{\partial}\cdot\vec{v}\right) + \frac{1}{2}\left(\partial_i v_j + \partial_j v_i - \frac{2}{3}\delta_{ij}\vec{\partial}\cdot\vec{v}\right). \tag{15.5.20}$$

Then we have expressed $\overleftrightarrow{V}$ and $1/2[\vec{\partial}\vec{v}+(\vec{\partial}\vec{v})^T]$ as the sum of their isotropic and deviatoric parts. Since an isotropic and a deviatoric tensor are orthogonal, there are no cross terms when we calculate $\overleftrightarrow{V}\langle 2\rangle\frac{1}{2}[\vec{\partial}\vec{v} + (\vec{\partial}\vec{v})^T]$ from (15.5.19) and (15.5.20). Therefore

$$\overleftrightarrow{V}\langle 2\rangle\vec{\partial}\vec{v} = \left(\vec{\partial}\cdot\vec{v}\right)\left(\chi\vec{\partial}\cdot\vec{v} + \gamma\partial_t\theta\right) + \frac{\mu}{2}\|\vec{\partial}\vec{v} + \left(\vec{\partial}\vec{v}\right)^T - \frac{2}{3}\overleftrightarrow{I}_P\,\vec{\partial}\cdot\vec{v}\|^2.$$

Thus for a viscous fluid (15.5.18) requires

$$\theta\vec{H}\cdot\vec{\partial}\left(\frac{1}{\theta}\right) + \chi\left(\vec{\partial}\cdot\vec{v}\right)^2 + \frac{\mu}{2}\left\|\vec{\partial}\vec{v} + \left(\vec{\partial}\vec{v}\right)^T - \frac{2}{3}\overleftrightarrow{I}_P\left(\vec{\partial}\cdot\vec{v}\right)\right\|^2 + \gamma\left(\vec{\partial}\cdot\vec{v}\right)\partial_t\theta \geq 0. \tag{15.5.21}$$

We can arrange experiments in which $\vec{\partial}\theta = \vec{0}$, and $\partial_t\theta = 0$, and $\vec{v}^E(\vec{r},t) = \varepsilon\vec{r} + \zeta(\vec{r}\cdot\hat{y}_3)\hat{y}_1$. Then $\vec{\partial}\vec{v} = \varepsilon I_p + \zeta\hat{y}_3\hat{y}_1$, $\vec{\partial}\cdot\vec{v} = 3\varepsilon$, and (15.5.21) becomes $9\varepsilon^2\chi + 2\mu\zeta^2 \geq 0$. Since this inequality must hold for $\varepsilon = 0$, $\zeta \neq 0$ and for $\varepsilon \neq 0$, $\zeta = 0$, we must have $\chi \geq 0$, $\mu \geq 0$. We can also arrange experiments with $\vec{v}^E(\vec{r},t) = \varepsilon\vec{r}$, $\vec{\partial}\theta = \vec{0}$, and $\partial_t\theta$ of any value. In such an experiment, (15.5.21) becomes

$$9\varepsilon^2\chi + 3\varepsilon\gamma\partial_t\theta \geq 0. \tag{15.5.22}$$

If $\gamma \neq 0$, take $\varepsilon > 0$ and $\gamma \partial_t \theta < -3\varepsilon\chi$. Then (15.5.22) is violated. Therefore we must have $\gamma = 0$. The fact that (15.5.12) are observed experimentally is one of the arguments for (15.5.18) as an extension of the second law of thermodynamics.

Part III

Exercises

Set One

Exercise 1

Let $(\vec{b}_1, \cdots, \vec{b}_n)$ be an ordered basis for Euclidean vector space V. Let $(\vec{b}^1, \cdots, \vec{b}^n)$ be its dual basis. Let g_{ij} and g^{ij} be its covariant and contravariant metric matrices. Prove

a) $\vec{b}_i = g_{ij}\vec{b}^j$

b) $\vec{b}^i = g^{ij}\vec{b}_j$

c) $g_{ij}g^{jk} = \delta_i{}^k$

d) $g^{ij}g_{jk} = \delta^i{}_k$

Note that either c) or d) says that the two $n \times n$ matrices g_{ij} and g^{ij} are inverse to each other. Thus we have a way to construct $(\vec{b}^1, \cdots, \vec{b}^n)$ when $(\vec{b}_1, \cdots, \vec{b}_n)$ are given. This is the procedure:

Step 1. Compute $g_{ij} = \vec{b}_i \cdot \vec{b}_j$.

Step 2. Compute the inverse matrix to g_{ij}, the matrix g^{ij} such that $g_{ij}g^{jk} = \delta_i{}^k$ or $g^{ij}g_{jk} = \delta^i{}_k$ (either equation implies the other).

Step 3. $\vec{b}^i = g^{ij}\vec{b}_j$.

Exercise 2

V is ordinary Euclidean 3-space, with the usual dot and cross products. The function M is defined by one of the following six equations, where $\vec{v}_1$, $\vec{v}_2$ $\vec{v}_3$, $\vec{v}_4$ are arbitrary vectors in V. In each case, state whether M is multilinear, whether M is a tensor, and whether M is totally symmetric or totally antisymmetric. State why you think so.

i) $M(\vec{v}_1, \vec{v}_2) = \vec{v}_1 \cdot \vec{v}_2$.

ii) $M(\vec{v}_1, \vec{v}_2) = \vec{v}_1 \times \vec{v}_2$.

iii) $M(\vec{v}_1, \vec{v}_2) = (\vec{v}_1 \cdot \vec{v}_2)\vec{v}_2$.

iv) $M(\vec{v}_1, \vec{v}_2) = (\vec{v}_1 \times \vec{v}_2) \cdot \vec{v}_2$.

v) $M(\vec{v}_1, \vec{v}_2, \vec{v}_3) = (\vec{v}_1 \times \vec{v}_2) \cdot \vec{v}_3$.

vi) $M(\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4) = (\vec{v}_1 \cdot \vec{v}_2)(\vec{v}_3 \cdot \vec{v}_4)$.

Solutions

1.a) For any vector $\vec{v} \in V$, we know that

$$\vec{v} = \left(\vec{v} \cdot \vec{b}_j\right) \vec{b}^j$$

See equations (D-14) and (D-15).

Let $i \in \{1, \cdots, n\}$, and apply this result to the vector $\vec{v} = \vec{b}_i$. Thus we obtain

$$\vec{b}_i = \left(\vec{b}_i \cdot \vec{b}_j\right) \vec{b}^j, \text{ or } \vec{b}_i = g_{ij}\vec{b}^j.$$

1.b) For any $\vec{v} \in V$, $\vec{v} = (\vec{v} \cdot \vec{b}^j)\vec{b}_j$ (see equations (D-12) and (D-13). Apply this result to $\vec{v} = \vec{b}^i$. Thus $\vec{b}^i = (\vec{b}^i \cdot \vec{b}^j)\vec{b}_j$, or $\vec{b}^i = g^{ij}\vec{b}_j$. Note that part b) can also be obtained immediately by regarding $(\vec{b}^1, \cdots, \vec{b}^n)$ as the original basis and $(\vec{b}_1, \cdots, \vec{b}_n)$ as the dual basis.

1.c) Dot $\vec{b}^k$ into a) and use $\vec{b}_i \cdot \vec{b}^k = \delta_i{}^k$.

1d. Dot $\vec{b}_k$ into b) and use $\vec{b}^i \cdot \vec{b}_k = \delta^i{}_k$.

2.i) M is multilinear because a dot product is linear in each factor.

M is a tensor because its values are real numbers (scalars). (It is of order 2.)

M is totally symmetric because $\vec{v}_1 \cdot \vec{v}_2 = \vec{v}_2 \cdot \vec{v}_1$.

2.ii) M is multilinear because a cross product is linear in each factor.

M is not a tensor because its values are vectors in V, not scalars.

M is totally antisymmetric because $\vec{v}_1 \times \vec{v}_2 = -\vec{v}_2 \times \vec{v}_1$.

2.iii) M is not multilinear because $M(\vec{v}_1, 2\vec{v}_2) = 4M(\vec{v}_1, \vec{v}_2)$.

M is not a tensor and is neither totally symmetric nor totally antisymmetric because $M(\vec{v}_1, \vec{v}_2)$ is generally $\neq M(\vec{v}_2, \vec{v}_1)$ and $\neq -M(\vec{v}_2, \vec{v}_1)$.

2.iv) $M(\vec{v}_1, \vec{v}_2) = 0$, so M is multilinear.

M is a tensor because its values are real numbers. (It is of order 2.)

M is both totally symmetric and totally antisymmetric because

$$M(\vec{v}_1, \vec{v}_2) = M(\vec{v}_2, \vec{v}_1) = -M(\vec{v}_2, \vec{v}_1)$$

for all $\vec{v}_1$, $\vec{v}_2 \in V$.

2.v) M is multilinear because it is linear in each of $\vec{v}_1$, $\vec{v}_2$, $\vec{v}_3$ when the other two are fixed.

M is a tensor because its values are scalars. (It is of order 3.)

M is totally antisymmetric. To show this, we must show that

$$(12)M = (13)M = (23)M = -M.$$

We have

$$M(\vec{v}_2, \vec{v}_1, \vec{v}_3) = (\vec{v}_2 \times \vec{v}_1) \cdot \vec{v}_3 = -(\vec{v}_1 \times \vec{v}_2) \cdot \vec{v}_3 = -M(\vec{v}_1, \vec{v}_2, \vec{v}_3)$$

so $(12)M = -M$. And we have

$$M(\vec{v}_3, \vec{v}_2, \vec{v}_1) = (\vec{v}_3 \times \vec{v}_2) \cdot \vec{v}_1 = (\vec{v}_2 \times \vec{v}_1) \cdot \vec{v}_3 = -M(\vec{v}_1, \vec{v}_2, \vec{v}_3)$$

so $(13)M = -M$. Finally

$$M(\vec{v}_1, \vec{v}_3, \vec{v}_2) = (\vec{v}_1 \times \vec{v}_3) \cdot \vec{v}_2 = (\vec{v}_2 \times \vec{v}_1) \cdot \vec{v}_3 = -M(\vec{v}_1, \vec{v}_2, \vec{v}_3)$$

so $(23)M = -M$.

2.vi) M is multilinear because it is linear in each of $\vec{v}_1$, $\vec{v}_2$, $\vec{v}_3$, $\vec{v}_4$ when the other three are fixed.

M is a tensor because its values are scalars. (It is of order 4.)

M is neither totally symmetric nor totally antisymmetric. It is true that $(12)M = M$ and $(34)M = M$. However, $(23)M$ is neither M nor $-M$. To see this, we must find $\vec{u}_1$, $\vec{u}_2$, $\vec{u}_3$, $\vec{u}_4 \in V$ such that $M(\vec{u}_1, \vec{u}_3, \vec{u}_2, \vec{u}_4) \neq M(\vec{u}_1, \vec{u}_2, \vec{u}_3, \vec{u}_4)$, and we must find $\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4 \in V$ such that $M = (\vec{v}_1, \vec{v}_3, \vec{v}_2, \vec{v}_4) \neq -M = (\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4)$. Let $\hat{x}, \hat{y}, \hat{z}$ be an orthonormal basis for V and let $\vec{u}_1 = \vec{u}_2 = \vec{v}_1 = \vec{v}_2 = \hat{x}$, $\vec{u}_3 = \vec{u}_4 = \vec{v}_3 = \vec{v}_4 = \hat{y}$. Then $M(\vec{u}_1, \vec{u}_2, \vec{u}_3, \vec{u}_4) = 1$ and $M(\vec{u}_1, \vec{u}_2, \vec{u}_3, \vec{u}_4) = 0$. [Note: One might be tempted to say that $(\vec{v}_1 \cdot \vec{v}_2)\ (\vec{v}_3 \cdot \vec{v}_4)$ is "obviously" different from $\pm(\vec{v}_1 \cdot \vec{v}_3)(\vec{v}_2 \cdot \vec{v}_4)$. Yet if $\dim V = 1$, we do have $(\vec{v}_1 \cdot \vec{v}_2)(\vec{v}_3 \cdot \vec{v}_4) = (\vec{v}_1 \cdot \vec{v}_3)(\vec{v}_2 \cdot \vec{v}_4)$, and M is totally symmetric.

Extra Problems

In what follows, V is an n-dimensional Euclidean vector space and $L \in \mathcal{L}(V \to V)$.

1E. Recall that $l \in \mathcal{R}$ is an "eigenvalue" of L iff there is a *nonzero* $\vec{v} \in V$ such that $L(\vec{v}) = l\vec{v}$.

a) Show that if l is an eigenvalue of L and p is any positive integer, then l^p is an eigenvalue of L^p.

b) If $\pi \in S_q$ (i.e., $\pi : \{1, \cdots, q\} \to \{1, \cdots, q\}$ is a bijection), show that there is a positive integer $p \leq q!$ such that $\pi^p = e$, the identity permutation.

c) Regard $\pi \in S_q$ as a linear operator on $\otimes^q V$. Suppose $c \in \mathcal{R}$ is an eigenvalue of $\pi : \otimes^q V \to \otimes^q V$. Show that $c = \pm 1$.

2E.

a) For any $c \in \mathcal{R}$, show that $\det(cL) = c^n \det L$.

b) Recall that L^{-1} exists iff the only solution $\vec{u}$ of $L(\vec{u}) = \vec{0}$ is $\vec{u} = \vec{0}$. Use this fact to give a pedantically complete proof that $l \in \mathcal{R}$ is an eigenvalue of L iff $\det(L - lI_V) = 0$, where I_V is the identity operator on V.

3E.

a) Show that $\det(L - lI_V)$ is a polynomial in l with degree $\leq n$. That is, $\det(L - lI_V) = c_0 l^n + c_1 l^{n-1} + \cdots + c_n$. [Hint: Choose a basis $B = (\vec{b}_1, \cdots, \vec{b}_n)$ for V. Let $L_i{}^j$ be the matrix of L relative to B, i.e., $L(\vec{b}_i) = L_i{}^j \vec{b}_j$. Show that relative to B the matrix of $L - lI_V$ is $L_i{}^j - l\delta_i{}^j$. Finally, use $\det L = L_1{}^{i_1} \cdots L_n{}^{i_n} \varepsilon_{i_1 \cdots i_n}$ with $L - lI_V$ replacing L.]

b) Show that $c_0 = (-1)^n$ and $c_n = \det L$. (Thus the degree of $\det(L - lI_V)$ is n.)

c) $(-1)^{n-1} c_1$ is called the trace of L, written tr L. Calculate tr L in terms of the matrix $L_i{}^j$ of L relative to B.

Solutions to Extra Problems

1E.

a) By hypothesis, $\vec{v} \neq \vec{0}$ and $L(\vec{v}) = l\vec{v}$. Then $L^2(\vec{v}) = L[L(\vec{v})] = L(l\vec{v}) = lL(\vec{v}) = l^2\vec{v}$. And $L^3(\vec{v}) = L[L^2(\vec{v})] = L(l^2\vec{v}) = l^2 L(\vec{v}) = l^3\vec{v}$. Etc. (or use mathematical induction on p).

b) S_q has $q!$ members. Therefore, among the $q!+1$ permutations $e, \pi, \pi^2, \pi^3, \cdots, \pi^{q!}$, two must be the same, say $\pi^a = \pi^b, a < b$. Then $\pi^{-a}\pi^a = e = \pi^{-a}\pi^b = \pi^{b-a}$. Take $p = b - a$.

c) If c is an eigenvalue of π, and p is as in b) above, then c^p is an eigenvalue of π^p, so there is a nonzero $M \in \otimes^q V$ such that $eM = c^p M$. But $eM = M$, so $c^p = 1$. Since $c \in \mathcal{R}$, $c = \pm 1$.

2E.

a) Let $A \in \Lambda^n V$, $A \neq 0$. Let $(\vec{b}_1, \cdots, \vec{b}_n) = B$ be a basis for V. Then $A[(cL)(\vec{b}_1), \cdots, (cL)(\vec{b}_n)]$ $= \det(cL) A(\vec{b}_1, \cdots, \vec{b}_n)$. But $A[(cL)(\vec{b}_1), \cdots, (cL)(\vec{b}_n)] = A[cL(\vec{b}_1), \cdots, cL(\vec{b}_n)]$ $= c^n A[L(\vec{b}_1), \cdots, L(\vec{b}_n)] = c^n (\det L) A(\vec{b}_1, \cdots, \vec{b}_n)$. Since $A(\vec{b}_1, \cdots, \vec{b}_n) \neq 0$, $\det(cL) = c^n (\det L)$.

b) l is an eigenvalue of $L \Leftrightarrow \exists \vec{v} \in V \ni \vec{v} \neq \vec{0}$ and $L(\vec{v}) = l\vec{v}$. (Definition of eigenvalue.)

$$\begin{aligned} &\Leftrightarrow \quad L(\vec{v}) - lI_V(\vec{v}) = \vec{0} && \text{(def. of } I_V) \\ &\Leftrightarrow \quad (L - lI_V)(\vec{v}) = \vec{0} && \text{(def. of } L - lI_V). \end{aligned}$$

But $\exists \vec{v} \in V \ni \vec{v} \neq \vec{0}$ and $(L - lI_V)(\vec{v}) = \vec{0} \Leftrightarrow \det(L - lI_V) = 0$.

3E.

a) $(L - lI_V)(\vec{b}_i) = L(\vec{b}_i) - lI_V(\vec{b}_i) = L_i{}^j \vec{b}_j = l\vec{b}_i = (L_i{}^j - l\delta_i{}^j)\vec{b}_j$, so the matrix of $L - lI_V$ relative to B is $L_i{}^j - l\delta_i{}^j$. * Therefore $\det(L - lI_V) = (L_l{}^{i_1} - l\delta_l{}^{i_1}) \cdots (L_n{}^{i_n} - l\delta_n{}^{i_n})\varepsilon_{i_1 \cdots i_n}$. This is a sum of $n!$ terms. Each term is a product of n monomials in l, whose leading coefficients are 0 or -1. Therefore each term is a polynomial in l of degree $\leq n$. Hence so is the sum of all $n!$ terms.

b) c_n is the value of the polynomial when $l = 0$. Hence it is the value of $\det(L - lI_V)$ when $l = 0$, i.e., it is $\det L$.

b and c) To find c_0 and c_1, note that (*) in the solution 3E.a) is a sum of n^n terms, of which $n!$ have a nonzero factor $\varepsilon_{i_1 \cdots i_n}$. Among these nonzero terms, one has $i_1 = 1$, $i_2 = 2, \cdots, i_n = n$. It is (†)

$$\begin{aligned} \left(L_1{}^1 - l\delta_l{}^l\right) \quad \cdots \quad & \left(L_{(n)}{}^{(n)} - l\delta_{(n)}{}^{(n)}\right)\varepsilon_{1 \cdots n} \\ = \quad & \left(L_1{}^1 - l\right)\left(L_2{}^2 - l\right) \cdots \left(L_{(n)}{}^{(n)} - l\right). \end{aligned}$$

In all other nonzero terms, there must be at least two failures among $i_1 = 1$, $i_2 = 2, \cdots, i_n = n$, because $\{i_1, \cdots, i_n\} = \{1, \cdots, n\}$. Therefore, in each nonzero term in (*) except (†) above, at least two factors have no l, so the total degree as a polynomial in l is $\leq n-2$. Therefore, all the terms in l^n and l^{n-1} in $\det(L - lI_V)$ come from (†). But (†) is

$$(-1)^n l^n + (-1)^{n-l} \left[L_1{}^1 + L_2{}^2 + \cdots + L_{(n)}{}^{(n)} \right] l^{n-1} + \cdots .$$

Thus $c_0 = (-1)^n$ and $c_1 = (-1)^{n-1} L_i{}^i$. Hence $\operatorname{tr} L = L_i{}^i$.

Set Two

Exercise 3

Let $(\hat{x}_1, \hat{x}_2, \cdots, \hat{x}_n) = B$ be an ordered orthonormal basis for Euclidean space V. Let $B' = (\hat{x}_2, \hat{x}_1, \hat{x}_3, \cdots, \hat{x}_n)$ (i.e., just interchange $\hat{x}_1$ and $\hat{x}_2$). Show that B and B' are oppositely oriented.

Exercise 4

If V is Euclidean vector space, the "identity tensor" on V is defined to be that $\mathbf{I} \in V \otimes V$ such that for any $\vec{x}, \vec{y} \in V$

$$\mathbf{I}\,(\vec{x}, \vec{y}) = \vec{x} \cdot \vec{y}.$$

Suppose that $\vec{u}_1, \cdots \vec{u}_n$, $\vec{v}_1, \cdots \vec{v}_n$ are fixed vectors in V such that

$$\mathbf{I} = \sum_{i=1}^{n} \vec{u}_i \vec{v}_i.$$

a) Prove that $n \geq \dim V$. (Hence, if $\dim V \geq 2$, $\mathbf{I}$ is not a dyad).

b) Prove that if $n = \dim V$ then $(\vec{u}_1, \cdots, \vec{u}_n)$ is the basis of V dual to $(\vec{v}_1, \cdots \vec{v}_n)$.

Solutions

3. Let A be one of the two unimodular alternating tensors over V. Let $\pi \in S_n$. Then $A(\hat{x}_{\pi(1)}, \cdots, \hat{x}_{\pi(n)}) = (\operatorname{sgn} \pi) A(\hat{x}_1, \cdots, \hat{x}_n)$. Hence $(\hat{x}_1, \cdots, \hat{x}_n)$ and $(\hat{x}_{\pi(1)}, \cdots, \hat{x}_{\pi(n)})$ have the same or opposite orientation according as $\operatorname{sgn} \pi = +1$ or -1.

4. Change notation so as to use our restricted summation convention. Write $\mathbf{I} = \vec{u}_i \vec{v}^i$. Then for any $\vec{x}, \vec{y} \in V$ we have

$$\vec{x} \cdot \vec{y} = \mathbf{I}(\vec{x}, \vec{y}) = (\vec{u}_i \cdot \vec{x})(\vec{v}^i \cdot \vec{y}) = \vec{x} \cdot \left[\vec{u}_i \left(\vec{v}^i \cdot \vec{y}\right)\right]. \tag{23}$$

Thus

$$\vec{x} \cdot \left[\vec{y} - \left(\vec{y} \cdot \vec{v}^i\right) \vec{u}_i\right] = 0. \tag{24}$$

For any fixed $\vec{y} \in V$, (24) is true for all $\vec{x} \in V$. Thus $\vec{y} - (\vec{y} \cdot \vec{v}^i)\vec{u}_i = \vec{0}$, or

$$\vec{y} = \left(\vec{y} \cdot \vec{v}^i\right) \vec{u}_i, \qquad \text{for any } \vec{y} \in V. \tag{25}$$

a) (25) shows that $\{\vec{u}_1, \cdots, \vec{u}_n\}$ spans V. Hence, $n \geq \dim V$.

b) If $n = \dim V$, then $(\vec{u}_1, \cdots, \vec{u}_n)$ is an ordered basis for V. Let $(\vec{u}^1, \cdots, \vec{u}^n)$ be its dual basis. Choose fixed $j, k \in \{1, \cdots, n\}$ and set $\vec{x} = \vec{u}^j$, $\vec{y} = \vec{u}_k$ in (23). Then

$$\vec{u}^j \cdot \vec{u}_k = \left(\vec{u}_i \cdot \vec{u}^j\right) \left(\vec{v}^i \cdot \vec{u}_k\right)$$

or

$$\delta^j{}_k = \delta_i{}^j \left(\vec{v}^i \cdot \vec{u}_k\right) = \vec{v}^j \cdot \vec{u}_k.$$

Thus $(\vec{v}^1, \cdots, \vec{v}^n)$ is a sequence dual to $(\vec{u}_1, \cdots, \vec{u}_n)$. Since $(\vec{u}_1, \cdots, \vec{u}_n)$ is a basis for V, there is exactly one such dual sequence, namely $(\vec{u}^1, \cdots, \vec{u}^n)$. Hence $\vec{v}^i = \vec{u}^i$.

Exercise 5

Let V be an n-dimensional Euclidean space. Let A be any nonzero alternating tensor over V (that is, $A \in \Lambda^n V$ and $A \neq 0$). Let $(\vec{b}_1, \cdots, \vec{b}_n)$ be an ordered basis for V, with dual basis $(\vec{b}^1, \cdots, \vec{b}^n)$. Let $\triangle = A(\vec{b}_1, \cdots, \vec{b}_n)$. For each $i \in \{1, \cdots, n\}$ define

$$\vec{v}^i = \frac{1}{\triangle}(-1)^{i-1} A\langle n-1\rangle \left(\vec{b}_1 \cdots \cancel{\vec{b}_i} \cdots \vec{b}_n\right).$$

Show that $\vec{v}^i = \vec{b}^i$. When $n = 3$, express $(\vec{b}^1, \vec{b}^2, \vec{b}^3)$ explicitly in terms of $(\vec{b}_1, \vec{b}_2, \vec{b}_3)$, using only the dot and cross products, not A.

Exercise 6

Suppose $L \in \mathcal{L}(V \to V)$.

a) Show that there are unique mappings S, $K \in \mathcal{L}(V \to V)$ such that S is symmetric ($S^T = S$) K is antisymmetric ($K^T = -K$) and $L = S+K$. (Hint: then $L^T = S-K$.)

b) Suppose $\dim V = 3$ and A is one of the two unimodular alternating tensors over V. Suppose $K \in \mathcal{L}(V \to V)$ is antisymmetric. Show that there is a unique vector $\vec{k} \in V$ such that $\overleftrightarrow{K} = A \cdot \vec{k}$. Show that $\vec{k} = \frac{1}{2} A\langle 2 \rangle \overleftrightarrow{K}$ and that $\overleftrightarrow{K} = \vec{k} \cdot A$. [Hint: Take components relative to a positively-oriented ordered orthonormal basis. Use (3.1.2) and page D-9.]

Exercise 7

Let $(\vec{b}_1, \cdots, \vec{b}_n)$ be a basis for Euclidean vector space V, its dual basis being $(\vec{b}^1, \cdots, \vec{b}^n)$. Let $L \in \mathcal{L}(V \to V)$.

a) Show that $\overleftrightarrow{L} = \vec{b}_i L(\vec{b}^i) = \vec{b}^i L(\vec{b}_i)$.

b If L is symmetric, show that V has an orthonormal basis $\hat{x}_1, \cdots, \hat{x}_n$ with the property that there are real numbers $l_1, \cdots, l_n$ such that $\overleftrightarrow{L} = \sum_{v=1}^n l_{(v)} \hat{x}_v \hat{x}_v$. [Hint: Let $l_{(v)}$ be the eigenvalues of L. See page D-27.]

Solutions

5.a) We show that $\vec{b}_j \cdot \vec{v}^{\,i} = \delta_j{}^i$, and then appeal to the fact that the dual sequence of an ordered basis is unique. We have

$$\begin{aligned} \vec{b}_j \cdot \vec{v}^{\,i} &= \frac{1}{\triangle} (-1)^{i-1} \vec{b}_j \cdot A\langle n-1 \rangle \left(\vec{b}_1, \cdots \cancel{\vec{b}_i} \cdots \vec{b}_n \right) \\ &= \frac{(-1)^{i-1}}{\triangle} A\left(\vec{b}_j, \vec{b}_1, \cdots, \vec{b}_{i-1}, \vec{b}_{i+1}, \cdots, \vec{b}_n \right). \end{aligned}$$

It takes $i-1$ interchanges to change $(\vec{b}_j, \vec{b}_1, \cdots, \vec{b}_{i-1}, \vec{b}_{i+1}, \cdots, \vec{b}_n)$ into $(\vec{b}_1, \cdots, \vec{b}_{i-1}, \vec{b}_j, \vec{b}_{i+1}, \cdots, \vec{b}_n)$, and each changes the sign of A, so

$$\vec{b}_j \cdot \vec{v}^{\,i} = \frac{1}{\triangle} A\left(\vec{b}_1, \cdots, \vec{b}_{i-1}, \vec{b}_j, \vec{b}_{i+1}, \cdots, \vec{b}_n\right).$$

If $j \neq i$, $\vec{b}_j$ appears twice, so $A(\vec{b}_1, \cdots, \vec{b}_{i-1}, \vec{b}_j, \vec{b}_{i+1}, \cdots, \vec{b}_n) = 0$ and $\vec{b}_j \cdot \vec{v}^{\,i} = 0$. If $j = i$, then $A(\vec{b}_1, \cdots, \vec{b}_{i-1}, \vec{b}_i, \vec{b}_{i+1}, \cdots, \vec{b}_n)$ is $\triangle$, so $\vec{b}_j \cdot \vec{v}^{\,i} = 1$. QED.

5.b) When $n = 3$, $\triangle = A(\vec{b}_1, \vec{b}_2, \vec{b}_3) = \vec{b}_1 \cdot A\langle 2\rangle(\vec{b}_2\vec{b}_3) = \vec{b}_1 \cdot (\vec{b}_2 \times \vec{b}_3)$, and

$$\begin{aligned}
\vec{b}^{\,1} &= \frac{1}{\triangle} A\langle 2\rangle(\vec{b}_2\vec{b}_3) = \frac{1}{\triangle}\vec{b}_2 \times \vec{b}_3 \\
\vec{b}^{\,2} &= -\frac{1}{\triangle} A\langle 2\rangle(\vec{b}_2\vec{b}_3) = -\frac{1}{\triangle}\vec{b}_1 \times \vec{b}_3 = \frac{1}{\triangle}\vec{b}_3 \times \vec{b}_1 \\
\vec{b}^{\,3} &= \frac{1}{\triangle} A\langle 2\rangle(\vec{b}_1\vec{b}_2) = \frac{1}{\triangle}\vec{b}_1 \times \vec{b}_2.
\end{aligned}$$

6.a) **Uniqueness:** If $L = S + K$ with $S = S^T$ and $K^T = -K$ then $L^T = S - K$ so $S = 1/2(L + L^T)$ and $K = 1/2(L - L^T)$.

Existence: define $S = 1/2(L + L^T)$ and $K = 1/2(L - L^T)$. Then $S^T = S$, $K^T = -K$, and $L = S + K$.

6.b) **Uniqueness:** Suppose $\overleftrightarrow{K} = A \cdot \vec{k}$. Relative to any pooob we have $K_{ij} = \varepsilon_{ijk}k^k$ so $K_{ij}\varepsilon^{ijl} = \varepsilon^{ijl}\varepsilon_{ijk}k^k = 2\delta^l{}_k k^k = 2k^l$. Thus, $2\vec{k} = \overleftrightarrow{K}\langle 2\rangle A$. Also $2k^l = \varepsilon^{lij}K_{ij}$, so $2\vec{k} = A\langle 2\rangle \overleftrightarrow{K}$.

Existence: Given $\overleftrightarrow{K}$ antisymmetric, define $\vec{k} = 1/2A\langle 2\rangle \overleftrightarrow{K}$. We have relative to any pooob

$$\begin{aligned}
\left(A \cdot \vec{k}\right)_{ij} &= \varepsilon_{ijk}k^k = \frac{1}{2}\varepsilon^{klm}K_{lm} \\
&= \frac{1}{2}\left(\delta_i{}^l\delta_j{}^m - \delta_i{}^m\delta_j{}^l\right)K_{lm} \\
&= \frac{1}{2}K_{ij} - \frac{1}{2}K_{ji} = K_{ij}.
\end{aligned}$$

Hence, $\overleftrightarrow{K} = A \cdot \vec{k}$. Also $K_{ij} = \varepsilon_{ijk}k^k = k^k\varepsilon_{kij}$ so $\overleftrightarrow{K} = \vec{k} \cdot A$.

7.a) $\vec{b}^i \cdot \overleftrightarrow{L} = L(\vec{b}^i)$. Multiply by $\vec{b}_i$ and sum over i to get $\vec{b}_i L(\vec{b}^i) = \vec{b}_i[\vec{b}^i \cdot \overleftrightarrow{L}] = (\vec{b}_i\vec{b}^i) \cdot \overleftrightarrow{L} = \overleftrightarrow{I}_V \cdot \overleftrightarrow{L} = \overleftrightarrow{L}$. By viewing $\vec{b}^i$ as the original and $\vec{b}_i$ as the dual basis, one then also finds $\vec{b}^i L(\vec{b}_i) = \overleftrightarrow{L}$.

7.b) From page D-27, V has an orthonormal basis $\hat{x}_1, \cdots, \hat{x}_n$ consisting of eigenvectors of L. That is, there are numbers $l_1, \cdots, l_n \in \mathcal{R}$ such that $L(\hat{x}_1) = l_1\hat{x}_1, \cdots, L(\hat{x}_n) = l_n\hat{x}_n$. By 7a)

$$\overleftrightarrow{L} = \sum_{v=1}^{n} \hat{x}_v L(\hat{x}_v) = \sum_{v=1}^{n} \hat{x}_v \left(l_v \hat{x}_v\right) = \sum_{v=1}^{n} l_v \left(\hat{x}_v \hat{x}_v\right).$$

Extra problems

4E. Suppose $\{\vec{b}_1, \cdots, \vec{b}_m\}$ is a linearly independent subset of Euclidean space V (so $m \leq \dim V$; $m < \dim V$ is possible!). Let $(\vec{u}^1, \cdots, \vec{u}^m)$ and $(\vec{v}^1, \cdots, \vec{v}^m)$ be any m-tuples of vectors in V. Show that $\vec{u}^i \vec{b}_i = \vec{v}^i \vec{b}_i$ implies $\vec{u}^i = \vec{v}^i$.

5E. Suppose V is a Euclidean vector space and a, b, c, d, e are non-negative integers. Suppose $A \in \otimes^a V$, $C \in \otimes^c V$, $E \in \otimes^e V$. We want to study whether it is true that

$$A\langle b\rangle \left(C\langle d\rangle E\right) = \left(A\langle b\rangle C\right) \langle d\rangle E. \tag{26}$$

i) Using p 54 of the notes, find conditions on a, b, c, d, e which imply (26).

ii) If $\dim V \geq 2$, show that whenever a, b, c, d, e violate the conditions found in i), it is possible to have $A \in \otimes^a V$, $C \in \otimes^c V$ and $E \in \otimes^l V$ which violate (26). (Hint: use polyads.)

6E. In oriented Euclidean 3-space (V, A), show that $\vec{u}$, $\vec{v}$, $\vec{w}$ can be chosen so that $\vec{u} \times (\vec{v} \times \vec{w}) \neq (\vec{u} \times \vec{v}) \times \vec{w}$.

Solutions to Extra Problems

4E. Let $U = sp\{\vec{b}_1, \cdots, \vec{b}_m\}$. Then U is a subspace of V and hence a Euclidean vector space under V's dot product. Also, $(\vec{b}_1, \cdots, \vec{b}_m)$ is an ordered basis for U, so in U it has a dual basis $(\vec{b}^1, \cdots, \vec{b}^m)$. If $\vec{u}^{\,i}\vec{b}_i = \vec{v}^{\,i}\vec{b}_i$ then $(\vec{u}^{\,i}\vec{b}_i) \cdot \vec{b}^j = (\vec{v}^{\,i}\vec{b}_i) \cdot \vec{b}^j$, so $\vec{u}^{\,i}(\vec{b}_i \cdot \vec{b}^j) = \vec{v}^{\,i}(\vec{b}_i \cdot \vec{b}^j)$, so $\vec{u}^{\,i}\delta_i{}^j = \vec{v}^{\,i}\delta_i{}^j$. Then $\vec{u}^{\,j} = \vec{v}^{\,j}$. QED.

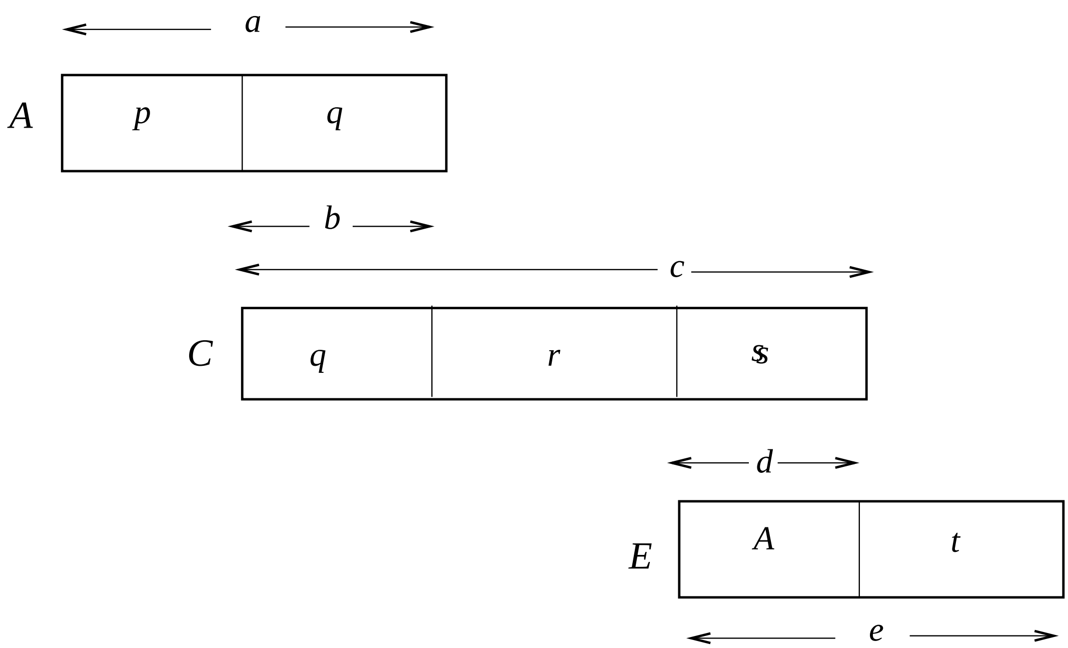


Figure Ex-1: Exercise 5E.i

5E.i) The picture that makes (26) work is in figure Ex-1 There must be $p, q, r, s, t \geq 0$ such that $a = p + q$, $b = q$, $c = q + r + s$, $d = s$, $e = s + t$. solving for p, q, r, s, t gives $p = a - b$, $q = b$, $r = c - b - d$, $s = d$, $t = e - d$. The conditions that $p, q, r, s, t, a, b, c, d, e \geq 0$ reduce to

$$0 \leq b \leq a, \quad 0 \leq d \leq e, \quad b + d \leq c.$$

5E.ii) The question is, is it possible to violate the inequalities in i) in such a way that $A\langle b\rangle(C\langle d\rangle E)$ and $(A\langle b\rangle C)\langle d\rangle E$ are both defined. Let $\langle\|\rangle(P) =$ order of tensor P. Thus $\langle\|\rangle(A) = a$, $\langle\|\rangle(A\langle b\rangle C) = a + c - b$, etc. A product $P\langle q\rangle R$ is defined only if $\langle\|\rangle(P) \geq q$ and $\langle\|\rangle(R) \geq q$. If $A\langle b\rangle(C\langle d\rangle E)$ is defined we must have $a \geq b$, $\langle\|\rangle(C\langle d\rangle E) = c + e - 2d \geq b$ and $c \geq d$, $e \geq d$. If $(A\langle b\rangle C)\langle d\rangle E$ is defined we must have $d \leq e$ and $\langle\|\rangle(A\langle b\rangle C) \geq d$, or $a + c - 2b \geq d$, and also $b \leq a$, $b \leq c$. If both triple products are defined we must have $b \leq a$, $b \leq c$, $d \leq c$, $d \leq e$, $b + 2d \leq c + e$ and $2b + d \leq a + c$. The only way i) can fail if all these hold is to have $b + d > c$. Choose $\vec{\alpha}_1, \cdots, \vec{\alpha}_a$, $\vec{\gamma}_1, \cdots \vec{\gamma}_c$, $\vec{\varepsilon}_1, \cdots \vec{\varepsilon}_e \in V$ and let $A = \vec{\alpha}_1 \cdots \vec{\alpha}_a$, $C = \vec{\gamma}_1 \cdots \vec{\gamma}_c$, $E = \vec{\varepsilon}_1 \cdots \vec{\varepsilon}_e$. Then we have figure Ex-5ii And $A\langle b\rangle C = (\vec{\alpha}_{a-b+1} \cdot \vec{\gamma}_1) \cdots (\vec{\alpha}_a \cdot$

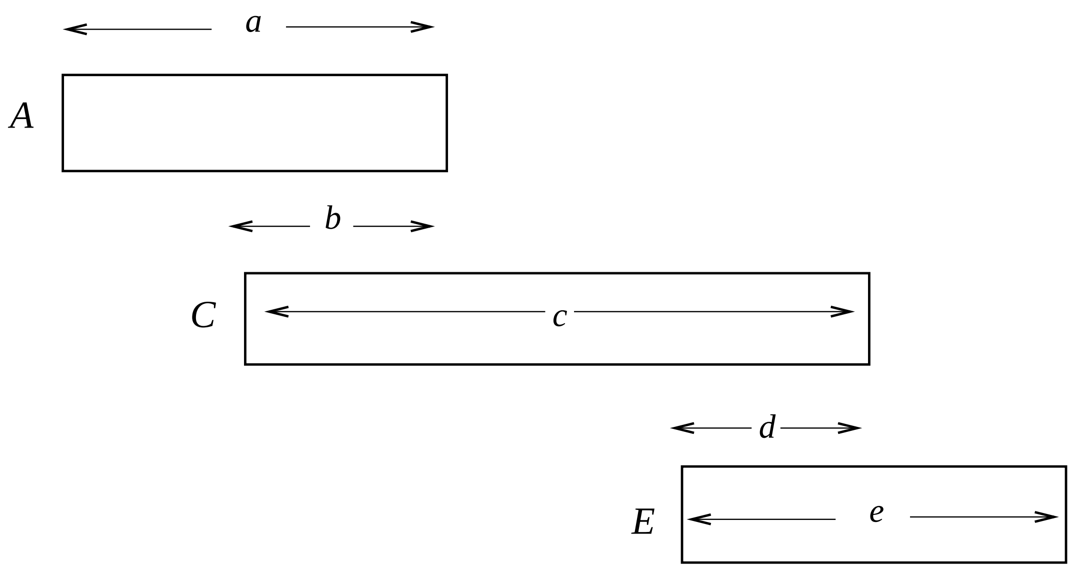


Figure Ex-2: Figure Ex-5.ii

$\vec{\gamma}_b)\vec{\alpha}_1\cdots\vec{\alpha}_{a-b}\vec{\gamma}_{b+1}\cdots\vec{\gamma}_c$ $(A\langle b\rangle C)\langle d\rangle E = (\vec{\alpha}_{a-b+1}\cdot\vec{\gamma}_1)\cdots(\vec{\alpha}_a\cdot\vec{\gamma}_b)(\vec{\alpha}_{a+c-2b-d+1}\cdot\vec{\varepsilon}_1)$ $\cdots(\vec{\alpha}_{a-b}\cdot\vec{\varepsilon}_{d+b-c})$ $(\vec{\gamma}_{b+1}\cdot\vec{\varepsilon}_{d+b-c+1})\cdots(\vec{\gamma}_c\cdot\vec{\varepsilon}_d)$ $\vec{\alpha}_1\cdots\vec{\alpha}_{a+c-2b-d}$ $\vec{\varepsilon}_{d+1}\cdots\vec{\varepsilon}_e$.

Similarly $C\langle d\rangle E = (\vec{\gamma}_{c-dn}\cdot\vec{\varepsilon}_1)\cdots(\vec{\gamma}_c\cdot\vec{\varepsilon}_d)\vec{\gamma}_1\cdots\vec{\gamma}_{c-d}\vec{\varepsilon}_{d+1}\cdots\vec{\varepsilon}_e$ and $A\langle b\rangle(C\langle d\rangle E) =$ $(\vec{\alpha}_{a-b+1}\cdot\vec{\gamma}_1)\cdots(\vec{\alpha}_{a-b+c-d}\cdot\vec{\delta}_{c-d})(\vec{\alpha}_{a-b+c-d+1}\cdot\vec{\varepsilon}_{d+1})\cdots(\vec{\alpha}_a\cdot\vec{\epsilon}_{b+2d-c})$ $(\vec{\gamma}_{c-d+1}\cdot\vec{\epsilon}_1)\cdots(\vec{\gamma}_c\cdot\vec{\epsilon}_d)\vec{\alpha}_1\cdots\vec{\alpha}_{a-b}\vec{\epsilon}_{b+2d-c+1}\cdots\vec{\varepsilon}_e$.

Choose all $\vec{\alpha}$, $\vec{\gamma}$, $\vec{\varepsilon}$ so the dot products are $\neq 0$. Now $a-b > a+c-2b-d$, so $(A\langle b\rangle C)\langle d\rangle E = A\langle b\rangle(C\langle d\rangle E)$ implies that $\vec{\alpha}_{a-b}$ and $\vec{\varepsilon}_{2d+b-c}$ are linearly dependent. By choosing them linearly independent we have $(A\langle b\rangle C)\langle d\rangle E \neq A\langle b\rangle(C\langle d\rangle E)$.

6E.

$$\begin{aligned}
\vec{u}\times(\vec{v}\times\vec{w}) &= (\vec{u}\cdot\vec{w})\vec{v}-(\vec{u}\cdot\vec{v})\vec{w}\\
(\vec{u}\times\vec{v})\times\vec{w} &= \vec{w}\times(\vec{v}\times\vec{u}) = (\vec{w}\cdot\vec{u})\vec{v}-(\vec{w}\cdot\vec{v})\vec{u}.
\end{aligned}$$

These are equal iff $(\vec{u}\cdot\vec{v})\vec{w} = (\vec{v}\cdot\vec{w})\vec{u}$. Choose $\vec{u}$ and $\vec{w}$ mutually $\perp$ and of unit length, with $\vec{v} = \vec{u}+\vec{w}$, and this equation will be violated.

Exercise 8

Let U be a Euclidean vector space. Let D be an open subset of U. Let $\overleftrightarrow{T} \in U \otimes U$. For each $\vec{u} \in U$, let $\vec{f}(\vec{u}) = \overleftrightarrow{T} \cdot \vec{u}$ and $g(\vec{u}) = \vec{u} \cdot \vec{u}$. Show that at every $\vec{u} \in D$, $\vec{f} : D \to U$ and $g : D \to \mathcal{R}$ are differentiable. Do so by finding $\vec{\nabla}\vec{f}(\vec{u})$, $\vec{\nabla}g(\vec{u})$, and the remainder functions, and by showing that $R(\vec{h}) \to 0$ as $\vec{h} \to \vec{0}$.

Exercise 9

Suppose that D_f and D_g are open subsets of Euclidean vector spaces U and V respectively. Suppose that $\vec{f} : D_f \to D_g$ and $\vec{g} : D_g \to D_f$ are inverse to one another. Suppose there is a $\vec{u} \in D_f$ where $\vec{f}$ is differentiable and that $\vec{g}$ is differentiable at $\vec{v} = \vec{f}(\vec{u})$. Show that $\dim U = \dim V$. [Hint: See page 108.]

Exercise 10

Suppose V is a Euclidean vector space, (U, A) is an oriented real three-dimensional Euclidean space, D is an open subset of U, and $\vec{f} : D \to U$ and $\vec{g} : D \to V$ are both differentiable at $\vec{u} \in D$.

i) Which proposition in the notes shows that $\vec{f}\vec{g} : D \to U \otimes V$ is differentiable at $\vec{u}$?

ii) Show that at $\vec{u}$

$$\vec{\nabla} \times \left(\vec{f}\vec{g}\right) = \left(\vec{\nabla} \times \vec{f}\right)\vec{g} - \vec{f} \times \vec{\nabla}\vec{g}.$$

[Note: In this equation, one of the expressions is undefined. Define it in the obvious way and then prove the equation.]

Solutions

8a.

$$\vec{f}\left(\vec{u} + \vec{h}\right) = \overleftrightarrow{T} \cdot \left(\vec{u} + \vec{h}\right) = \overleftrightarrow{T} \cdot \vec{u} + \overleftrightarrow{T} \cdot \vec{h}$$

$$= \vec{f}(\vec{u}) + \vec{h} \cdot \overset{\leftrightarrow}{T}^T .$$

Therefore $\vec{\nabla}\vec{f}(\vec{u}) = \overset{\leftrightarrow}{T}^T$ and $\vec{R}(\vec{h}) = \vec{0}$ for all $\vec{h}$.

8b.

$$\begin{aligned} g\left(\vec{u}+\vec{h}\right) &= \left(\vec{u}+\vec{h}\right)\cdot\left(\vec{u}+\vec{h}\right) \\ &= \vec{u}\cdot\vec{u} + 2\vec{h}\cdot\vec{u} + \vec{h}\cdot\vec{h} \\ &= g\left(\vec{u}\right) + \vec{h}\cdot[2\vec{u}] + \left\|\vec{h}\right\|\left\|\vec{h}\right\|. \end{aligned}$$

Therefore $\vec{\nabla}g(\vec{u}) = 2\vec{u}$, and $R(\vec{h}) = \left\|\vec{h}\right\|$. As $\vec{h} \to 0$, $R(\vec{h}) \to 0$ because $\vec{h} \to \vec{0}$ means $\left\|\vec{h}\right\| \to 0$.

9. Define $\overset{\leftrightarrow}{P} = \vec{\nabla}\vec{f}(\vec{u})$ and $\overset{\leftrightarrow}{Q} = \vec{\nabla}\vec{g}(\vec{v})$. By example 9.2.24, $\overset{\leftrightarrow}{P} \cdot \overset{\leftrightarrow}{Q} = \overset{\leftrightarrow}{I}_U$ and $\overset{\leftrightarrow}{Q} \cdot \overset{\leftrightarrow}{P} = \overset{\leftrightarrow}{I}_V$. Therefore, by 7.4.20, the linear mappings $P : U \to V$ and $Q : V \to U$ satisfy $Q \circ P = I_U$ and $P \circ Q = I_V$. By iii) on page D-6, P and Q are bijections. Since they are linear, they are isomorphisms of U to V and V to U. By preliminary exercise 3, $\dim U = \dim V$.

10.i) Method 1 on page 105 or method 2 on page 107, applied to $\vec{f}\,\hat{1} : D \to U \otimes \mathcal{R}^1$ and $\hat{1}\vec{g} : D \to \mathcal{R}^1 \otimes V$.

10.ii) For $\vec{u} \in U$ and $\overset{\leftrightarrow}{P} \in U \otimes V$ we define $\vec{u} \times \overset{\leftrightarrow}{P} := A\langle 2\rangle \left[\vec{u}\,\overset{\leftrightarrow}{P}\right]$. Then, relative to any basis $\vec{b}_1, \vec{b}_2, \vec{b}_3$ for U and $\vec{\beta}_1, \cdots, \vec{\beta}_n$ for V, we have $(\vec{u} \times \overset{\leftrightarrow}{P})^i{}_l = A^{ijk}\mu_j P_{kl}$. Then

$$\begin{aligned} \left[\vec{\nabla} \times \left(\vec{f}\vec{g}\right)\right]^i{}_l &= A^{ijk}\partial_j\left(\vec{f}\vec{g}\right)_{kl} = A^{ijk}\partial_j\left(f_k g_l\right) \\ &= A^{ijk}\left[\left(\partial_j f_k\right) g_l + f_k \partial_j g_l\right] \\ &= \left(\vec{\nabla}\times\vec{f}\right)^i g_l - A^{ikj} f_k \partial_j g_l \\ &= \left[\left(\vec{\nabla}\times\vec{f}\right)\vec{g}\right]^i{}_l - A^{ijk} f_j \left(\vec{\nabla}\vec{g}\right)_{kl} \\ &= \left[\left(\vec{\nabla}\times\vec{f}\right)\vec{g}\right]^i{}_l - \left(\vec{f}\times\vec{\nabla}\vec{g}\right)^i_l \\ &= \left[\left(\vec{\nabla}\times\vec{f}\right)\vec{g} - \vec{f}\times\vec{\nabla}\vec{g}\right]^i{}_l. \qquad \text{QED.} \end{aligned}$$

Extra Problems

7E. Let U be a Euclidean vector space. Let $\overleftrightarrow{T} \in U \otimes U$. Let D be the set of all $\vec{u} \in U$ such that $\vec{u} \cdot \overleftrightarrow{T} \cdot \vec{u} > 0$. If $\vec{u} \in D$, define $f(\vec{u}) = \ln(\vec{u} \cdot \overleftrightarrow{T} \cdot \vec{u})$.

i) Show that D is open.

ii) Show that $f : D \to \mathcal{R}$ is differentiable at every point $\vec{u} \in D$, and find $\vec{\nabla} f(\vec{u})$.

8E. Let V be the set of all infinite sequences of real numbers, $\vec{x} = (x_1, x_2, \cdots)$, such that $\sum_{n=1}^{\infty} x_n^2$ converges. If $\vec{y} = (y_1, y_2, \cdots)$ is also in V and $a, b \in \mathcal{R}$, define $a\vec{x} + b\vec{y} = (ax_1 + by_1, ax_2 + by_2, \cdots)$. Define $\vec{x} \cdot \vec{y} = \sum_{n=1}^{\infty} x_n y_n$. [The sum converges by the Cauchy test because, for any M and N, Schwarz's inequality implies $|\sum_{n=m}^{N} x_n y_n| \leq (\sum_{n=m}^{N} x_n^2)^{1/2} (\sum_{n=m}^{N} y_n^2)^{1/2}$. Furthermore, $|\vec{x} \cdot \vec{y}|^2 = |\lim_{N\to\infty} \sum_{n=1}^{N} x_n y_n|^2 \stackrel{\leq}{=} \lim_{N\to\infty} (\sum_{n=1}^{N} x_n^2)(\sum_{n=1}^{n} y_n^2) = \|\vec{x}\|^2 \|\vec{y}\|^2$, so Schwarz's inequality works in V.] V is a real dot-product space but not a Euclidean space, because it is not finite dimensional. If $\vec{v}_0, \vec{v}_1, \vec{v}_2, \cdots$ is a sequence of vectors in V, define $\lim_{n\to\infty} \vec{v}_n = \vec{v}_0$ by the obvious extension of definition 9.1.30, i.e. $\lim_{n\to\infty} \|\vec{v}_0 - \vec{v}_n\| = 0$.

i) Suppose $\lim_{n\to\infty} \vec{v}_n = \vec{v}_0$. Show that for any $\vec{x} \in V$, $\lim_{n\to\infty} \vec{x} \cdot \vec{v}_n = \vec{x} \cdot \vec{v}_0$.

ii) Construct a sequence $\vec{v}_1, \vec{v}_2, \cdots$ in V such that $\lim_{n\to\infty} \vec{x} \cdot \vec{v}_n = 0$ for every $\vec{x} \in V$, but $\lim_{n\to\infty} \vec{v}_n$ does not exist.

(For differentiability, V has two different definitions, leading to Frechet and Gateau derivation respectively based on norms and on components).

Solutions

7E. If $\overleftrightarrow{T} = \overleftrightarrow{0}$, D is empty and i) and ii) are trivial. Therefore assume $\| \overleftrightarrow{T} \| > 0$. Define $g(\vec{u}) = \vec{u} \cdot \overleftrightarrow{T} \cdot \vec{u}$.

i) $g(\vec{u}+\vec{h}) - g(\vec{u}) = \vec{h}\cdot \overleftrightarrow{T} \cdot\vec{u} + \vec{u}\cdot \overleftrightarrow{T} \cdot\vec{h}\cdot \overleftrightarrow{T} \cdot\vec{h}$. Therefore, by the triangle and generalized Schwarz inequalities,

$$\left|g\left(\vec{u}+\vec{h}\right) - g\left(\vec{u}\right)\right| \leq \|\vec{h}\|\| \overleftrightarrow{T} \| \left(2\|\vec{u}\| + \|\vec{h}\|\right).$$

Assume $\vec{u} \in D$ so $g(\vec{u}) > 0$. Then $\vec{u} \neq \vec{0}$. Let ε be the smaller of $\|\vec{u}\|$ and $g(\vec{u})/(3\|\vec{u}\|\| \overleftrightarrow{T} \|)$. If $\|\vec{h}\| < \varepsilon$ then $|g(\vec{u}+\vec{h}) - g(\vec{u})| < \varepsilon \| \overleftrightarrow{T} \|(2\|\vec{u}\| + \|\vec{h}\|)$ $< 3\varepsilon\| \overleftrightarrow{T} \|\|\vec{u}\| \leq g(\vec{u})$. Hence $g(\vec{u}+\vec{h}) - g(\vec{u}) > -g(\vec{u})$, so $g(\vec{u}+\vec{h}) > 0$. That is, if $\vec{u} \in D$ and $\varepsilon =$ smaller of $\vec{u}$ and $g(\vec{u})/(3\|\vec{u}\|\| \overleftrightarrow{T} \|)$ then the open ball $B(\vec{u}, \varepsilon) \subseteq D$. Hence D is open,

ii) By i) above, $g(\vec{u}+\vec{h}) = g(\vec{u}) + \vec{h}\cdot(\overleftrightarrow{T} + \overleftrightarrow{T}^T)\cdot\vec{u} + \|\vec{h}\|R(\vec{h})$ where $R(\vec{h}) = (\vec{h}\cdot \overleftrightarrow{T} \cdot\vec{h})/\|\vec{h}\|$ for $\vec{h} \neq \vec{0}$ and $= \vec{0}$ for $\vec{h} = \vec{0}$. Thus g is differentiable at $\vec{u}$, with $\vec{\nabla} g(\vec{u}) = (\overleftrightarrow{T} + \overleftrightarrow{T}^T)\cdot\vec{u}$. By the chain rule, $f \circ g$ is differentiable, with $f(v) = \ln v$, and its gradient is $\partial_v f(v)\vec{\nabla} g(\vec{u})$ where $v = g(\vec{u})$. This is $v^{-1}\vec{\nabla} g(\vec{u})$, so

$$\left[\vec{\nabla}\ln\left(\vec{u}\cdot \overleftrightarrow{T} \cdot\vec{u}\right)\right](\vec{u}) = \frac{\left(\overleftrightarrow{T} + \overleftrightarrow{T}^T\right)\cdot\vec{u}}{\vec{u}\cdot \overleftrightarrow{T} \cdot\vec{u}}.$$

8E. i)

$$|(\vec{x}\cdot\vec{v}_n - \vec{x}\cdot\vec{v}_0)| = |\vec{x}\cdot(\vec{v}_n - \vec{v}_0)| \leq \|\vec{x}\|\|\vec{v}_n - \vec{v}_0\| \to 0$$

if $\|\vec{v}_n - \vec{v}_0\| \to 0$.

8E. ii) Let $\vec{v}_n = (0, \cdots, 0, 1, 0, \cdots)$, all 0 except for a 1 in the n'th place. If $\vec{x} \in V$, $\vec{x}\cdot\vec{v}_n = x_n \to 0$ as $n \to \infty$ because $\sum_1^\infty x_n^2$ converges. But if $\|\vec{v}_n - \vec{v}\| \to 0$ for some $\vec{v} \in V$ then there is an N so large that if $n > N$ then $\|\vec{v}_n - \vec{v}\| < 1/2$. If m, $n > N$ then $\|\vec{v}_m - \vec{v}_n\| = \|\vec{v}_m - \vec{v} + \vec{v} - \vec{v}_n\| \leq \|\vec{v}_m - \vec{v}\| + \|\vec{v}_n - \vec{v}\| = 1$. But $\|\vec{v}_m - \vec{v}_n\| = \sqrt{2}$.

Exercise 11

A mass distribution m occupies a subset D of ordinary Euclidean 3-space U. Let $\hat{n}$ be any unit vector in U and denote by $\hat{n}\mathcal{R}$ the straight line through U's origin $\vec{0}$ in the direction of $\hat{n}$. Let $\tilde{w}\ (\vec{r})$ be the perpendicular distance of the point $\vec{r}$ from the axis $\hat{n}\mathcal{R}$, so that the moment of inertia of the mass distribution about that axis is $J(\hat{n}) = \int_D dm(\vec{r})\tilde{w}(\vec{r})^2$. Define $\overleftrightarrow{T}$: $D \to U \otimes U$ by requiring for each $\vec{r} \in D$ that $\overleftrightarrow{T}\ (\vec{r}) = r^2\ \overleftrightarrow{I}_U - \vec{r}\vec{r}$. Here $r^2 = \vec{r} \cdot \vec{r}$ and $\overleftrightarrow{I}_U$ is the identity tensor in $U \otimes U$. The tensor $\overleftrightarrow{J} = \int_D dm(\vec{r})\ \overleftrightarrow{T}\ (\vec{r})$ is called the inertia tensor of the mass distribution, and is usually written $\overleftrightarrow{J} = \int_D dm(\vec{r})[r^2\ \overleftrightarrow{I}_U - \vec{r}\vec{r}]$.

a) Show that $\overleftrightarrow{J}$ is symmetric.

b) Show that $J(\hat{n}) = \hat{n} \cdot \overleftrightarrow{J} \cdot \hat{n}$.

c) Show that to calculate $\overleftrightarrow{J}$ it is necessary to calculate only six particular integrals of real-valued functions on D with respect to the mass distribution m. (Once these six functions are calculated, the moment of inertia of the body about any axis is found from 11b.)

Exercise 12

In problem 11, suppose the mass distribution rotates rigidly about the axis $\hat{n}\mathcal{R}$ with angular velocity $\vec{\Omega} = \Omega\hat{n}$, so that the mass at position $\vec{r}$ has velocity $\vec{v}(\vec{r}) = \vec{\Omega} \times r$. The angular momentum of the mass distribution about $\vec{0}$ is defined as $\vec{L} = \int_D dm(\vec{r})[\vec{r} \times \vec{v}(\vec{r})]$, and the kinetic energy of the mass distribution is defined as $K = 1/2 \int_D dm(\vec{r})v(\vec{r})^2$, where v^2 means $\vec{v} \cdot \vec{v}$. Show the following:

a) $\vec{L} = \overleftrightarrow{J} \cdot \vec{\Omega}$

b) $K = 1/2\, \vec{\Omega} \cdot \overleftrightarrow{J} \cdot \vec{\Omega}$

c) $\vec{L}$ need not be $J(\hat{n})\vec{\Omega}$.

Exercise 13

The second moment tensor of the mass distribution in problem 11 is defined as $\overleftrightarrow{M} = \int_D dm(\vec{r})\, \vec{r}\,\vec{r}$.

a) Show that tr $\overleftrightarrow{M} = \int_D dm(\vec{r}) r^2$.

b) Express $\overleftrightarrow{J}$ in terms of $\overleftrightarrow{M}$, tr $\overleftrightarrow{M}$ and $\overleftrightarrow{I}_U$.

c) Express $\overleftrightarrow{M}$ in terms of $\overleftrightarrow{J}$, tr $\overleftrightarrow{J}$ and $\overleftrightarrow{I}_U$. ($\overleftrightarrow{J}$ can be observed by noting the reaction of the mass distribution to torques. Then 13c) gives $\overleftrightarrow{M}$ and 13a) gives $\int_D dm(\vec{r}) r^2$ from such observations).

d) If the mass distribution is a cube of uniform density ρ, side 2a and center at $\vec{0}$, find $J(\hat{n})$ when $\hat{n}\mathcal{R}$ is the axis through one of the corners of the cube.

Exercise 14

Let $\hat{y}_1$, $\hat{y}_2$, $\hat{y}_3$ be an orthonormal basis for physical space P. Let c be a fixed scalar. For any $\vec{r} \in P$, write $\vec{r} = r_i \hat{y}_i$. Consider a steady fluid motion whose Eulerian description $\vec{v}^E$ is given by

$$\vec{v}^E(\vec{r}, t) = r_1 \hat{y}_1 + c\, r_2 \hat{y}_2 = \vec{r} \cdot (\hat{y}_1 \hat{y}_1 + c \hat{y}_2 \hat{y}_2)$$

for all $\vec{r} \in P$ and all $t \in \mathcal{R}$.

a) Find the Lagrangian description of this motion, using t_0-position labels.

b) Show that the paths of the individual particles lie in planes r_3 = constant, and are hyperbolas if $c = -1$ and straight lines if $c = +1$.

c) Find the label function $\vec{\mathbf{x}}^E : P \times \mathcal{R} \to P$. (It will depend on which fixed t_0 is used to establish t_0-position labels.)

d) Find the Eulerian and Lagrangian descriptions of the particle acceleration $\vec{a}$.

Solutions

11 a) The permutation operator (12): $P \otimes P \to P \otimes P$ is linear, and (12) $\overleftrightarrow{T}(\vec{r}) = \overleftrightarrow{T}(\vec{r})$, so

$$(12) \quad \int_D dm(\vec{r})\, \overleftrightarrow{T}(\vec{r}) = \int_D dm(\vec{r}) \quad (12)\, \overleftrightarrow{T}(\vec{r}) = \int_D dm(\vec{r})\, \overleftrightarrow{T}(\vec{r}).$$

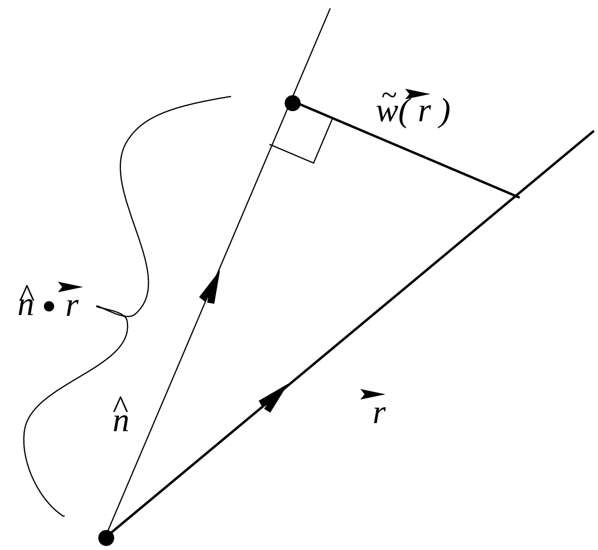


Figure Ex-3:

b)

$$\begin{aligned}
\tilde{w}\left(\vec{r}\right)^2 &= r^2 - \left(\hat{n}\cdot\vec{r}\right)^2 \\
&= r^2\hat{n}\cdot \overleftrightarrow{I}_P \cdot\hat{n} - \hat{n}\cdot\hat{r}\hat{r}\cdot\hat{n} \\
&= \hat{n}\cdot\left[r^2 \overleftrightarrow{I}_P - \vec{r}\,\vec{r}\right]\cdot\hat{n}.
\end{aligned}$$

Hence

$$J(\hat{n}) = \int_D dm(\vec{r})\hat{n}\cdot\left[r^2 I_P - \vec{r}\,\vec{r}\right]\cdot\hat{n} \tag{27}$$

$$= \hat{n}\cdot\left[\int_D dm\left(\vec{r}\right)\left(r^2 \overleftrightarrow{I}_P - \vec{r}\,\vec{r}\right)\right]\cdot\hat{n} = \hat{n}\cdot \overleftrightarrow{J} \cdot\hat{n}. \tag{28}$$

c) Choose a pooob, $\hat{x}_1$, $\hat{x}_2$, $\hat{x}_3$. Then $\overleftrightarrow{J} = J_{ij}\hat{x}_i\hat{x}_j$ and since $\overleftrightarrow{J}^T = \overleftrightarrow{J}$, $J_{ij} = J_{ji}$. Thus $\overleftrightarrow{J}$ is known if we know J_{11}, J_{22}, J_{33}, J_{12}, J_{23}, J_{31} relative to one pooob. But

$$J_{ij} = \int_D dm\left(\hat{r}\right)\left[r^2\delta_{ij} - r_i r_j\right].$$

12 a)

$$\begin{aligned}
\vec{L} &= \int_D dm\left(\vec{r}\right)\vec{r}\times\left[\vec{\Omega}\times\vec{r}\right] = \int_D dm\left(\vec{r}\right)\vec{r}\left[\vec{\Omega} - \vec{r}\left(\vec{r}\cdot\vec{\Omega}\right)\right] \\
&= \int_D dm\left(\vec{r}\right)\left[\left(r^2 \overleftrightarrow{I}_P - \vec{r}\,\vec{r}\right)\cdot\vec{\Omega}\right] \\
&= \int_D dm\left(\vec{r}\right)\left(r^2 \overleftrightarrow{I}_P - \vec{r}\vec{r}\right)\cdot\vec{\Omega} = \overleftrightarrow{J}\cdot\vec{\Omega}.
\end{aligned}$$

b)

$$\begin{aligned}
K &= \frac{1}{2}\int_D dm\,(\vec{r})\left(\vec{\Omega}\times\vec{r}\right)\cdot\left(\vec{\Omega}\times\vec{r}\right)\\
&= \frac{1}{2}\int_D dm\,(\vec{r})\,\Omega\cdot\left[\vec{r}\times\left(\vec{\Omega}\times\vec{r}\right)\right]\\
&= \frac{1}{2}\int_D dm\,(\vec{r})\,\vec{\Omega}\cdot\left[r^2\vec{\Omega}-\vec{r}\left(\vec{r}\cdot\vec{\Omega}\right)\right]\\
&= \frac{1}{2}\int_D dm\,(\vec{r})\left[\vec{\Omega}\cdot\left(r^2\overset{\leftrightarrow}{I}_P-\vec{r}\,\vec{r}\right)\cdot\vec{\Omega}\right]\\
&= \frac{1}{2}\vec{\Omega}\cdot\left[\int_D dm\,(\vec{r})\left(r^2\overset{\leftrightarrow}{I}_P-\vec{r}\vec{r}\right)\right]\cdot\vec{\Omega}\\
&= \frac{1}{2}\vec{\Omega}\cdot\overset{\leftrightarrow}{J}\cdot\vec{\Omega}.
\end{aligned}$$

c) Let the mass distribution consist of a single mass point m at position $\vec{r}$. Then

$$\begin{aligned}
\overset{\leftrightarrow}{J} &= m\left(r^2\overset{\leftrightarrow}{I}_P-\vec{r}\,\vec{r}\right)\\
\vec{L} &= \overset{\leftrightarrow}{J}\cdot\vec{\Omega}=m\Omega\left(r^2\overset{\leftrightarrow}{I}_P-\vec{r}\,\vec{r}\right)\cdot\hat{n}\\
&= m\Omega\left[r^2\hat{n}-\vec{r}\,(\vec{r}\cdot\hat{n})\right]\\
J\,(\hat{n})\,\vec{\Omega} &= \left(\hat{n}\cdot\overset{\leftrightarrow}{J}\cdot\hat{n}\right)\hat{n}\Omega=m\Omega\left[r^2-(\hat{n}\cdot\vec{r}\,)^2\right]\hat{n}.
\end{aligned}$$

If $\vec{L}=J(\hat{n})\vec{\Omega}$ then $[r^2-(\hat{n}\cdot\vec{r})^2]\hat{n}=r^2\hat{n}-\vec{r}(\vec{r}\cdot\hat{n})$, so $(\vec{r}\cdot\hat{n})^2\hat{n}=(\vec{r}\cdot\hat{n})\vec{r}$. To violate this condition choose $\hat{n}$ so it is neither parallel nor perpendicular to $\vec{r}$.

13 a) $\mathrm{tr}:\ P\otimes P\to\mathcal{R}$ is linear, so

$$\begin{aligned}
\mathrm{tr}\,\overset{\leftrightarrow}{M} &= tr\int_D dm\,(\vec{r})\,\vec{r}\,\vec{r}=\int_D dm\,(\vec{r})\,\mathrm{tr}\,\vec{r}\vec{r}\\
&= \int_D dm\,(\vec{r})\,r^2.
\end{aligned}$$

b) $\overset{\leftrightarrow}{J}=\int dm(\vec{r})\left[r^2\overset{\leftrightarrow}{I}_P-\vec{r}\,\vec{r}\right]=\left[\int_D dm\,(\vec{r})\,r^2\right]\overset{\leftrightarrow}{I}_P-\int_D dm\,(\vec{r})\,\vec{r}\,\vec{r}$

$$* \quad \therefore \overset{\leftrightarrow}{J}=\overset{\leftrightarrow}{I}_P\left(\mathrm{tr}\,\overset{\leftrightarrow}{M}\right)-\overset{\leftrightarrow}{M}\,.$$

c) From the above equation, since $\mathrm{tr}: P\otimes P\to\mathcal{R}$ is linear,

$$\mathrm{tr}\,\overset{\leftrightarrow}{J}=\left(\mathrm{tr}\,\overset{\leftrightarrow}{I}_P\right)\left(\mathrm{tr}\,\overset{\leftrightarrow}{M}\right)-\mathrm{tr}\,\overset{\leftrightarrow}{M}=3\,\mathrm{tr}\,\overset{\leftrightarrow}{M}-\mathrm{tr}\,\overset{\leftrightarrow}{M}=2\,\mathrm{tr}\,\overset{\leftrightarrow}{M}\,.$$

Hence, $\mathrm{tr}\,\overset{\leftrightarrow}{M}=1/2\,\mathrm{tr}\,\overset{\leftrightarrow}{J}$ and $\overset{\leftrightarrow}{M}=1/2\,\overset{\leftrightarrow}{I}_P\,(\mathrm{tr}\,\overset{\leftrightarrow}{J})-\overset{\leftrightarrow}{J}$.

d) Let $\hat{x}_1$, $\hat{x}_2$, $\hat{x}_3$ be a pooob parallel to the edges of the cube. Then

$$\begin{aligned} M_{ij} &= \int_{cube} dx_1\, dx_2\, dx_3 \;\; x_i x_j \rho \\ &= \rho \int_{-a}^{a} dx_1 \int_{-a}^{a} dx_2 \int_{-a}^{a} dx_3\, x_i\, x_j. \end{aligned}$$

Thus $M_{ij} = 0$ if $i \neq j$ and $M_{11} = M_{22} = M_{33} = 4a^2\rho \int_{-a}^{a} x^2 = 8a^5\rho/3$. Therefore $\overleftrightarrow{M} = 8a^5\rho/3 \overleftrightarrow{I}_P$ and tr $\overleftrightarrow{M} = 8a^5\rho$ so from $(*)$ above

$$\overleftrightarrow{J} = \frac{16a^5\rho}{3} \overleftrightarrow{I}_P \,.$$

Then for *any* $\hat{n}$, $\hat{n} \cdot \overleftrightarrow{J} \cdot \hat{n} = 16a^5\rho/3$.

14 a)

$$\begin{aligned} \frac{dr^1}{dt} &= r^1, \quad \frac{dr^2}{dt} = cr^2, \quad \frac{dr^3}{dt} = 0 \text{ so} \\ r^1 &= A^1 e^{(t-t_0)}, \quad r^2 = A^2 e^{c(t-t_0)}, \quad r^3 = A^3. \end{aligned}$$

The label $\vec{x} = x^i \hat{y}_i$ belongs to the particle which was at that position at $t = t_0$. For this particle when $t = t_0$, $r^i = x^i$, so $x^i = A^i$. Thus

$$(*) \quad r^1 = x^1 e^{(t-t_0)}, \quad r^2 = x^2 e^{c(t-t_0)}, \quad r^3 = x^3.$$

$$\begin{aligned} \vec{r} = r^i \hat{y}_i &= e^{(t-t_0)} x^1 \hat{y}_1 + e^{c(t-t_0)} x^2 \hat{y}_2 + x^3 \hat{y}_3 \\ &= \vec{x} \cdot \left[\hat{y}_1\hat{y}_1 e^{(t-t_0)} + \hat{y}_2\hat{y}_2 e^{c(t-t_0)} + \hat{y}_3\hat{y}_3 \right]. \\ \mathbf{r}^L(\vec{x}, t) &= \vec{x} \cdot \left[\hat{y}_1\hat{y}_1 e^{(t-t_0)} + \hat{y}_2\hat{y}_2 e^{c(t-t_0)} + \hat{y}_3\hat{y}_3 \right]. \end{aligned}$$

b) $\hat{y}_3 \cdot \vec{r}$ is independent of time for each particle, so each particle lies in a plane $\hat{y}_3 \cdot \vec{r} =$ constant. And for a particular particle, $r^1 r^2 = x^1 x^2 e^{(1+c)(t-t_0)} = x^1 x^2$ if $c = -1$, which is a hyperbola. And $r^2/r^1 = e^{(c-1)(t-t_0)} x_2/x_1 = x_2/x_1$, if $c = +1$. Then this is a straight line through the $\hat{y}_3$ axis.

c) From $(*)$ above, $x^1 = r^1 e^{-(t-t_0)}$, $x^2 = r^2 2e^{-c(t-t_0)}$, $x^3 = r^3$, so

$$\begin{aligned} \vec{x} &= x^i \hat{y}_i = r^1 e^{-(t-t_0)} \hat{y}_1 + r^2 e^{-c(t-t_0)} + r^3 \hat{y}_3 \\ &= \vec{r} \cdot \left[\hat{y}_1\hat{y}_1 e^{-(t-t_0)} + \hat{y}_2\hat{y}_2\, e^{-c(t-t_0)} + \hat{y}_3\hat{y}_3 \right]. \\ \mathbf{x}^E(\vec{r}, t) &= \vec{r} \cdot \left[\hat{y}_1\hat{y}_1\, e^{-(t-t_0)} + \hat{y}_2\hat{y}_2\, e^{-c(t-t_0)} + \hat{y}_3\hat{y}_3 \right]. \end{aligned}$$

d)

$$\begin{aligned}
\vec{a}^L(\vec{x},t) &= D_t^2\,\vec{\mathbf{r}}^{\,L}(\vec{x},t) = \vec{x}\cdot\left[\hat{y}_1\hat{y}_1\,e^{(t-t_0)} + \hat{y}_2\hat{y}_2\,e^{c(t-t_0)}\right] \\
\vec{a}^E(\vec{r},t) &= \vec{\mathbf{x}}^{\,E}(\vec{r},t)\cdot\left[\hat{y}_1\hat{y}_1\,e^{(t-t_0)} + c^2\hat{y}_2\hat{y}_2\,e^{c(t-t_0)}\right]
\end{aligned}$$

$$\begin{aligned}
\vec{a}^E(\vec{r},t) &= \vec{r}\cdot\left[\hat{y}_1\hat{y}_1\,e^{-(t-t_0)} + \hat{y}_2\hat{y}_2\,e^{-c(t-t_0)} + \hat{y}_3\hat{y}_3\right] \\
&\quad \cdot\left[\hat{y}_1\hat{y}_1\,e^{(t-t_0)} + c^2\hat{y}_2\hat{y}_2\,e^{c(t-t_0)}\right] \\
&= \vec{r}\cdot\left[\hat{y}_1\hat{y}_1 + c^2\hat{y}_2\hat{y}_2\right].
\end{aligned}$$

Another way to get the same result is to observe that

$$\vec{v}^E(\vec{r},t) = \vec{r}\cdot(\hat{y}_1\hat{y}_1 + c\,\hat{y}_2\hat{y}_2)$$

has no explicit t dependence, so $\partial_t\vec{v}^E(\vec{r},t) = \vec{0}$. Also $\vec{\partial}\vec{v}^E(\vec{r},t) = \hat{y}_1\hat{y}_1 + c\,\hat{y}_2\hat{y}_2$, so $\vec{v}^E\cdot\vec{\partial}\vec{v}^E(\vec{r},t) = \vec{r}\cdot(\hat{y}_1\hat{y}_1 + c^2\hat{y}_2\hat{y}_2)$. But

$$\begin{aligned}
\vec{a}^E(\vec{r},t) &= D_t\vec{v}^E(\vec{r},t) \\
&= \partial_t\vec{v}^E(\vec{r},t) + \vec{v}^E\cdot\vec{\partial}\vec{v}^E(\vec{r},t) \\
&= \vec{0} + \vec{r}\cdot\left(\hat{y}_1\hat{y}_1 + c^2\hat{y}_2\hat{y}_2\right).
\end{aligned}$$

Exercise 15

a) Given a continuum and a function $f : \mathcal{R} \to W$ which depends on t alone (W is a Euclidean vector space), invent a physical quantity which can reasonably be called f, and show that $(D_t f)^L(t) = (\partial_t f)^E(t) = \delta_t f(t)$.

b) A continuum undergoes rigid body motion. At time t the position, velocity and acceleration of the pivot particle are $\vec{R}(t)$, $\vec{V}(t) = \delta_t\vec{R}(t)$ and $\vec{A}(t) = \delta_t\vec{V}(t)$, and the angular velocity relative to the pivot particle is $\vec{\Omega}(t)$. Find the Eulerian description of particle acceleration in the material.

Exercise 16

Several decades ago, before the big bang was accepted, Fred Hoyle suggested that the expanding universe be explained by the continuous creation of matter. Then mass is not quite conserved. Suppose that a certain region of space is occupied by a continuum, and that at position $\vec{r}$ at time t new matter is being created at the rate of $\gamma^E(\vec{r},t)$ kilograms per cubic meter per second. Make the necessary changes in the derivation of the Eulerian form of the law of conservation of mass and in equations (13.1.12) and (13.1.17).

Exercise 17

At time t_0 in a certain material, the Cauchy stress tensor $\overleftrightarrow{S}$ is *symmetric*.

a) Let K' be an open set in the region occupied by the material, and suppose its boundary $\partial K'$ is piecewise smooth. Show that the total force $\vec{F}$ and the total torque $\vec{L}$ about $\vec{0}$ exerted on K' by the stress on $\partial K'$ are the same as those exerted by a fictitious body force with density $\vec{\partial}\cdot \overleftrightarrow{S}$ newtons / m^3 acting in K'.

b) Write the Taylor series expansion of $\overleftrightarrow{S}$ about $\vec{0}$ in physical space P as $\overleftrightarrow{S}(\vec{r}) = \overleftrightarrow{S}^{(0)} + \vec{r}\cdot S^{(1)} + 1/2(\vec{r}\vec{r})\langle 2\rangle S^{(2)} + \cdots$ where $\overleftrightarrow{S}^{(0)}$, $\overleftrightarrow{S}^{(1)}$, $\overleftrightarrow{S}^{(2)}$, $\cdots$ are constant tensors in $P\otimes P$, $\otimes^3 P$, $\otimes^4 P$, $\cdots$ and $(12)S^{(2)} = S^{(2)}$. Neglect all terms except those involving $\overleftrightarrow{S}^{(0)}$, $\overleftrightarrow{S}^{(1)}$, and $\overleftrightarrow{S}^{(2)}$, and calculate $\vec{F}$ and $\vec{L}$ of a) for $K' = B(\vec{0},c)$, the solid ball of radius c centered on $\vec{0}$.

Hint 1: Do all calculations relative to an orthonormal basis in P.

Hint 2: Use the symmetry of $B(\vec{0},c)$ to evaluate $\int_B dV(\vec{r}) r_i r_j$.

Solutions

15 a) The physical quantity has as its Lagrangian description $f^L(\vec{x},t) = f(t)$. Then its Eulerian description is $f^E(\vec{r},t) = f(t)$. Since $\vec{\partial}\vec{f} = 0$, $D_t f = \partial_t f + \vec{v}\cdot\vec{\partial}\vec{f} = \partial_t f$.

b)

$$\begin{aligned}\vec{v}^E(\vec{r},t) &= \vec{V}(t)+\vec{\Omega}(t)\times\left[\vec{r}-\vec{R}(t)\right]\\ &= \vec{V}(t)+\left[\vec{r}-\vec{R}(t)\right]\cdot\overleftrightarrow{\Omega}(t).\end{aligned}$$

$\vec{\partial}\vec{v}^E = \overleftrightarrow{\Omega}$ so

$$\begin{aligned}\vec{v}^E\cdot\vec{\partial}\vec{v}^E &= \vec{v}^E\cdot\overleftrightarrow{\Omega} = \vec{\Omega}\times\vec{v}^E = \vec{\Omega}\times\vec{V}+\vec{\Omega}\times\left[\vec{\Omega}\times\left(\vec{r}-\vec{R}\right)\right]\\ \partial_t\vec{v}^E &= \vec{A}(t)-\vec{V}(t)\cdot\overleftrightarrow{\Omega}+\left[\vec{r}-\vec{R}\right]\cdot\partial_t\overleftrightarrow{\Omega}\\ &= \vec{A}-\vec{\Omega}\times\vec{V}+\partial_t\vec{\Omega}\times\left(\vec{r}-\vec{R}\right).\\ \left(D_t\vec{v}\right)^E &= \vec{a}^E=\left(\partial_t\vec{v}+\vec{v}\cdot\vec{\partial}\vec{v}\right)^E=\vec{A}+\partial_t\vec{\Omega}\times\left(\vec{r}-\vec{R}\right)+\vec{\Omega}\times\left[\vec{\Omega}\times\left(\vec{r}-\vec{R}\right)\right]\\ \vec{a}^E(\vec{r},t) &= \vec{A}(t)+\left[\vec{r}-\vec{R}(t)\right]\cdot\left[\partial_t\overleftrightarrow{\Omega}+\overleftrightarrow{\Omega}\cdot\overleftrightarrow{\Omega}\right].\end{aligned}$$

16. The amount of mass in $K'(t)$ is

$$M\left[K'(t)\right]=\int_{K'(t)}dV_P(\vec{r})\rho^E(\vec{r},t).$$

It is a mathematical identity (derivative over moving volume) that

$$\begin{aligned}\frac{d}{dt}M\left[K'(t)\right] &= \int_{K'(t)}dV_P\partial_t\rho^E+\int_{\partial K'(t)}dA_p\hat{n}_p\cdot\left(\vec{v}^E\rho^E\right)\\ &= \int_{K'(t)}dV_p\left[\partial_t\rho+\vec{\partial}\cdot(\rho\vec{v})\right]^E(\vec{r},t).\end{aligned}$$

Hoyle's new physical law is that

$$\frac{d}{dt}M\left[K'(t)\right]=\int_{K'(t)}dV_p\,\gamma^E(\vec{r},t).$$

Therefore

$$\int_{K'(t)}dV_p\left[\partial_t\rho+\vec{\partial}\cdot(\rho\vec{v})-\gamma\right]^E(\vec{r},t)=0.$$

Since this is true for every open set $K'(t)$ with piecewise smooth boundary, the vanishing integral theorem gives

$$\partial_t\rho+\vec{\partial}\cdot(\rho\vec{v})=\gamma.$$

Then for any other physical quantity $\vec{f}$,

$$
\begin{aligned}
\frac{d}{dt}\int_{K'(t)} dV_P \left(\rho\vec{f}\right)^E &= \int_{K'(t)} dV_P \left[(\partial_t\rho)\,\vec{f}+\rho\left(\partial_t\vec{f}\right)\right]^E \\
&\quad + \int_{K'(t)} dA_P\hat{n}\cdot\left[\vec{v}\left(\rho\vec{f}\right)\right]^E \\
&= \int_{K'(t)} dV_P \left[(\partial_t\rho)\,\vec{f}+\rho\left(\partial_t\vec{f}\right)+\vec{\partial}\cdot\left(\rho\vec{v}\vec{f}\right)\right]^E \\
&= \int_{K'(t)} dV_P \left\{(\partial_t\rho)\,\vec{f}+\rho\left(\partial_t\vec{f}\right)+\left[\vec{\partial}\cdot(\rho\vec{v})\right]\vec{f}+\rho\vec{v}\cdot\vec{\partial}f\right\}^E \\
&= \int_{K'(t)} dV_P \left\{\rho\left(\partial_t\vec{f}+\vec{v}\cdot\partial_t\vec{f}\right)+\left[\partial_t\rho+\vec{\partial}\cdot(\rho\vec{v})\right]\vec{f}\right\}^E \\
&= \int_{K'(t)} dV_P \left[\rho D_t\vec{f}+\gamma\vec{f}\right]^E
\end{aligned}
$$

17 a)

$$
\int_{\partial K'} dA\hat{n}\cdot\overset{\leftrightarrow}{S}=\int_{K'} dV\left(\vec{\partial}\cdot\overset{\leftrightarrow}{S}\right)
$$

$$
\left[\int_{\partial K'} dA\,\vec{r}\times\left(\hat{n}\cdot\overset{\leftrightarrow}{S}\right)\right]_i = -\left[\int_{K'} dA\,\vec{n}\times\overset{\leftrightarrow}{S}\times\vec{r}\right]_i
$$

$$
\begin{aligned}
&= -\int_{K'} dV\left[\vec{\partial}\cdot\left(\overset{\leftrightarrow}{S}\times\vec{r}\right)\right]_i = -\int_{K'} dV\,\partial_j\left(S_{jk}r_l\varepsilon_{kli}\right) \\
&= -\varepsilon_{ikl}\int_{K'} dV\left(r_l\partial_j S_{jk}+\delta_{jl}S_{jk}\right) \\
&= -\varepsilon_{ikj}\int_{K'} dV\left(r_j\partial_l S_{lk}+\delta_{jl}S_{lk}\right) \\
&= \varepsilon_{ijk}\int_{K'} dV\left(r_j\partial_l S_{lk}+S_{jk}\right) \\
&= \int_{K'} dV\left(\vec{r}\times\vec{\partial}\cdot\overset{\leftrightarrow}{S}\right)_i .
\end{aligned}
$$

b)

$$
\begin{aligned}
S_{ij} &= S^{(0)}_{ij}+r^k S^{(1)}_{kij}+\frac{1}{2}r^k r^l S^{(2)}_{klij} \\
\partial_i S_{ij} &= S^{(1)}_{iij}+\frac{1}{2}\left(\delta^k{}_i r^l+\delta^l{}_i r^k\right)S^{(2)}_{klij} \\
&= S^{(1)}_{iij}+\frac{1}{2}\left(r^l S^{(2)}_{ilij}+r^k S^{(2)}_{kiij}\right) \\
\partial_i S_{ij} &= S^{(1)}_{iij}+r^k S^{(2)}_{kiij}=v_j+r^k w_{kj}
\end{aligned}
$$

(defines v_j and w_{kj}).

$$
\begin{aligned}
\vec{\partial}\cdot \overleftrightarrow{S} &= \vec{v} + \vec{r}\cdot \overleftrightarrow{w}\,. \\
\vec{r}\times\vec{\partial}\cdot \overleftrightarrow{S} &= \vec{r}\times\vec{v} + \vec{r}\times\left(\vec{r}\cdot \overleftrightarrow{w}\right) \\
\left(\vec{r}\times\vec{\partial}\cdot \overleftrightarrow{S}\right)_i &= (\vec{r}\times\vec{v})_i + \varepsilon_{ijk} r_j r_l w_{lk} \\
\int_B dV \vec{\partial}\cdot \overleftrightarrow{S} &= |B|\vec{v} + \left(\int_B dV\vec{r}\right)\cdot\vec{w}. \quad \int_B dV\vec{r} = \overleftrightarrow{0} \text{ so} \\
\vec{F} &= \frac{4\pi}{3}c^3\vec{v} = \frac{4\pi}{3}c^3\,\mathrm{tr}_{12}\, S^{(1)} \\
\vec{L} &= \int_B dV\vec{r}\times\vec{\partial}\cdot \overleftrightarrow{S}, \quad L_i = \int_B dV\varepsilon_{ijk}\left(r_j v_k + r_j r_l w_{lk}\right).
\end{aligned}
$$

$\int_B dV r_j = 0$ and $\int_B dV r_j r_l = 4\pi/15\, c^5\delta_{jl}$ so $L_i = 4\pi/15\, c^5\varepsilon_{ijk}w_{jk}$, $\vec{L} = 4\pi/15\, c^5$ $A\langle 2\rangle \mathrm{tr}_{23} S^{(2)}$.

Exercise 18

a) The angular momentum of a single dysprosium atom is $15h/4\pi$ where h is Planck's constant. The density of dysprosium is $8.54 gm/cm^3$, and its atomic weight is 162.50. Compute the angular momentum of a stationary ball of solid dysprosium with radius r cm if all the dysprosium atoms are aligned in the same direction. If the dysprosium atoms had no angular momentum, how rapidly would the ball have to rotate as a rigid body in order to have the same angular momentum?

b) Let $\hat{y}_1$, $\hat{y}_2$, $\hat{y}_3$ be a positively oriented ordered orthonormal basis for physical space P. A permanent alnico magnet has magnetization density $\vec{M} = M\hat{y}_1$ where $M = 10^6$ amp/meter (close to the upper limit for permanent magnets). The magnet is held at rest in a uniform magnetic field $\vec{B} = B\hat{y}_2$ with $B = 1$ tesla (about the upper limit for commercial electromagnets). Then the volume torque on the magnet is $\vec{m} = \vec{M}\times\vec{B}$ joules /meter^3. Suppose that the magnet's intrinsic angular momentum density $\vec{l}$ does not change with time, and that its torque stress tensor $\overleftrightarrow{M}$ vanishes. Let $\overleftrightarrow{S}$ be the Cauchy stress tensor in the magnet. Find $1/2(\overleftrightarrow{S} - \overleftrightarrow{S}^T)$, the antisymmetric part of $\overleftrightarrow{S}$, in atmospheres of pressure (one atmosphere $= 1.013\times 10^5$ newtons/meter^2 $=$ 1.013×10^6 dynes /cm^2). If the symmetric part of $\overleftrightarrow{S}$, $1/2(\overleftrightarrow{S} + \overleftrightarrow{S}^T)$, vanishes, sketch

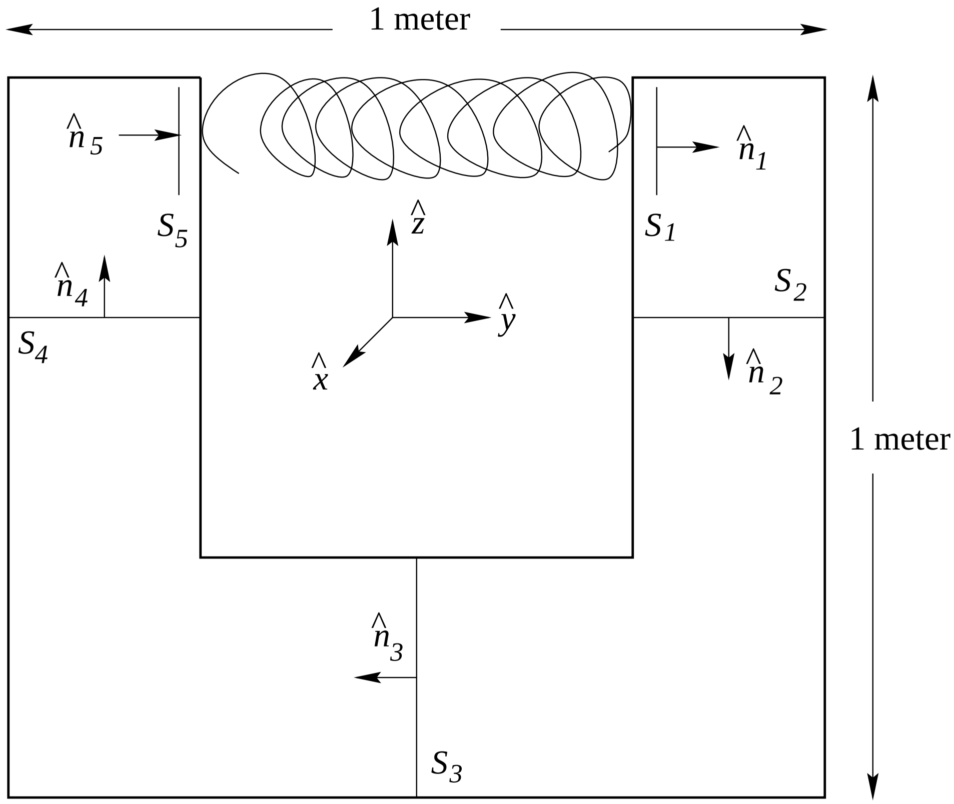


Figure Ex-4:

the stresses acting on the surface of a small spherical ball of material in the magnet.

Exercise 19

a) Give the obvious definition of the stuff "y component of momentum". Give its $(\Phi, \vec{\mathbf{F}})$, its $(\phi, \vec{F})$, and its creation rate γ.

b Show that in a material at rest with no body forces, the material flux density $\vec{F}$ of y momentum satisfies $\vec{\partial} \cdot \vec{F} = 0$. (If $\vec{F}$ were the velocity of a fluid, the fluid would be incompressible.)

c) An aluminum casting 10 cm thick and 1 meter on each side is shaped as in figure Ex-4. A steel spring is compressed and placed between the jaws of the casting as shown. The spring and casting are at rest and gravity is to be neglected. Roughly *sketch* the field lines of $\vec{F}$ (the material flux density of $\hat{y}$ momentum) in the casting. (If $\vec{F}$ were a force field, the field lines would be the lines of force; if $\vec{F}$ were a velocity field, they would be the streamlines.)

Hint for c): *Estimate qualitatively* the sign of the $\hat{y}$ component of the force exerted by the aluminum just in front of $(S_i, \hat{n}_i)$ on the aluminum just behind it for the surfaces $i = 1, 2, 3, 4, 5$ sketched in the figure.

Note for c): The $\hat{x}$ axis points out of the paper, the $\hat{y}$ axis is parallel to the axis of the spring, and the $\hat{z}$ axis points up, as shown in the figure.

Solutions

18 a) There are $6.025 \times 10^{23}/162.50 = 3.7 \times 10^{21}$ atoms in a gram of Dysprosium. Each has angular momentum $(15/4\pi)(6.625 \times 10^{-27})$ erg sec. If all are aligned, one gram of Dy has angular momentum $l = 2.932 \times 10^{-5}$ erg sec/gm. A Dy ball of radius r cm has intrinsic angular momentum $(4\pi/3)\rho r^3 l$. If this angular momentum were not intrinsic, but due to an angular velocity ω of rigid rotation, the angular momentum would be $(2/5)r^2(4\pi/3)\rho r^3\omega$. Thus $l = (2/5)\, r^2\omega$, or $\omega = (5/2)\, l/r^2 = 7.33 \times 10^{-5}/r^2$ radians/sec, with r in cm.

b)

$$
\begin{aligned}
\vec{0} &= \vec{m} + A\langle 2\rangle \overleftrightarrow{S}, \text{ so relative to the pooob } \hat{y}_1, \hat{y}_2, \hat{y}_3 \\
0 &= m_i + \varepsilon_{ijk} S_{jk}. \text{ Then} \\
0 &= m_i \varepsilon_{ilm} + \varepsilon_{ilm}\varepsilon_{ijk} S_{jk} \\
0 &= m_i \varepsilon_{ilm} + \left(\delta_{lj}\delta_{mk} - \delta_{lk}\delta_{mj}\right) S_{jk} = m_i\varepsilon_{ilm} + S_{lm} - S_{ml}.
\end{aligned}
$$

$$
(*) \quad \frac{1}{2}\ (S_{lm} - S_{ml}) = -\frac{1}{2}\,\varepsilon_{lmi} m_i.
$$

In our problem, $\vec{m} = MB\hat{y}_3$ so

$$
\frac{1}{2}\,(S_{12} - S_{21}) \ = \ -\frac{1}{2} MB, \quad \frac{1}{2}\,(S_{13} - S_{31}) = \frac{1}{2}\,(S_{23} - S_{32}) = 0, \text{ and}
$$

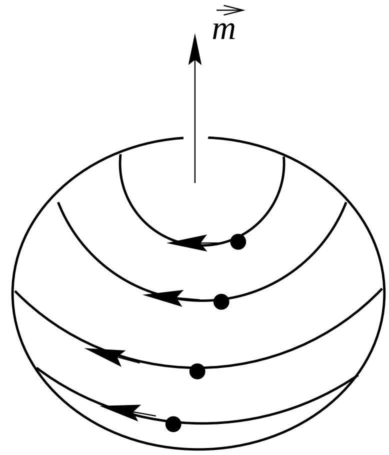


Figure Ex-5:

$$\begin{aligned}\frac{1}{2}\left(\overleftrightarrow{S}-\overleftrightarrow{S}^{T}\right) &= \frac{1}{2}\left(\overleftrightarrow{S}-\overleftrightarrow{S}^{T}\right)_{ij}\hat{y}_i\hat{y}_j, \text{ so}\\ \frac{1}{2}\left(\overleftrightarrow{S}-\overleftrightarrow{S}^{T}\right) &= \frac{1}{2}MB\left(\hat{y}_2\hat{y}_1-\hat{y}_1\hat{y}_2\right).\end{aligned}$$

$$\begin{aligned}\frac{1}{2}MB &= \frac{10^6}{2}\text{ joules/meter}^3 = \frac{10^6}{2}\text{ newtons/meter}^2\\ &= \frac{10^6}{2\times 1.013\times 10^5}\text{ atm} = 4.94\text{ atmospheres.}\end{aligned}$$

For any $\overleftrightarrow{S}$, $\overleftrightarrow{S}= 1/2(\overleftrightarrow{S} + \overleftrightarrow{S}^{T}) + 1/2(\overleftrightarrow{S} - \overleftrightarrow{S}^{T})$. If $1/2(\overleftrightarrow{S} + \overleftrightarrow{S}^{T}) = 0$, then $\overleftrightarrow{S}= 1/2(\overleftrightarrow{S} - \overleftrightarrow{S}^{T})$. From (*), $\overleftrightarrow{S}= -1/2\, A\cdot\vec{m}$. The stress $\vec{S}(\hat{n})$ is $\hat{n}\cdot\overleftrightarrow{S}= -1/2\hat{n}\cdot A\cdot\vec{m}$ so $S_j(\hat{n}) = -1/2 n_i\varepsilon_{ijk}m_k = 1/2\varepsilon_{jik}n_i m_k = 1/2(\hat{n}\times\vec{m})_j$. Thus

$$\vec{S}(\hat{n}) = \frac{1}{2}\hat{n}\times\vec{m}.$$

The stress is directed along lines of latitude about the north pole $\vec{m}$, points west, and has magnitude equal to the sine of the colatitude of $\hat{n}$ relative to $\vec{m}$.

19 a)

$$\begin{aligned}\Phi = \rho\vec{v}\cdot\vec{y} \qquad\qquad & \phi = \vec{v}\cdot\hat{y}\\ \vec{\mathbf{F}} = \left(\rho\vec{v}\vec{v}-\overleftrightarrow{S}\right)\cdot\hat{y} \quad & \vec{F} = -\overleftrightarrow{S}\cdot\hat{y}\end{aligned}$$

$$\gamma = \rho D_t\phi + \vec{\partial}\cdot\left(-\overleftrightarrow{S}\cdot\hat{y}\right) = \left(\rho D_t\vec{v} - \vec{\partial}\cdot\overleftrightarrow{S}\right)\cdot\hat{y} = \vec{f}\cdot\hat{y}.$$

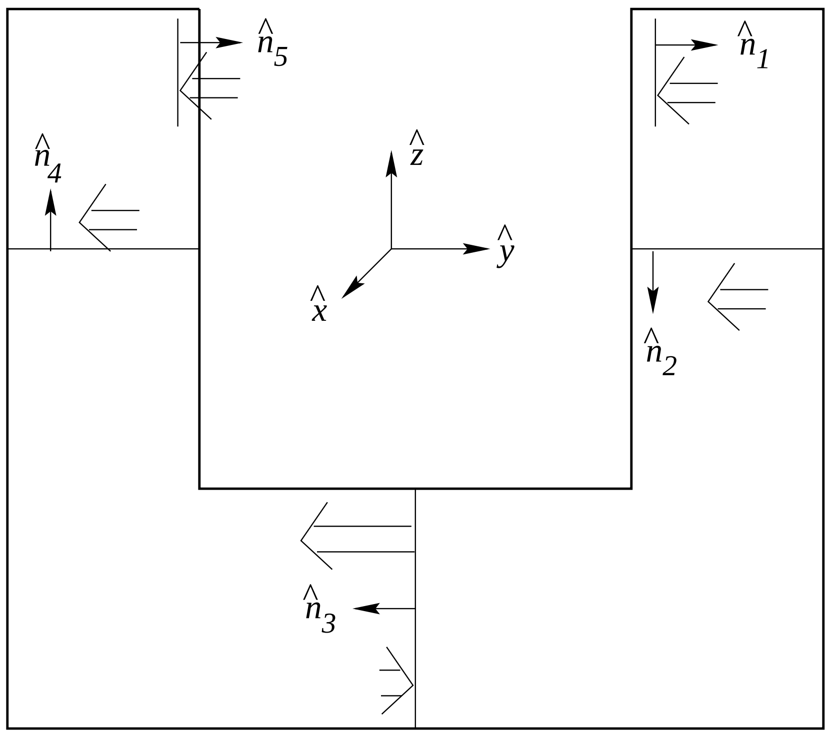


Figure Ex-6:

b) If $\vec{f}=0$ and $\vec{v}=\vec{0}$ then $\vec{\partial}\cdot\overleftrightarrow{S}=\vec{0}$ so $(\vec{\partial}\cdot\overleftrightarrow{S})\cdot\hat{y}=\vec{\partial}\cdot(\overleftrightarrow{S}\cdot\hat{y})=0$.

c) $\vec{F}=-\overleftrightarrow{S}\cdot\hat{y}$. The double arrows show $\hat{n}_i\cdot\overleftrightarrow{S}$ on the five surface sections. Then

$$\text{on } S_1,\quad \hat{n}_1\cdot\vec{F}=\hat{y}\cdot\left(-\overleftrightarrow{S}\cdot\hat{y}\right)=-\hat{y}\cdot\overleftrightarrow{S}\cdot\hat{y}=-\left(\hat{n}_1\cdot\overleftrightarrow{S}\right)\cdot\hat{y}>0$$

$$\text{on } S_2,\quad \hat{n}_2\cdot\vec{F}=-\hat{z}\cdot\left(-\overleftrightarrow{S}\cdot\hat{y}\right)=-\left(-\hat{z}\cdot\overleftrightarrow{S}\right)\cdot\hat{y}=-\left(\hat{n}_2\cdot\overleftrightarrow{S}\right)\cdot\hat{y}>0$$

$$\text{on } S_3,\quad \hat{n}_3\cdot\vec{F}=-\hat{y}\cdot\left(-\overleftrightarrow{S}\cdot\hat{y}\right)=-\left(-\hat{n}_3\cdot\overleftrightarrow{S}\right)\cdot\hat{y}\ \begin{matrix}>0\text{ above}\\<0\text{ below}\end{matrix}$$

$$\text{on } S_4,\quad \hat{n}_4\cdot\vec{F}=\hat{z}\cdot\left(-\overleftrightarrow{S}\cdot\hat{y}\right)=-\left(\hat{n}_4\cdot\overleftrightarrow{S}\right)\cdot\hat{y}>0$$

$$\text{on } S_5,\quad \hat{n}_5\cdot\vec{F}=\hat{y}\cdot\left(-\overleftrightarrow{S}\cdot\hat{y}\right)=-\left(\hat{n}_5\cdot\overleftrightarrow{S}\right)\cdot\hat{y}>0.$$

Thus the flow lines are as in figure Ex-7.

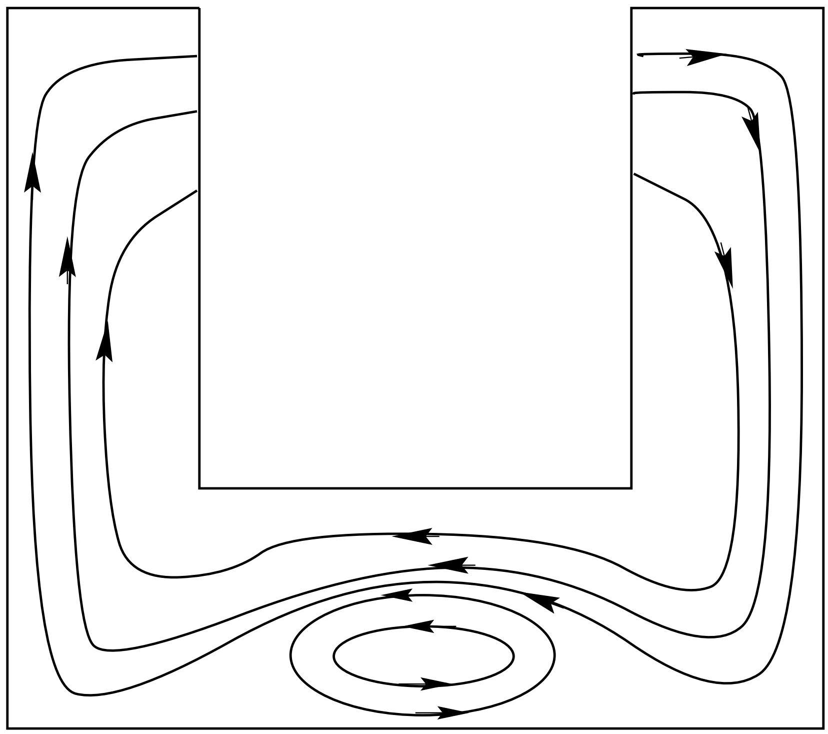


Figure Ex-7:

Exercise 20

Geological evidence indicates that over the last 5 million years, the west side of the San Andreas fault has moved north relative to the east side at an average rate of about 6 cm/year. On a traverse across the fault, Brune, Henyey and Roy (1969) (J.G.R. 74, 3821) found no detectable extra geothermal heat flow due to the fault. They estimate that they would have detected any anomalous heat flow larger than 1.3×10^{-2} watts/meter^2. Use (13.7.45) to obtain an upper bound on the northward component of the average stress exerted by the material just west of the fault on the material just east of the fault. Assume that the fault extends vertically down from the surface to a depth D, and that the shape of the heat flow anomaly due to the fault is a tent function whose width is D on each side of the fault.

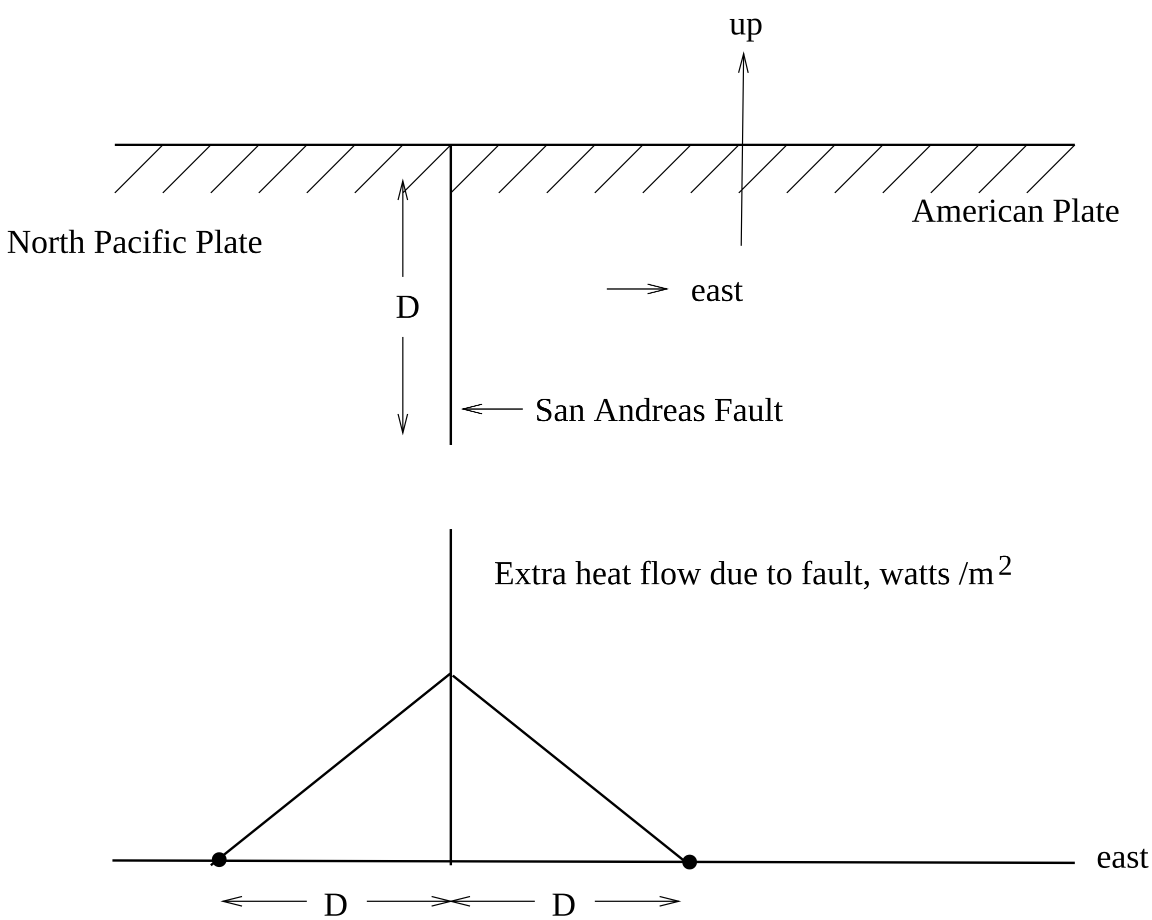

up
North Pacific Plate
American Plate
D
east
San Andreas Fault
Extra heat flow due to fault, watts /m2
east
D
D

Figure Ex-8:

Solutions

20. We assume that the heat flow is steady. Since we are interested only in averages, we may assume that $[\overleftrightarrow{S}]_-^+$ is constant on the fault, as is $[\hat{\nu}\cdot\vec{H}]_-^+$. D is probably < 20 km (see Brune et al. 1969) so we may treat the fault as infinitely long. Then $\vec{H}$ lies in east-west vertical planes. Consider a length L along the fault. The amount of heat produced by this section of the fault in $LD[\hat{\nu}\cdot\vec{H}]_-^+$. It all flows out of the surface of the earth in a rectangle of length L along the fault and width $2D$ across the fault, the profile being the triangular one sketched as the bottom figure in the exercise. The area under that triangle is Dh where h is the triangle's height, so the total heat flow out in the length L and width $2D$ is LDh. This must equal $LD[\hat{\nu}\cdot\hat{H}]_-^+$ in the steady state. Thus $h = [\hat{\nu}\cdot\hat{H}]_-^+$. Since no anomaly was detected, $h \le 1.3\times 10^{-2} w/m^2$. Thus $[\hat{\nu}\cdot\hat{H}]_-^+ = \hat{\nu}\cdot[\overleftrightarrow{S}]_-^+\cdot[\vec{v}]_-^+ \le 1.3\times 10^{-2} w/m^2$. Let $\hat{x}$ be a unit vector pointing north, and let $\hat{y} = \hat{\nu}$ point west. Then $[\vec{v}]_-^+ = 6\hat{x}\, cm/yr = 6\times 10^{-2}\hat{x}$ meters/yr $= 2\times 10^{-9}\hat{x}$ meters/sec. The heat production must be positive, so

$$0 \le \hat{\nu}\cdot\left[\overleftrightarrow{S}\right]_-^+\cdot\hat{x}\left(2\times 10^{-9}\right) \le 1.3\times 10^{-2},$$

or

$$0 \le \left[\hat{\nu}\cdot\overleftrightarrow{S}\right]_-^+\cdot\hat{x} \le 65\times 10^5 \text{ pascals.}$$

The material just west of the fault exerts on the material just east of the fault, a stress $[\hat{\nu}\cdot\overleftrightarrow{S}]_-^+$. The northward component of this stress is $[\hat{\nu}\cdot\overleftrightarrow{S}]_-^+\cdot\hat{x}$, and it lies between 0 and 65 bars (10^5 pascal).

See Brune et al., for a discussion of energy loss from $\vec{v}\cdot[\overleftrightarrow{S}]_-^+\cdot[\vec{v}]_-^+$ by seismic radiation. This exercise works only for the parts of the fault where most energy is not radiated, i.e., the parts of the fault which creep.

Exercise 21

Below are the Eulerian descriptions of three motions of continua. In these descriptions, $\{\hat{x},\hat{y},\hat{z}\}$ is a right-handed orthonormal basis for physical space P, α is a constnat real number, and $\vec{\Omega}$ is a constant vector. For each motion, find the local strain rate tensor $\dot{\overleftrightarrow{\sigma}}$, its eigenvectors, and its eigenvalues. Also find the local rotation rate tensor $\dot{\overleftrightarrow{\omega}}$, and the local angular velocity $\dot{\vec{w}}$.

a) $\vec{v}^E(\vec{r},t) = \alpha\vec{r} \cdot (\hat{x}\hat{x} - \hat{y}\hat{y}) = \alpha(x\hat{x} - y\hat{y})$

b) $\vec{v}^E(\vec{r},t) = \alpha\vec{r} \cdot (\hat{y}\hat{x}) = \alpha y\hat{x}$

c) $\vec{v}^E(\vec{r},t) = \vec{\Omega} \times \vec{r}$.

Exercise 22

Let $\vec{v}$, $\dot{\overleftrightarrow{\sigma}}$, $\dot{\overleftrightarrow{\omega}}$ be the Eulerian descriptions of the velocity, local strain rate tensor and local rotation rate tensor in a certain continuum. Let $\hat{x}_1$, $\hat{x}_2$ $\hat{x}_3$ be a right-handed orthonormal basis in physical space P. Let position vectors be written $\vec{r} = r_i\hat{x}_i$. Let ∂_i denote the partial derivative with respect to r_i. Take tensor components relative to $\hat{x}_1, \hat{x}_2, \hat{x}_3$.

a) Show that $\partial_i \dot{\omega}_{jk} = \partial_j \dot{\sigma}_{ki} - \partial_k \dot{\sigma}_{ij}$

b) Show that $\partial_i\partial_j \dot{\sigma}_{kl} - \partial_j\partial_k \dot{\sigma}_{li} + \partial_k\partial_l \dot{\sigma}_{ij} - \partial_l\partial_i \dot{\sigma}_{jk} = 0$.

c) Show that if $\dot{\overleftrightarrow{\sigma}}: P \to P \otimes P$ is any tensor-valued function of $\vec{r}$ such that everywhere $\dot{\sigma}_{ij} = \dot{\sigma}_{ji}$ and b) holds, then a function $\dot{\overleftrightarrow{\omega}}: P \to P \otimes P$ can be found which satisfies $\dot{\omega}_{ij} = -\dot{\omega}_{ji}$ and also a) above. Show that this function is unique up to an additive constant tensor $\dot{\overleftrightarrow{\Omega}}$ which is antisymmetric.

d) Show that if $\dot{\overleftrightarrow{\sigma}}$ is the function in c), there is a velocity function $\vec{v} : P \to P$ whose local strain rate tensor is $\dot{\overleftrightarrow{\sigma}}$. Show that any two such functions $\vec{v}_1$ and $\vec{v}_2$ differ by the velocity field of a rigid motion; that is, there are fixed vectors $\vec{\Omega}$ and $\vec{V}$ such that for all $\vec{r} \in P$, $\vec{v}_2(\vec{r}) - \vec{v}_1(\vec{r}) = \vec{V} + \vec{\Omega} \times \vec{r}$.

Hint for c) and d): Suppose $g_i : P \to \mathcal{R}$ for $i = 1, 2, 3$. You know from elementary vector calculus that there is a "potential function" $f : P \to \mathcal{R}$ such that $g_i = \partial_i f$ iff $\partial_i g_j = \partial_j g_i$ everywhere; and if this is so, then $f(\vec{r}) = f(\vec{0}) + \int_0^{\vec{r}} dr_i g_i$ (line integral).

Exercise 23

Use the notation beginning on page 256 for discussing small disturbances of an HTES of an elastic material whose reference state is HTES_0.

a) Using only θ_0, $\overleftrightarrow{S}_0$, B_0, $\overleftrightarrow{W}_0$, ρ_0, $\overleftrightarrow{\overleftrightarrow{Q}}_0$, explicitly express $U - U_0$ as a function of δN and $\delta \overleftrightarrow{G}$, correct to second order (i.e., including the terms $(\delta N)^2$, $\delta N \delta \overleftrightarrow{G}$ and $\delta \overleftrightarrow{G} \delta \overleftrightarrow{G}$.)

b) The tensor $\overleftrightarrow{\overleftrightarrow{F}}_0$, is called the isentropic or adiabatic stiffness tensor because if $\delta N = 0$ (no change in entropy) then

$$\delta \overleftrightarrow{S} = \delta \overleftrightarrow{G} \langle 2 \rangle \overleftrightarrow{\overleftrightarrow{F}} .$$

Find the isothermal stiffness tensor $\overleftrightarrow{\overleftrightarrow{F}}'_0$, the tensor such that if $\delta\theta = 0$ then

$$\delta \overleftrightarrow{S} = \delta \overleftrightarrow{G} \langle 2 \rangle \overleftrightarrow{\overleftrightarrow{F}}'_0 .$$

Express $\overleftrightarrow{\overleftrightarrow{F}}'_0$ in terms of $\overleftrightarrow{\overleftrightarrow{F}}_0$, $\overleftrightarrow{W}_0$ B_0, ρ_0.

Exercise 24

Using the notation on page 260, give the results of 23 a), b) when HTES_0 is isotropic.

Solutions

21 a)

$$\vec{\partial}\vec{v}^E = \alpha\left(\hat{x}\hat{x} - \hat{y}\hat{y}\right) = \left(\vec{\partial}\vec{v}^E\right)^T ,$$

so

$$\dot{\overleftrightarrow{\sigma}} = \alpha\left(\hat{x}\hat{x} - \hat{y}\hat{y}\right) ,$$

eigenvectors and eigenvalues being $(\hat{x}, \alpha)$, $(\hat{y}, -\alpha)$, $(\hat{z}, -0)$. $\dot{\overleftrightarrow{\omega}} = \overleftrightarrow{0}$, so $\dot{\vec{\omega}} = \vec{0}$,

b)

$$\begin{aligned}\vec{\partial}\vec{v}^E &= \alpha\hat{y}\hat{x}, \quad \left(\vec{\partial}\vec{v}^E\right)^T = \alpha\hat{x}\hat{y},\\ \dot{\overleftrightarrow{\sigma}} &= \frac{\alpha}{2}\left(\hat{x}\hat{y}+\hat{y}\hat{x}\right), \quad \dot{\overleftrightarrow{\omega}} = \frac{\alpha}{2}\left(\hat{y}\hat{x}-\hat{x}\hat{y}\right)\end{aligned}$$

eigenvectors and eigenvalues of $\dot{\overleftrightarrow{\sigma}}$ are

$$\left(\frac{\hat{x}+\hat{y}}{\sqrt{2}}, \frac{\alpha}{2}\right), \quad \left(\frac{\hat{x}+\hat{y}}{\sqrt{2}}, -\frac{\alpha}{2}\right), \quad (\hat{z}, 0).$$

$$\dot{\vec{w}} = \frac{1}{2}A\langle 2\rangle\ \dot{\overleftrightarrow{\omega}} = \frac{1}{2}\frac{\alpha}{2}\left(\hat{y}\times\hat{x}-\hat{x}\times\hat{y}\right) = -\frac{1}{2}\alpha\hat{z}$$

c) Let $\overleftrightarrow{\Omega} = A\cdot\vec{\Omega}$. Then

$$\begin{aligned}\vec{v}^E = A\langle 2\rangle\vec{\Omega}\vec{r} &= -\vec{\Omega}\cdot A\cdot\vec{r}\\ &= -\overleftrightarrow{\Omega}\cdot\vec{r} = \vec{r}\cdot\overleftrightarrow{\Omega}.\end{aligned}$$

Then $\vec{\partial}\vec{v}^E = \overleftrightarrow{\Omega}$, $\left(\vec{\partial}\vec{v}^E\right)^T = -\overleftrightarrow{\Omega}$, so $\dot{\overleftrightarrow{\sigma}} = \overleftrightarrow{0}$ (so all eigenvalues are 0 and any vector is an eigenvector)

$$\dot{\overleftrightarrow{\omega}} = \overleftrightarrow{\Omega}, \quad \dot{\vec{\omega}} = \vec{\Omega}.$$

22. a)

$$\begin{aligned}\dot{\omega}_{jk} &= \frac{1}{2}\left(\partial_j v_k - \partial_k v_j\right)\\ \dot{\sigma}_{ij} &= \frac{1}{2}\left(\partial_i v_j + \partial_j v_i\right)\\ \partial_k\dot{\sigma}_{ij} &= \frac{1}{2}\left(\partial_k\partial_i v_j + \partial_k\partial_j v_i\right)\end{aligned}$$

Replace (ijk) by (kij) and this becomes

$$\begin{aligned}\partial_j\dot{\sigma}_{ki} &= \frac{1}{2}\left(\partial_j\partial_k v_i + \partial_j\partial_i v_k\right).\\ \partial_i\dot{\omega}_{jk} &= \frac{1}{2}\left(\partial_i\partial_j v_k - \partial_i\partial_k v_j\right).\end{aligned}$$

Comparing these three equations gives a).

b) $\partial_i\partial_l\dot\omega_{jk} = \partial_l\partial_i\dot\omega_{jk}$ so, by a),

$$\partial_i\left(\partial_j\dot\sigma_{kl} - \partial_k\dot\sigma_{lj}\right) = \partial_l\left(\partial_j\dot\sigma_{ki} - \partial_k\dot\sigma_{ij}\right)$$

or, since $\dot\sigma_{kl} = \dot\sigma_{lk}$ and $\partial_i\partial_j = \partial_j\partial_i$,

$$\partial_j\partial_i\dot\sigma_{kl} - \partial_i\partial_k\dot\sigma_{lj} + \partial_k\partial_i\dot\sigma_{li} - \partial_l\partial_j\dot\sigma_{ik} = 0.$$

Now make the replacements $j \to i$, $i \to j$, and one proves b).

c) Fix i and j. a) can be written

$$(*) \qquad \partial_k\dot\omega_{ij} = \partial_i\dot\sigma_{jk} - \partial_j\dot\sigma_{ki}.$$

Then b) implies $\partial_k\partial_l\dot\omega_{ij} = \partial_l\partial_k\dot\omega_{ij}$, which are exactly the integrability conditions for (*). They insure that the line integral

$$\int_{\vec 0}^{\vec r} dr^k\left(\partial_i\dot\sigma_{jk} - \partial_j\dot\sigma_{ki}\right)$$

will be independent of path. Thus the solution of (*) exists, and has the form

$$\dot\omega_{ij}(\vec r) = -\dot\omega_{ij}(\vec 0) + \int_{\vec 0}^{\vec r} dr^k\left(\partial_i\dot\sigma_{jk} - \partial_j\dot\sigma_{ik}\right).$$

The integral is antisymetric in ij. If $\dot\omega_{ij}(\vec 0)$ is chosen antisymmetric in ij, $\dot\omega_{ij}(\vec r)$ will be antisymmetric for all $\vec r$.

d) Let $\dot\omega_{ij}(\vec r)$ be given as in c). We want to try to solve

$$\partial_i v_j = \dot\sigma_{ij} + \dot\omega_{ij}.$$

This will be possible, and the solution will be

$$v_j(\vec r) = V_j + \int_{\vec 0}^{\vec r} dr^i\left(\dot\sigma_{ij} + \dot\omega_{ij}\right),$$

as long as for each fixed j the integrability condition $\partial_k\partial_i v_j = \partial_i\partial_k v_j$ is satisfied, i.e.,

$$(\dagger) \qquad \partial_k\left(\dot\sigma_{ij} + \dot\omega_{ij}\right) = \partial_i\left(\dot\sigma_{kj} + \dot\omega_{kj}\right).$$

But $\dot\omega_{ij}$ was obtained by solving (*) on the preceding page. Substituting (*) in (†) makes (†) an identity.

23. a)

$$\begin{aligned}
U - U_0 &= \delta N \partial_N U + \delta \overleftrightarrow{G} \langle 2 \rangle \vec{\partial}_{\overleftrightarrow{G}} U \\
&\quad + \frac{1}{2} (\delta N)^2 \partial_N^2 U + \delta N \delta \overleftarrow{G} \langle 2 \rangle \vec{\partial}_{\overleftrightarrow{G}} \partial_N U \\
&\quad + \frac{1}{2} \left(\delta \overleftrightarrow{G} \, \delta \overleftrightarrow{G} \right) \langle 4 \rangle \vec{\partial}_{\overleftrightarrow{G}} \vec{\partial}_{\overleftrightarrow{G}} U + \text{ third order terms} \\
&= \delta N \theta_0 + \delta \overleftrightarrow{G} \langle 2 \rangle \frac{\overleftrightarrow{S}_0}{\rho_0} + \frac{1}{2} (\delta N)^2 B_0 \\
&\quad + \delta N \delta \overleftrightarrow{G} \langle 2 \rangle \frac{\overleftrightarrow{W}_0}{\rho_0} + \frac{1}{2} \left(\delta \overleftrightarrow{G} \, \delta \overleftrightarrow{G} \right) \langle 4 \rangle \frac{\overset{\leftrightarrow}{\overleftrightarrow{Q}}_0}{\rho_0} + \text{ third order term.}
\end{aligned}$$

b)

$$\begin{aligned}
\delta \theta &= \delta N B_0 + \delta \overleftrightarrow{G} \langle 2 \rangle \frac{\overleftrightarrow{W}_0}{\rho_0} \\
\delta \overleftrightarrow{S} &= \delta N \, \overleftrightarrow{W}_0 + \delta \overleftrightarrow{G} \langle 2 \rangle \, \overset{\leftrightarrow}{\overleftrightarrow{F}}_0 \, .
\end{aligned}$$

If $\delta\theta = 0$, then $\delta N = -\delta \overleftrightarrow{G} \langle 2 \rangle \frac{\overleftrightarrow{W}_0}{\rho_0 B_0}$, so

$$\delta \overleftrightarrow{S} = \delta \overleftrightarrow{G} \langle 2 \rangle \left(\overset{\leftrightarrow}{\overleftrightarrow{F}}_0 - \frac{\overleftrightarrow{W}_0 \overleftrightarrow{W}_0}{\rho_0 B_0} \right) .$$

Therefore

$$\overset{\leftrightarrow}{\overleftrightarrow{F}}{}'_0 = \overset{\leftrightarrow}{\overleftrightarrow{F}}_0 - \frac{\overleftrightarrow{W}_0 \overleftrightarrow{W}_0}{\rho_0 B_0}.$$

24. In an isotropic material,

$$\begin{aligned}
\overleftrightarrow{S}_0 &= -p_0 \, \overleftrightarrow{I} \\
\overleftrightarrow{W}_0 &= -\alpha \, \overleftrightarrow{I} \\
F_{ijkl} &= \lambda \delta_{ij} \delta_{kl} + \mu \left(\delta_{ik} \delta_{jl} + \delta_{il} \delta_{kj} \right) \\
Q_{ijkl} &= F_{ijkl} + \delta_{ij} S_{kl}^{(0)} - \delta_{jk} S_{il}^{(0)} \\
&= F_{ijkl} + p_0 \left(\delta_{jk} \delta_{il} - \delta_{ij} \delta_{kl} \right) \\
Q_{ijkl} &= (\lambda - p_0) \, \delta_{ij} \delta_{kl} + \mu \, \delta_{ik} \delta_{jl} + (\mu + p_0) \, \delta_{jk} \delta_{il}
\end{aligned}$$

a)

$$\begin{aligned}
U - U_0 &= \delta N \theta_0 - \frac{p_0}{\rho_0}\left(\operatorname{tr}\delta \overleftrightarrow{G}\right) + \frac{1}{2}(\delta N)^2 B_0 \\
&\quad -\frac{\alpha}{\rho_0}(\delta N)\left(\operatorname{tr}\delta \overleftrightarrow{G}\right) + \frac{1}{2\rho_0}\Big\{(\lambda - p_0)\left[\operatorname{tr}\delta \overleftrightarrow{G}\right]^2 + \mu\,\delta \overleftrightarrow{G}\langle 2\rangle \delta \overleftrightarrow{G} \\
&\quad + (\mu + p_0)\,\delta \overleftrightarrow{G}\langle 2\rangle \delta \overleftrightarrow{G}^T\Big\} + \text{ third order terms.}
\end{aligned}$$

b)

$$\begin{gathered}
\lambda' \delta_{ij}\delta_{kl} + \mu' \left(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}\right) \\
= \lambda \delta_{ij}\delta_{kl} + \mu\left(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}\right) - \tfrac{\alpha^2}{\rho_0 B_0}\delta_{ij}\delta_{kl}.
\end{gathered}$$

Thus $\mu' = \mu$, $\lambda' = \lambda - \alpha^2/\rho_0 B_0$.