Introduction to TENSOR ANALYSIS

H. D. BLOCK

Introduction to
Tensor Analysis

H. D. BLOCK *Professor of Applied Mathematics,*
Department of THEORETICAL AND APPLIED MECHANICS
Cornell University, Ithaca, New York

Corrected and Recopied 1978

Library of Congress Catalog Card Number: 62-9816

Printed in the United States of America

CONTENTS

PREFACE

This book is intended for those who have a working knowledge of elementary Vector Analysis and Matrix Theory. It is intended to be preparatory or collateral reading in connection with a more detailed study of one of the fields where tensors are used; for example: Continuum Mechanics, Hydrodynamics, Elasticity, Relativity, Vectorial Mechanics, Analytical Mechanics, Crystal Physics, Differential Geometry, or even Tensor Analysis itself from the physical point of view.

In books that have a physical orientation, the current definition of a tensor is that it is a set of quantities which transform according to a certain, rather complicated, law. I have found that this definition leaves the student rather dazed, unmotivated, and permanently insecure in his relations with tensors. On the other extreme, the purely mathematical presentation in terms of linear functionals, dual spaces, or multilinear algebras is usually too abstract and too remote from the student's background in physics or engineering to enable him to comprehend the subject in such a way that it becomes useful to him as a working tool

The development given here is based on the *polyadic*, an entity which has all but disappeared from Mathematics. Since the concept was introduced by Willard Gibbs (who thereby initiated the subject of tensor analysis). the author makes no claims to originality. It does appear that this method of introduction gives the student a feeling of "knowing what he is doing." After reading and understanding this book, my students tell me that they can read books or papers in the engineering and scientific fields without encountering any difficulties with the tensor aspects of the subject.

The mathematical theory of the exterior calculus is, for the sake of brevity, not included here, although it seems clear that the day is not far off when this will be incorporated into the first course in Calculus. Some references to this material are given at the end of the book. A lucid exposition of the exterior calculus was given by the late Professor A. N. Milgrim in his lectures at the Institute of Technology of the University of Minnesota. The writer hopes to assemble this material soon and to make it available in printed form.

Before starting to read this book the student should have been exposed to the elements of Vector Analysis and Matrix Theory at least to the extent that these topics are covered in the currently standard course in "Advanced Mathematics for Engineers and Scientists." For example the material in:

Mathematics of Physics and Modern Engineering
Sokolnikoff, I. S. and Redheffer, R. M.
New York: McGraw-Hill, 1958

is sufficient background. Other texts of a similar nature are

Advanced Engineering Mathematics
Wylie, C. R., Jr.
New York: McGraw-Hill, 1960

Applied Mathematics for Engineers and Physicists
Pipes, L. A.
New York and London, McGraw-Hill, 1958

Mathematical Methods for Scientists and Engineers
Smith, L. P.
New York, Prentice-Hall, 1953

The Mathematics of Physics and Chemistry
Margenau, H. and Murphy, G. M.
New York, Van Nostrand, 1943

Methods of Mathematical Physics
Jeffreys, H. and Jeffreys, B. S.
Third Edition, New York, Cambridge University Press, 1956

Elements of Pure and Applied Mathematics
Lass H.
New York, McGraw-Hill, 1957

Alternatively, the reader will have acquired the necessary background if he is familiar with the contents of the elementary courses on Vector Analysis and Linear Algebra and Matrix Theory. Some standard texts on Vector Analysis are:

Vector Analysis
Brand, L.
New York, Wiley, 1957

Vector Analysis
Coffin, J. G.
New York, J. Wiley, 1911

An Introduction to Vector Analysis
Hague, B.
London, Methuen, 5th Edition, 1951

Vector Analysis
Phillips, H. B.
New York, J. Wiley, 1933

Vector Analysis with Applications to Geometry and Physics
Schwartz, M; Green, Simon; Rutledge, W. A.
New York, Harper, 1960

In addition, the books cited by the following authors under "Further Reading", page 62, offer not only further reading, but also an excellent introduction to Vector Analysis.

Brand, L
Craig, H. V.
Gans, R.
Gibbs, J. W.
Hay, G. E.
Lass, H.
Rutherford, D. E.
Spiegel, M. R.
Weatherburn, C. E.
Wills, A. P.

Examples of texts covering Linear Algebra and Matrix Theory are:

Introduction to Modern Algebra and Matrix Theory
Beaumont, R. A. and Ball, R. W.
New York, Rinehart and Company, 1954

Linear Algebra
Hadley, G.
Reading, Mass., Addison-Wesley, 1961

Linear Algebra
Hoffman, K. and Kunze, R.
Englewood Cliffs, N. J., Prentice-Hall, 1961

Elementary Matrix Algebra
Hohn, F. E.
New York, MacMillan, 1960

Linear Algebra for Undergraduates
Murdoch, D. C.
New York, Wiley, 1957

Theory of Matrices
Perlis, S.
Cambridge, Mass., Addison Wesley Press, 1952

Linear Algebra and Matrix Theory
Stoll, R. R.
New York, McGraw-Hill, 1952

Vector Spaces and Matrices
Thrall, R. M. and Tornheim, L.
New York, Wiley, 1957

CHAPTER ONE

WHAT IS A VECTOR?

You will find in most books on tensor analysis in physics or engineering that a vector is defined as a set of numbers (its components in a particular coordinate system) and a recipe by which these components change when a different coordinate system is used as a reference. This seems to be a most unintuitive approach. Admittedly it is easy to understand the emphasis on components, since real numbers are all that can be read on a slide rule, which in a certain sense, contains all the answers to engineering problems. The same may also be said about more modern computers. Thus the components are, in the final analysis, the objects of interest. On the other hand, compare the divergence theorem in terms of vectors

$$\int_\tau \nabla \cdot \mathbf{V}\, d\tau = \int_\sigma \mathbf{V} \cdot \hat{\mathbf{n}}\, d\sigma \tag{1.1}$$

with the corresponding theorem in terms of the scalar components

$$\int_\tau \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}\right) dxdydz = \int_\sigma P dydz + Qdzdx + Rdxdy \tag{1.1'}$$

or Stokes' theorem in terms of vectors

$$\int_{\sigma} \nabla \times \mathbf{V} \cdot \hat{\mathbf{n}}\, d\sigma = \int_{C} \mathbf{V} \cdot d\mathbf{R} \tag{1.2}$$

with the corresponding theorem in terms of scalar components.

$$\int_{\sigma} \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right) dydz + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right) dzdx + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dxdy$$

$$= \int_{C} Pdx + Qdy + Rdz \tag{1.2'}$$

We believe it is clear that the vector formulations Eqs. (1.1) and (1.2) have much greater intuitive content than the scalar formulations Eqs. (1.1') and (1.2'). Thus the vector formulations are not only easier to remember and work with, but also offer the advantage that the engineer or scientist can, on the basis of their physical meaning, modify them to fit his needs in particular applications.

Thus there is much to be gained in the way of physical insight by using a vector approach. Consequently we shall regard the vector as the primary object of interest. We will find the usual necessary laws by which the components change when the reference frame is changed; but to make this the *definition* of a vector seems to us to be unnecessarily confusing.

Of course we shall also study scalar components, as this is the form in which numerical answers to problems appear; but for formulation of physical facts (which have no real dependence on the coordinate system) it seems to us to be generally preferable to use an equation which does not involve the coordinates with respect to a particular system. Such a formulation is called "invariant."

The above remarks also apply to tensors, but if the student has not met them yet, the cogency of these remarks is not apparent at this time.

CHAPTER TWO

ALGEBRA

The student will recall that in the study of vector analysis he first learned about the *algebra* of vectors and then studied vector fields. Similarly here, although we are mainly interested in *fields*, we will first consider the algebra of the objects of interest.

We take the position that a *vector is a vector*. In other words, we do not attempt to define a vector in more primitive terms, but only in terms of certain properties which will be postulated about them. Of course we can (and will) give examples, which constitute the applications and the real point of the theory, but fundamentally a vector is a vector. We can yield a little and say a vector is an element of a vector space. Now we must define a vector space.

An (abstract) vector space is a collection of objects (vectors) whose nature is unspecified except that they must have the following properties. Any two vectors can be added to yield a third vector; and the sum is independent of the order of the summands. In symbols: if $\mathbf{u}$ and $\mathbf{v}$ are vectors then $\mathbf{u} + \mathbf{v}$ is another vector and $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$. The associative law: $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ is also assumed; thus either of these expressions can be written unambiguously as $\mathbf{u} + \mathbf{v} + \mathbf{w}$. Any vector can be multiplied by a scalar (a real number*) to produce another vector. Thus, if $\mathbf{v}$ is a vector and c is a scalar, then $c\mathbf{v}$ is a vector. We also assume that $c\mathbf{v} = \mathbf{v}c$, that $1\mathbf{v} = \mathbf{v}$ and that $c(d\mathbf{v}) = (cd)\mathbf{v}$. We also require the distributive laws $(c + d)\mathbf{v} = c\mathbf{v} + d\mathbf{v}$ and $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$. We also postulate the existence of a vector $\mathbf{0}$ such that for every vector $\mathbf{v}, \mathbf{v} + \mathbf{0} = \mathbf{v}$, and $0\mathbf{v} = \mathbf{0}$.** It follows that $\mathbf{v} + (-1)\mathbf{v} = \mathbf{0}$, so that we write $\mathbf{u} - \mathbf{v}$ for $\mathbf{u} + (-1)\mathbf{v}$.

Examples.

a. The set of real numbers, using ordinary addition, satisfies the above conditions and is thus a particular vector space.
b. The set of complex numbers, using ordinary addition of complex numbers, is also a vector space.
c. The set of vectors in everyday (three dimensional) space, as studied in engineering, constitutes a vector space.
d. The set of n-tuples: $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, where the x_i are real numbers, with the law of addition

$$\mathbf{x} + \mathbf{y} = (x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n)$$
$$= (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$$

and the law for multiplication by scalars

$$c\mathbf{x} = c(x_1, x_2, \ldots, x_n) = (cx_1, cx_2, \ldots, cx_n)$$

is also a vector space.
e. The set of functions $\mathbf{v} = v(t)$ defined and differentiable on the interval $0 \leqq t \leqq 1$, with the rule of addition $\mathbf{u} + \mathbf{v} = (u + v)(t) = u(t) + v(t)$, and the rule for multiplication by scalars given by $c\mathbf{v} = (cv)(t) = c[v(t)]$ is also a vector space. The student should verify this.
f. The set of functions $\mathbf{v} = v(x, y, z)$ of three variables defined and analytic over a bounded region R of 3-space and vanishing on the boundary, with the definition of addition and multiplication by scalars analogous to that given in (e), is also a vector space.

---

*The vector space we define here is a "vector space over the real numbers." The scalars might instead be taken to be the complex numbers, in which case we would have a "vector space over the complex numbers." Similarly one might have a "vector space over the quaternions," and so on.

**That is, when any vector is multiplied by the *number* zero, the result is the *zero vector*.

CHAPTER THREE

DYADS, DYADICS, TENSORS

Suppose we have an abstract vector space and suppose we start writing down vectors in this space as if they were ordinary algebraic quantities. For example we might write down **uv**. This would be called (following Gibbs) a *dyad*. If we formed a linear combination of such terms $c_1 \mathbf{u}_1 \mathbf{v}_1 + c_2 \mathbf{u}_2 \mathbf{v}_2 + \ldots + c_k \mathbf{u}_k \mathbf{v}_k$ we would have (in the language of Gibbs) a *dyadic expression* or, more briefly, a *dyadic*. Today we call it a tensor of order two. A product of three vectors **uvw** is a *triad*, and a linear combination $\Sigma c_k \mathbf{u}_k \mathbf{v}_k \mathbf{w}_k$ a *triadic*. Today we call it a tensor of order three. Similarly **uvwz** is a tensor of order four, and so on. Such products are called "open" products or "tensor" products of the vectors. They are also called *polyads*, and linear combinations of them are called *polyadics*.

In this way we generate an algebraic system consisting of open products of vectors and their linear combinations. All the usual multiplicative rules of algebra except for the commutative law, are to hold. For example, for any vectors **u**, **v**, **w** and scalars *c*, *d*, the associative laws

$$c(\mathbf{uv}) = (c\mathbf{u})\mathbf{v} = c\mathbf{uv}$$
$$(\mathbf{uv})\mathbf{w} = \mathbf{u}(\mathbf{vw}) = \mathbf{uvw}$$
$$(c\mathbf{u})(d\mathbf{v}) = (cd)(\mathbf{uv})$$

and the mixed distributive laws

$$
\begin{aligned}
\mathbf{u}(\mathbf{v} + \mathbf{w}) &= \mathbf{uv} + \mathbf{uw} \\
(\mathbf{u} + \mathbf{v})\mathbf{w} &= \mathbf{uw} + \mathbf{vw}, \\
c(\mathbf{uv} + \mathbf{wz}) &= c\mathbf{uv} + c\mathbf{wz}
\end{aligned}
$$

and the unitary law $1(\mathbf{uv}) = \mathbf{uv}$ are valid but we do *not* assume that $\mathbf{uv} = \mathbf{vu}$. The scalars, however, do commute and $\mathbf{u}(c\mathbf{v}) = (c\mathbf{u})\mathbf{v} = c\mathbf{uv}$.

Except for the fact that the commutative law is not assumed to hold and the fact that we do not divide by vectors, we can manipulate these objects just as in ordinary algebra. We could, as we did for vectors, supply a logically rigorous definition of (abstract) tensors in terms of their algebraic properties cited above. These properties would define these objects, just as the rules of chess define a knight or a rook. We find that the "concrete" definition in terms of open products of vectors is easier for the student to grasp.

Now that we have this collection of objects, the reader may ask, what do they represent? For the present they are fundamental objects, like the vectors, and we do not try to define them in more primitive terms. Later we shall investigate the physical significance of such objects.

CHAPTER FOUR

LINEAR INDEPENDENCE; DIMENSION

A set of vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are said to be *linearly dependent* if a linear combination $c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \ldots + c_n \mathbf{v}_n = \mathbf{0}$, with not all of the c_i's = 0. {This is equivalent to saying that one of the $\mathbf{v}_i$'s can be expressed as a linear combination of the others, for if $c_1 \neq 0$, then $\mathbf{v}_1 = \frac{-1}{c_1}[c_2 \mathbf{v}_2 + c_3 \mathbf{v}_3 + \ldots + c_n \mathbf{v}_n]$. Conversely if $\mathbf{v}_1 = d_2\mathbf{v}_2 + \ldots + d_n\mathbf{v}_n$, then

$$\sum_{j=1}^{n} d_j \mathbf{v}_j = \mathbf{0}$$

with $d_1 = -1 \neq 0$.} A set is *linearly independent* if it is not linearly dependent.

If there exists in a vector space a set of n linearly independent vectors but no set of $(n + 1)$ linearly independent vectors, then the vector space is said to be *n-dimensional*.

For example, the reader may verify that the vector spaces of the Examples on page 4 are, respectively, of dimension 1, 2 (over the reals), 3, n, ∞, ∞.

CHAPTER FIVE

COMPONENTS: CHANGE OF BASIS

In an n-dimensional vector space a set of n linearly independent vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ is called a *basis*. Any vector $\mathbf{v}$ in the space can be expressed as a unique linear combination of the basis vectors: $\mathbf{v} = v^1\mathbf{e}_1 + \ldots + v^n\mathbf{e}_n$, or using the summation convention, (in which an index occurring as both a subscript and a superscript is automatically summed from 1 to n, unless otherwise specified) $\mathbf{v} = v^i\mathbf{e}_i$.

Exercise. Prove the above. [Hint. To prove that the representation is always possible, show first that there are scalars c_i not all zero such that $c_0\mathbf{v} + c_1\mathbf{e}_1 + c_2\mathbf{e}_2 + \ldots + c_n\mathbf{e}_n = \mathbf{0}$. Next show that $c_0 \neq 0$; hence, solve for $\mathbf{v}$. To prove uniqueness, suppose $\mathbf{v} = v^i\mathbf{e}_i = V^i\mathbf{e}_i$. Hence $(v^i - V^i)\mathbf{e}_i = \mathbf{0}$. Now use the linear independence of the basis vectors.]

The scalars v^i are called the *components* of the vector $\mathbf{v}$ with respect to the basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$.

Suppose we have another basis $\{\bar{\mathbf{e}}_1, \bar{\mathbf{e}}_2, \ldots, \bar{\mathbf{e}}_n\}$. The same vector $\mathbf{v}$ can also be expressed as a linear combination of the elements in the new basis:

$$\mathbf{v} = \bar{v}^i\bar{\mathbf{e}}_i$$

Since $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is a basis, we can express each of the new basis elements $\bar{\mathbf{e}}_i$ in terms of the $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ and conversely. Thus

$$\bar{\mathbf{e}}_j = a^i{}_j \mathbf{e}_i \tag{5.1}$$

We regard $a^i{}_j$ as the element in the i^{th} row and j^{th} column of a matrix A. If we apply the usual rule of matrix multiplication of a matrix by the column of vectors

$$\begin{pmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \vdots \\ \mathbf{e}_n \end{pmatrix}$$

we might regard Eq. (5.1) as stating

$$\bar{\mathbf{e}}_* = A^T \mathbf{e}_*, \text{ where } \mathbf{e}_* = \begin{pmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \vdots \\ \mathbf{e}_n \end{pmatrix}, \ \bar{\mathbf{e}}_* = \begin{pmatrix} \bar{\mathbf{e}}_1 \\ \bar{\mathbf{e}}_2 \\ \vdots \\ \bar{\mathbf{e}}_n \end{pmatrix} \tag{5.2}$$

Let us see what effect the transformation of basis (5.1) has on the coordinates. Since the same vector $\mathbf{v}$ is involved

$$\mathbf{v} = v^i \mathbf{e}_i = \bar{v}^j \bar{\mathbf{e}}_j = \bar{v}^j a^i{}_j \mathbf{e}_i$$

Since the vectors $\mathbf{e}_i$ are linearly independent, we have

$$v^i = a^i{}_j \bar{v}^j \tag{5.3}$$

or

$$v^* = A\bar{v}^*, \text{ where } v^* = \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}, \ \bar{v}^* = \begin{pmatrix} \bar{v}^1 \\ \vdots \\ \bar{v}^n \end{pmatrix}$$

Since Eq. (5.1) expresses the new (barred) basis vectors in terms of the old (unbarred) we would like to do the same for the components. Thus

$$\bar{v}^* = A^{-1} v^* \tag{5.4}$$

Compare Eq. (5.4) with Eq. (5.2). If A is an orthogonal matrix ($A^{-1} = A^T$), then the components follow the same transformation law as the base vectors. In the general case, the law of transformation of the components is given by the inverse transposed matrix of the law describing the basis change. These components are therefore called *contravariant*. [Most writers call them the components of a contravariant vector. Under certain circumstances this terminology is justified. (Cf. the "Remark" in chapter 9) But from the way we shall develop the subject the vector is the vector. It is the components which are contravariant.]

If we let $B = A^{-1}$ and let $b^i{}_j$ be the element in the i^{th} row and j^{th} column of B, then Eq. (5.4) can be written

$$\bar{v}^i = b^i{}_j v^j \quad \text{or} \quad \bar{v}^* = Bv^* \tag{5.5}$$

This follows from Eq. (5.4) immediately. To give the student a little practice at the kind of manipulation to come, we derive Eq. (5.5) from Eq. (5.3) as follows. Multiply each side of Eq. (5.3) by $b^k{}_i$ and sum on i. Then, letting $\delta_j{}^i$ have the value $+1$ if $i = j$ and the value 0 if $i \neq j$ (the "Kronecker delta") we get

$$b^k{}_i v^i = \bar{v}^j a^i{}_j b^k{}_i = \bar{v}^j \delta^k{}_j = \bar{v}^k, \text{ as in Eq. (5.5).}$$

By the same method we can solve Eq. (5.1) for the $\mathbf{e}_i$ in terms of the $\bar{\mathbf{e}}_i$

$$b^j{}_k \bar{\mathbf{e}}_j = b^j{}_k a^i{}_j \mathbf{e}_i = \delta^i{}_k \mathbf{e}_i = \mathbf{e}_k$$

Hence

$$\mathbf{e}_i = b^j{}_i \bar{\mathbf{e}}_j \tag{5.6}$$

In terms of a given basis $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ any dyadic $\mathbf{T} = c^{ij}\mathbf{u}_i\mathbf{v}_j$, where $\mathbf{u}_i = u^j{}_i\mathbf{e}_j$; $\mathbf{v}_i = v^j{}_i\mathbf{e}_j$ and c^{ij} are scalars, can be expressed in the form

$$\mathbf{T} = c^{ij} u^k{}_i \mathbf{e}_k v^s{}_j \mathbf{e}_s = c^{ij} u^k{}_i v^s{}_j \mathbf{e}_k \mathbf{e}_s$$

i.e. $\mathbf{T}$ can be expressed in the canonical form

$$\mathbf{T} = t^{ij}\mathbf{e}_i\mathbf{e}_j$$

The scalars t^{ij} are called the components with respect to the given basis. In terms of a new basis $\{\bar{\mathbf{e}}_1, \bar{\mathbf{e}}_2, \ldots, \bar{\mathbf{e}}_n\}$ we have the same tensor $\mathbf{T}$ expressed as

$$\mathbf{T} = \bar{t}^{ij}\bar{\mathbf{e}}_i\bar{\mathbf{e}}_j$$

where $\bar{t}^{ij}$ are the components in terms of the new basis.

How are the new components $\bar{t}^{ij}$ related to the old components t^{ij}? If the transformation of the basis is given by Eq. (5.1) $[\bar{e}_* = A^T e_*]$,

$$\mathbf{T} = \bar{t}^{ij}\bar{\mathbf{e}}_i\bar{\mathbf{e}}_j = t^{rs}\mathbf{e}_r\mathbf{e}_s = t^{rs} b^i{}_r \bar{\mathbf{e}}_i b^j{}_s \bar{\mathbf{e}}_j$$

from which

$$\bar{t}^{ij} = b^i{}_r b^j{}_s t^{rs} \tag{5.7}$$

If the elements t^{ij} are regarded as the elements of a matrix T, and $\bar{t}^{ij}$ of a matrix $\bar{T}$, then Eq. (5.7) can be written

$$\bar{T} = BTB^T = A^{-1}T(A^{-1})^T \tag{5.7'}$$

Equation (5.7), or (5.7'), gives the law of transformation of the components t^{ij} of the tensor **T** when the basis vectors are changed according to Eq. (5.1). They are called the contravariant components of the tensor (not, as is usual, the components of a contravariant tensor). (Cf. "Remarks" in Chapter 9.)

Similarly for a third order tensor $t^{ijk}\mathbf{e}_i\mathbf{e}_j\mathbf{e}_k$ the law of transformation of the components is

$$\bar{t}^{ijk} = b^i{}_r b^j{}_s b^k{}_\mu t^{rs\mu} \tag{5.8}$$

For third and higher order tensors the matrix notation, as in Eq. (5.7') is no longer convenient, and the notation of Eqs. (5.7) and (5.8) is employed instead.

To correlate this with the field development later, suppose we consider a vector **x** written in the form $\mathbf{x} = x^i\mathbf{e}_i = \bar{x}^i\bar{\mathbf{e}}_i$. Then, by Eqs. (5.3) and (5.5)

$$x^i = a^i{}_j\bar{x}^j \tag{5.9}$$

$$\bar{x}^i = b^i{}_j x^j \tag{5.10}$$

Then Eqs. (5.5), (5.7), and (5.8) can be written in the form

$$\bar{v}^i = \frac{\partial\bar{x}^i}{\partial x^j}v^j \tag{5.11}$$

$$\bar{t}^{ij} = \frac{\partial\bar{x}^i}{\partial x^r}\frac{\partial\bar{x}^j}{\partial x^s}t^{rs} \tag{5.12}$$

and

$$\bar{t}^{ijk} = \frac{\partial\bar{x}^i}{\partial x^r}\frac{\partial\bar{x}^j}{\partial x^s}\frac{\partial\bar{x}^k}{\partial x^\mu}t^{rs\mu} \tag{5.13}$$

since

$$b^i{}_j = \frac{\partial\bar{x}^i}{\partial x^j}$$

Notice in the equations of this section how the free indices on each side of the equations balance (like dimensions do) and how the dummy indices occur always in a subscript and superscript pair. This will make it easier to remember and to check equations.

The notation in this section was chosen so as to render it easy to remember in later sections. Thus A (the first letter of the alphabet) was used in Eq. (5.1) to describe the first change of basis. For the corresponding law of transformation of the components, which is "backwards" to this [Eq. (5.5)], the letter B is used. The notation introduced in chapter 5 will be used throughout.

CHAPTER SIX

LINEAR TRANSFORMATIONS

In general, a system of equations

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & & \vdots \\ c_{n1} & \cdots & c_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

or, in matrix notation $\quad y = Cx$

may arise from a change of coordinates (as in the previous section), in which case no vectors move, only the names are changed. On the other hand such a system might also represent a linear mapping or deformation of the space, in which the vector whose components are x is sent into a new vector whose components, with respect to the same basis, are now y. Considerable confusion has arisen in this connection since the former interpretation is often called a transformation of coordinates and the latter a transformation of the vectors. By linquistic attrition the word *transformation* has come to be used for both. By the word "transformation" physicists generally mean a change in coordinates, leaving the vectors unchanged; mathematicians generally mean a change in the vectors, leaving the coordinate system unchanged; engineers mean either one.

The words "alias" (only the name is changed) and "alibi" (the vector goes somewhere else) have been proposed to distinguish these two interpretations, but seem to be too literary to be adopted widely. We will discuss in this section linear transformations in the alibi sense and to make matters clear we will call them *linear mappings*. After that we will come back to linear transformations in the alias sense and for clarity will call them a *linear change of bases*, or *linear change in coordinates*.

Suppose we have a mapping defined on a vector space. That is, with each vector $\mathbf{v}$ we associate a transformed vector $\mathbf{w} = \mathcal{J}(\mathbf{v})$. The function $\mathcal{J}$ is said to be *additive* if $\mathcal{J}(\mathbf{u}+\mathbf{v}) = \mathcal{J}(\mathbf{u}) + \mathcal{J}(\mathbf{v})$; *homogeneous* if $\mathcal{J}(c\mathbf{u}) = c\,\mathcal{J}(\mathbf{u})$ where c is any scalar. If $\mathcal{J}$ is additive and homogeneous, then $\mathcal{J}$ is called a *linear* mapping.

Let $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ be a given basis, and let $\mathcal{J}$ be a linear mapping. Then

$$\mathcal{J}(\mathbf{v}) = \mathcal{J}(v^i\mathbf{e}_i) = v^i\,\mathcal{J}(\mathbf{e}_i) \tag{6.1}$$

Hence if we know what $\mathcal{J}$ does to the basis vectors we know what it does to every vector.

Since the $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ are a basis, we can express the vector $\mathcal{J}(\mathbf{e}_i)$ as a linear combination of the basis elements:

$$\mathcal{J}(\mathbf{e}_i) = f^j{}_i\mathbf{e}_j \tag{6.2}$$

If we let $\mathbf{w} = w^i\mathbf{e}_i = \mathcal{J}(\mathbf{v})$, then $w^i\mathbf{e}_i = v^i f^j{}_i\mathbf{e}_j$, or

$$w^i = f^i{}_j v^j \tag{6.3}$$

That is, in terms of the given basis, the transformation $\mathcal{J}$ can be represented by the components changing according to Eq. (6.3). In other words if F represents the matrix $f^i{}_j$, then the mapping $\mathcal{J}$ can be specified by the statement that the vector with the components v^i goes over into a vector with components w^i given by Eq. (6.3); i.e.

$$w = Fv \tag{6.4}$$

This is the alibi interpretation of the matrix F. So far all components have referred to the given basis $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$. Suppose now we consider the same transformation $\mathcal{J}$ but relative to a new basis $\{\bar{\mathbf{e}}_1, \bar{\mathbf{e}}_2, \ldots, \bar{\mathbf{e}}_n\}$ given in terms of the old basis by Eq. (5.1). Now

$$\mathbf{w} = \bar{w}^i \bar{\mathbf{e}}_i = w^i \mathbf{e}_i = v^j f^i{}_j \mathbf{e}_i = v^j f^i{}_j b^k{}_i \bar{\mathbf{e}}_k = a^j{}_r \bar{v}^r f^i{}_j b^k{}_i \bar{\mathbf{e}}_k$$

Hence

$$\bar{w}^i = a^j{}_r f^k{}_j b^i{}_k \bar{v}^r$$

In matrix notation

$$\bar{w} = BFA\bar{v} = A^{-1} FA\bar{v} = BFB^{-1}\bar{v}$$

That is, the matrix representing the transformation in the new coordinate system ($\bar{v} = Bv$) is BFB^{-1}, where F is the matrix representing the same transformation in the original coordinate system. Two such matrices are said to be *similar*. In other words, two similar matrices represent the same transformation, but described in terms of different coordinate systems.

The action of $\mathcal{F}$ on certain vectors will be to map them into a scalar multiple of themselves. This is always true for the vector $\mathbf{0}$. Any non-zero vector $\mathbf{v}$ such that $\mathcal{F}(\mathbf{v}) = \lambda\mathbf{v}$ is called an *eigenvector* of $\mathcal{F}$ and the scalar λ an *eigenvalue*. In terms of the matrix representation Eq. (6.4) we have

$$Fv = \lambda v \tag{6.5}$$

The eigenvalues λ satisfy the equation

$$|F - \lambda I| = 0 \tag{6.6}$$

If another basis were to be used, the eigenvalues and eigenvectors would not be changed, because these depend on the mapping, not on the coordinate system. Therefore the eigenvalues satisfy also the characteristic equation

$$|BFB^{-1} - \lambda I| = 0 \tag{6.7}$$

[These facts may also be seen directly, since Eq. (6.5) may be written $FA\bar{v} = \lambda A\bar{v}$, where $v = A\bar{v}$, i.e. $BFB^{-1}\bar{v} = \lambda\bar{v}$, or $\bar{F}\bar{v} = \lambda\bar{v}$. Also Eq. (6.7) can be written

$$|BFB^{-1} - \lambda I| = |B(F - \lambda I)B^{-1}| = |F - \lambda I| .]$$

Now

$$|F - \lambda I| = (-1)^n [\lambda^n - \mathcal{I}_1 \lambda^{n-1} + \mathcal{I}_2 \lambda^{n-2} \ldots + (-1)^n \mathcal{I}_n] = |BFB^{-1} - \lambda I| \tag{6.8}$$

The quantities $\mathcal{I}_1, \ldots, \mathcal{I}_n$ are the same no matter which coordinate system is used. These are called invariants of the mapping $\mathcal{F}$. $\mathcal{I}_j$ is the sum of the principal minors of order j. In particular $\mathcal{I}_1$ is the trace, $\mathcal{I}_n$ is the determinant. These quantities do not change under similarity transformations of the matrix. They are a characteristic of the *mapping,* rather than only of the matrix. The sum of the principal minors of order j, $\mathcal{I}_j$, is also equal to the sum of the products of all the eigenvalues taken j at a time. In particular the trace, $\mathcal{I}_1$, is the sum of the eigenvalues; the determinant, $\mathcal{I}_n$, is the product of the eigenvalues.

CHAPTER SEVEN

INNER PRODUCT: METRIC TENSOR

An *inner,* or *dot, product* is a real* valued function of two vectors which is a linear function of each vector. That is, for any pair of vectors $\mathbf{u}$, $\mathbf{v}$ there is associated a scalar which we denote by $\mathbf{u} \cdot \mathbf{v}$. For this function to be an inner product, it must have the following properties.

1. Symmetric: $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$.
2. Additive: $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$.
3. Homogeneous $(c\mathbf{u}) \cdot \mathbf{v} = c(\mathbf{u} \cdot \mathbf{v})$, where c is any scalar.
4. Positive: $\mathbf{v} \cdot \mathbf{v} > 0$, for $\mathbf{v} \neq \mathbf{0}$.
   It follows that $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$ and that $\mathbf{u} \cdot (c\mathbf{v}) = c(\mathbf{u} \cdot \mathbf{v})$.

Examples. We can define an inner product for each of the examples given on page 4 as follows:

a. The inner product is simply the product of the two numbers.

b. $(a + ib) \cdot (c + id) = ac + bd$.

---

*Under other circumstances the inner product is sometimes taken to be complex valued. In this case the symmetric condition [(1) above] is replaced by the Hermitian Symmetry: $x \cdot y = \overline{y \cdot x}$.

c. $\mathbf{V}_1 \cdot \mathbf{V}_2 = |\mathbf{V}_1||\mathbf{V}_2| \cos(\mathbf{V}_1, \mathbf{V}_2)$.

Or $(x\mathbf{i} + y\mathbf{j} + z\mathbf{k})\cdot(\xi\mathbf{i} + \eta\mathbf{j} + \zeta\mathbf{k}) = x\xi + y\eta + z\zeta$.

d. $(x_1, x_2, \dots, x_n) \cdot (y_1, y_2, \dots, y_n) = \sum_i x_i y_i$.

e. $\mathbf{u}\cdot\mathbf{v} = \int_0^1 u(t)v(t)\,dt$.

f. $\mathbf{u}\cdot\mathbf{v} = \iiint_R u(x, y, z)v(x, y, z)\,dxdydz$.

The student should verify that the defining properties [(1)–(4) above] hold in each case.

By property (4), $\mathbf{v}\cdot\mathbf{v} \geqq 0$, so that $\sqrt{\mathbf{v}\cdot\mathbf{v}}$ is real; we call it the *length* (or norm) of $\mathbf{v}$

$$||\mathbf{v}|| = \sqrt{\mathbf{v}\cdot\mathbf{v}}$$

[The other features characterizing the norm are given in Eqs. (7.13) and (7.14) below.]

Let $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ be a basis. Let $\mathbf{e}_i \cdot \mathbf{e}_j = g_{ij}$. (7.5)

The length of an arbitrary vector $\mathbf{v} = v^i\mathbf{e}_i$ is then given by

$$||\mathbf{v}||^2 = v^i\mathbf{e}_i \cdot v^j\mathbf{e}_j = v^i v^j g_{ij} \tag{7.6}$$

The numbers g_{ij} are called the (covariant) *metric coefficients* (with respect to the given basis). Note that in view of Eqs. (7.5) and property (1), $g_{ij} = g_{ji}$. We denote the matrix by g.

Note that $||\mathbf{e}_i|| = \sqrt{g_{ii}}$. Sometimes it is desired to express the vector $\mathbf{v}$ in terms of the normalized basis $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \dots, \hat{\mathbf{e}}_n\}$ where $\hat{\mathbf{e}}_i = \dfrac{\mathbf{e}_i}{||\mathbf{e}_i||}$, so that $||\hat{\mathbf{e}}_i|| = 1$.

Thus $\mathbf{v} = v^i\mathbf{e}_i = V^i\hat{\mathbf{e}}_i$, where

$$V^i = ||\mathbf{e}_i||v^i = \sqrt{g_{ii}}\, v^i \text{ (no sum on } i)$$

Such components V^i are called the "physical components" of $\mathbf{v}$ with respect to the given basis.

Suppose that $\{\bar{\mathbf{e}}_1, \bar{\mathbf{e}}_2, \dots, \bar{\mathbf{e}}_n\}$ is another basis as in Eq. (5.1). Let $\bar{g}_{ij} = \bar{\mathbf{e}}_i \cdot \bar{\mathbf{e}}_j$. How does $\bar{g}_{ij}$ compare with the g_{ij}? We have $\bar{g}_{ij} = \bar{\mathbf{e}}_i \cdot \bar{\mathbf{e}}_j = a^r{}_i\mathbf{e}_r \cdot a^s{}_j\mathbf{e}_s$, or

$$\bar{g}_{ij} = a^r{}_i a^s{}_j g_{rs} \tag{7.7}$$

Notice how Eq. (7.7) differs from Eq. (5.7); or, for comparison with Eq. (5.10), we can rewrite Eq. (7.7) [using Eq. (5.9)] as

$$\bar{g}_{ij} = \frac{\partial x^r}{\partial \bar{x}^i} \frac{\partial x^s}{\partial \bar{x}^j} g_{rs}$$

This is the way the covariant metric coefficients transform when a new basis is introduced. The mnemonic "Co go below; and the bars also" is helpful.

(The reader who is interested in getting on with the main development should proceed directly to Chapter 8.) To illustrate the beauty of the abstract method let us derive Schwartz's inequality. Since we know nothing about the vectors, aside from what has been explicitly postulated about them, the derivation must be based on only these few properties.

Let **u** and **v** be vectors and λ a scalar. Then from property (4) we have $(\mathbf{u} + \lambda\mathbf{v})\cdot(\mathbf{u} + \lambda\mathbf{v}) > 0$, if $\mathbf{u} \neq -\lambda\mathbf{v}$.

Using properties (3) and (2) we get

$$||\mathbf{u}||^2 + 2\lambda\mathbf{u}\cdot\mathbf{v} + \lambda^2 ||\mathbf{v}||^2 > 0 \qquad (7.8)$$

for any choice of λ. Let us choose λ so as to make the left side as small as possible. Set

$$\frac{d}{d\lambda}(||\mathbf{u}||^2 + 2\lambda\mathbf{u}\cdot\mathbf{v} + \lambda^2 ||\mathbf{v}||^2) = 2\mathbf{u}\cdot\mathbf{v} + 2\lambda\, ||\mathbf{v}||^2 = 0$$

That is, let us select $\lambda = \dfrac{-\mathbf{u}\cdot\mathbf{v}}{||\mathbf{v}||^2}$. Then Eq. (7.8) gives

$$||\mathbf{u}||^2 - 2\frac{(\mathbf{u}\cdot\mathbf{v})^2}{||\mathbf{v}||^2} + \frac{(\mathbf{u}\cdot\mathbf{v})^2}{||\mathbf{v}||^2} > 0, \text{ for } \mathbf{v} \neq \mathbf{0}$$

Hence

$$|\mathbf{u}\cdot\mathbf{v}| < ||\mathbf{u}|| \cdot ||\mathbf{v}|| \quad \text{if } \mathbf{u} \neq \lambda\mathbf{v} \text{ and } \mathbf{v} \neq \mathbf{0}$$

In every case we have Schwartz's inequality

$$|\mathbf{u}\cdot\mathbf{v}| \leq ||\mathbf{u}|| \cdot ||\mathbf{v}|| \qquad (7.9)$$

Since the examples given earlier, (pg. 4, 17) satisfy the specified conditions (1)—(4), it follows that Schwartz's inequality must be valid also in these particular cases. Thus in particular *we have proved* [from (d), (e), and (f)]

$$\left| \sum_i x_i y_i \right| \leq \sqrt{\sum_i x^2_i} \sqrt{\sum_i y^2_i} \qquad (7.10)$$

for any real numbers $(x_1, \ldots, x_n)$, $(y_1, \ldots, y_n)$;

$$\left| \int_0^1 u(t)v(t)dt \right| \leq \sqrt{\int_0^1 u^2(t)dt} \sqrt{\int_0^1 v^2(t)dt} \qquad (7.11)$$

$$\left[\iiint_R u(x,y,z)v(x,y,z)dxdydz \right]^2 \leq \iiint_R u^2(x,y,z)dxdydz \cdot \iiint_R v^2(x,y,z)dxdydz \qquad (7.12)$$

The reader who is not impressed with the neatness of this proof should try writing out a proof of Eqs. (7.10), (7.11), (7.12) directly.

The norm has the following properties

$$||c\mathbf{v}|| = |c| \, ||\mathbf{v}|| \qquad (7.13)$$

$$||\mathbf{u} + \mathbf{v}|| \leq ||\mathbf{u}|| + ||\mathbf{v}|| \qquad (7.14)$$

The first follows from the fact that $||c\mathbf{v}||^2 = c\mathbf{v} \cdot c\mathbf{v} = c^2 ||\mathbf{v}||^2$ and the second from the fact that $||\mathbf{u} + \mathbf{v}||^2 = (\mathbf{u} + \mathbf{v}) \cdot (\mathbf{u} + \mathbf{v}) = \mathbf{u} \cdot (\mathbf{u} + \mathbf{v}) + \mathbf{v} \cdot (\mathbf{u} + \mathbf{v}) \leq ||\mathbf{u}|| \, ||\mathbf{u} + \mathbf{v}|| + ||\mathbf{v}|| \, ||\mathbf{u} + \mathbf{v}|| = (||\mathbf{u}|| + ||\mathbf{v}||) \, ||\mathbf{u} + \mathbf{v}||$. Now divide by $||\mathbf{u} + \mathbf{v}||$ to obtain (7.14).

Replacing $\mathbf{v}$ in Eq. (7.14) by $-\mathbf{v}$ we get

$$||\mathbf{u} - \mathbf{v}|| \leq ||\mathbf{u}|| + ||\mathbf{v}|| \qquad (7.15)$$

By letting $\mathbf{w} = \mathbf{u} - \mathbf{v}$ in Eq. (7.15) we see that $||\mathbf{w}|| \leq ||\mathbf{w} + \mathbf{v}|| + ||\mathbf{v}||$. It follows that

$$||\mathbf{w} + \mathbf{v}|| \geq \Big| \, ||\mathbf{w}|| - ||\mathbf{v}|| \, \Big| \qquad (7.16)$$

The distance between two vectors $\mathbf{u}$ and $\mathbf{v}$ is defined as $||\mathbf{u} - \mathbf{v}||$. We then have the triangle inequality (Cf. Fig. 1)

$$||\mathbf{u} - \mathbf{v}|| \leq ||\mathbf{u} - \mathbf{w}|| + ||\mathbf{w} - \mathbf{v}|| \qquad (7.17)$$

as follows from taking $\mathbf{u} - \mathbf{w}$ for $\mathbf{u}$ and $\mathbf{w} - \mathbf{v}$ for $\mathbf{v}$ in Eq. (7.14).

Since

$$|\mathbf{u} \cdot \mathbf{v} - \mathbf{u} \cdot \mathbf{w}| = |\mathbf{u} \cdot (\mathbf{v} - \mathbf{w})| \leq ||\mathbf{u}|| \, ||\mathbf{v} - \mathbf{w}||$$

we see that the inner product is a continuous function of its arguments; that is, if $\mathbf{v} \to \mathbf{w}$ (this means $||\mathbf{v} - \mathbf{w}|| \to 0$) then $\mathbf{u} \cdot \mathbf{v} \to \mathbf{u} \cdot \mathbf{w}$.

Similarly, from Eq. (7.16) it follows that the norm is also a continuous function of its argument; that is, if $\mathbf{v} \to \mathbf{w}$ (i.e., $||\mathbf{v} - \mathbf{w}|| \to 0$), then $||\mathbf{v}|| \to ||\mathbf{w}||$. Thus we have a definition of "distance" which satisfies all the usual properties. We can then proceed with many of the usual techniques of analysis in these abstract vector spaces.

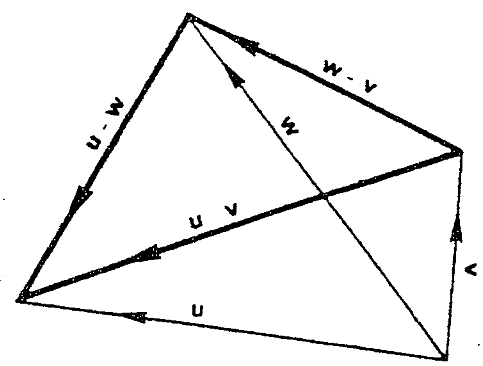


Fig.1. Triangle Inequality.

CHAPTER EIGHT

RECIPROCAL BASES; COVARIANT COMPONENTS; MIXED COMPONENTS

Suppose we have a vector space with an inner product and $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is a basis. The *reciprocal basis* is a set of vectors $\{\mathbf{e}^1, \mathbf{e}^2, \ldots, \mathbf{e}^n\}$ such that

$$\mathbf{e}^i \cdot \mathbf{e}_j = \delta^i{}_j$$

Example: In three-space let $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$, be three non-coplanar vectors, then the reciprocal basis is given by

$$\mathbf{e}^1 = \frac{\mathbf{e}_2 \times \mathbf{e}_3}{\lambda}, \quad \mathbf{e}^2 = \frac{\mathbf{e}_3 \times \mathbf{e}_1}{\lambda}, \quad \mathbf{e}^3 = \frac{\mathbf{e}_1 \times \mathbf{e}_2}{\lambda}$$

where

$$\lambda = \mathbf{e}_1 \times \mathbf{e}_2 \cdot \mathbf{e}_3$$

Let $\mathbf{e}^i \cdot \mathbf{e}^j = g^{ij}$ and $\mathbf{e}_i \cdot \mathbf{e}_j = g_{ij}$. Then, since $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is a basis, $\mathbf{e}^i = c^{ij}\mathbf{e}_j$, where c^{ij} are scalars. Dotting $\mathbf{e}^k$ into each side we get $g^{ik} = c^{ik}$ so that

$$\mathbf{e}^i = g^{ij}\mathbf{e}_j \tag{8.1}$$

Dotting $\mathbf{e}_k$ into each side we get $\delta^i{}_k = g^{ij}g_{jk}$, so that the matrices (g_{ij}) and (g^{ij}) are inverses. Since we have denoted the matrix (g_{ij}) by g, it follows that $(g^{ij}) = g^{-1}$.

We have assumed that there existed a reciprocal set. To prove that this is indeed the case define (g^{ij}) as the matrix inverse* to (g_{ij}) and define the $\mathbf{e}^i$ by Eq. (8.1). Then it follows that $\mathbf{e}^i \cdot \mathbf{e}_j = g^{ik}\mathbf{e}_k \cdot \mathbf{e}_j = g^{ik}g_{kj} = \delta^i{}_j$ so that we have our reciprocal set. A reciprocal basis is also called a dual basis.

To solve for the $\mathbf{e}_i$ in terms of the $\mathbf{e}^i$ we multiply each side of Eq. (8.1) by g_{ki} and sum on i. We get

$$g_{ki}\mathbf{e}^i = g_{ki}g^{ij}\mathbf{e}_j = \delta^j{}_k\mathbf{e}_j = \mathbf{e}_k$$

That is

$$\mathbf{e}_i = g_{ij}\mathbf{e}^j \tag{8.2}$$

[Compare the structure of Eq. (8.1) and (8.2).]

Hence the $\mathbf{e}^i$ also form a basis. Any vector $\mathbf{v}$ can also be expanded in terms of them

$$\mathbf{v} = v_i\mathbf{e}^i$$

The v_i are called the *covariant* components of the vector with respect to the given basis.

How are the components v_i related to the components v^i? Since the same vector is involved,

$$\mathbf{v} = v_i\mathbf{e}^i = v^i\mathbf{e}_i = v^i g_{ij}\mathbf{e}^j$$

Hence

$$v_i = g_{ij}v^j \tag{8.3}$$

This is the formula for "lowering indices of vectors." We shall see more of this shortly. Note that $v_i = \mathbf{v}\cdot\mathbf{e}_i$ and $v^i = \mathbf{v}\cdot\mathbf{e}^i$.

If instead of the basis $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$, we had started with the basis $\{\bar{\mathbf{e}}_1, \bar{\mathbf{e}}_2, \ldots, \bar{\mathbf{e}}_n\}$, then the dual basis would be $\{\bar{\mathbf{e}}^1, \ldots \bar{\mathbf{e}}^n\}$ where

$$\bar{\mathbf{e}}^i = \bar{g}^{ij}\bar{\mathbf{e}}_j \tag{8.4}$$

From Eq. (7.7) $\bar{g} = A^T g A$

Hence

$$\bar{g}^{-1} = A^{-1}g^{-1}(A^T)^{-1} = Bg^{-1}B^T$$

*We leave it as an exercise to prove that g is non-singular so that g^{-1} exists.

Therefore

$$\bar{g}^{ij} = b^i{}_k b^j{}_r g^{kr} = \frac{\partial \bar{x}^i}{\partial x^k} \frac{\partial \bar{x}^j}{\partial x^r} g^{kr} \tag{8.5}$$

The g^{ij} are therefore called contravariant metric tensor components. Now from Eq. (8.5) we get

$$\bar{\mathbf{e}}^i = \bar{g}^{ij}\bar{\mathbf{e}}_j = b^i{}_k b^j{}_r g^{kr}\bar{\mathbf{e}}_j = b^i{}_k b^j{}_r g^{kr} a^s{}_j \mathbf{e}_s$$

$$= b^i{}_k b^j{}_r g^{kr} a^s{}_j g_{st}\mathbf{e}^t = b^i{}_k \delta^s{}_r g^{kr} g_{st}\mathbf{e}^t = b^i{}_k g^{ks} g_{st}\mathbf{e}^t$$

$$= b^i{}_k \delta^k{}_t \mathbf{e}^t = b^i{}_j \mathbf{e}^j$$

The change in the dual basis is thus described by

$$\bar{\mathbf{e}}^i = b^i{}_j \mathbf{e}^j \qquad \text{or} \qquad \bar{\mathbf{e}}^* = B\mathbf{e}^* \tag{8.6}$$

If we multiply each side by $a^k{}_i$ we find also that

$$\mathbf{e}^i = a^i{}_j \bar{\mathbf{e}}^j \tag{8.7}$$

Note that this is somewhat the reverse of the relation relating the basis vectors $\mathbf{e}_i$ and $\bar{\mathbf{e}}_i$ Eqs. [(5.1), (5.3)] except that the matrices are not transposed. The upper and lower indices, however, provide a foolproof notation.

If $\mathbf{v} = \bar{v}_i \bar{\mathbf{e}}^i = v_i \mathbf{e}^i$, how are the $\bar{v}_i$ related to the v_i?

$$\mathbf{v} = \bar{v}_i \bar{\mathbf{e}}^i = v_i \mathbf{e}^i = v_i a^i{}_j \bar{\mathbf{e}}^j$$

i.e.

$$\bar{v}_j = a^i{}_j v_i \qquad \text{or} \qquad \bar{v}_* = A^T v_* \tag{8.8}$$

Since Eq. (8.8) is analogous to the original change of basis [Eq. (5.1)], the v_i are called the *covariant* components of the vector $\mathbf{v}$.

Similarly if we consider a tensor $\mathbf{T}$, we can write it as linear combinations of polyads formed by using either the $\mathbf{e}_i$ or the $\mathbf{e}^i$ or combinations.

For example, say

$$\mathbf{T} = t^{ijk}\mathbf{e}_i\mathbf{e}_j\mathbf{e}_k = t^{ij}{}_k \mathbf{e}_i\mathbf{e}_j\mathbf{e}^k = t^i{}_{jk}\mathbf{e}_i\mathbf{e}^j\mathbf{e}^k = t_{ijk}\mathbf{e}^i\mathbf{e}^j\mathbf{e}^k \tag{8.9}$$

It's all the same tensor. How are the components related when $\mathbf{T}$ is referred partially to the base vectors $\mathbf{e}_i$ and partially to the reciprocal base vectors $\mathbf{e}^i$?

Note that here we are *not* changing basis vectors; we have *one* basis $\{\mathbf{e}_i\}$ and its dual basis $\{\mathbf{e}^i\}$ and open products made up of assortments from each set. The notation of upper and lower indices $t^{\cdots}_{\cdots}$ is not entirely foolproof, since, for example, the tensor $t^i{}_j\mathbf{e}_i\mathbf{e}^j$ is different from the tensor $t^i{}_j\mathbf{e}^j\mathbf{e}_i$ in general. A notation such as

$$\mathbf{T} = t^{i\cdot}_{\cdot j}\,\mathbf{e}_i\mathbf{e}^j = t^{\cdot i}_{j\cdot}\;\mathbf{e}^j\mathbf{e}_i$$

can be used; but for simplicity we shall use simply $t^i{}_j$ with the understanding that we mean

$$\mathbf{T} = t^i{}_j\mathbf{e}_i\mathbf{e}^j$$

More generally we use

$$\mathbf{T} = t^{i_1 \dots i_r}{}_{j_1 \dots j_s}\,\mathbf{e}_{i_1}\dots\mathbf{e}_{i_r}\mathbf{e}^{j_1}\dots\mathbf{e}^{j_s}$$

with the understanding that in the polyads we write all the base vectors first and then all the dual vectors as shown. If we wish to alter the order of some basis and reciprocal basis vectors we shall have to extend the notation on the components $t^{\cdots}_{\cdots}$. Now

$$\mathbf{T} = t^{ijk}\mathbf{e}_i\mathbf{e}_j\mathbf{e}_k = t^{ij}{}_r\mathbf{e}_i\mathbf{e}_j\mathbf{e}^r = t^{ij}{}_r\mathbf{e}_i\mathbf{e}_j g^{rk}\mathbf{e}_k$$

Hence

$$t^{ijk} = t^{ij}{}_r g^{rk} \tag{8.10}$$

Similarly

$$\mathbf{T} = t^{ij}{}_k\mathbf{e}_i\mathbf{e}_j\mathbf{e}^k = t^i{}_{rk}\mathbf{e}_i\mathbf{e}^r\mathbf{e}^k = t^i{}_{rk}\mathbf{e}_i g^{jr}\mathbf{e}_j\mathbf{e}^k$$

Hence

$$t^{ij}{}_k = t^i{}_{rk}g^{rj} \tag{8.11}$$

Similarly

$$\mathbf{T} = t^i{}_{jk}\mathbf{e}_i\mathbf{e}^j\mathbf{e}^k = t_{rjk}\mathbf{e}^r\mathbf{e}^j\mathbf{e}^k = t_{rjk}g^{ri}\mathbf{e}_i\mathbf{e}^j\mathbf{e}^k$$

so that

$$t^i{}_{jk} = t_{rjk}g^{ri} \tag{8.12}$$

Similarly

$$t_{ijk} = g_{ir}t^r{}_{jk} \tag{8.13}$$

$$t^k{}_{ij} = g_{ir}t^{kr}{}_j \tag{8.14}$$

$$t^{ij}{}_k = g_{kr}t^{ijr} \tag{8.15}$$

Equations (8.10)–(8.15) are generalizations of the formula Eq. (8.3) above. From these we hope it is clear how indices may be raised or lowered. In words, the g^{ij} and the g_{ij} can be used to raise or lower indices; one index is a dummy to be summed out against a mate in the opposite position on the $t^{:::}_{:::}$; the other index provides the new free index, thus restoring the total number of free indices to its initial value, but replacing an upper by a lower, or vice versa. With the understanding that these $t^{:::}_{:::}$ are the scalar multipliers of base vectors $\mathbf{e}_i$ or reciprocal base vectors $\mathbf{e}^i$ with the matching index in reverse (lower or upper) position, and that all the $\mathbf{e}_i$'s come first, followed by the $\mathbf{e}^i$'s, we see that the tensor itself is *unchanged* by these operations on the components. For example, the tensor $t^{ij}\mathbf{e}_i\mathbf{e}_j$ is the same as the tensor $t^i{}_j\mathbf{e}_i\mathbf{e}^j$, since

$$t^i{}_j\mathbf{e}_i\mathbf{e}^j = g_{kj}t^{ik}\mathbf{e}_i\mathbf{e}^j = t^{ik}\mathbf{e}_i g_{kj}\mathbf{e}^j = t^{ik}\mathbf{e}_i\mathbf{e}_k$$

The components $t^{i_1 i_2 \ldots i_r}_{j_1 j_2 \ldots j_s}$ are called *mixed components* of a tensor of order $(r + s)$; r-times contravariant and s-times covariant; or briefly: $\binom{r}{s}$ tensor components. The tensor itself is

$$t^{i_1 i_2 \ldots i_r}_{j_1 j_2 \ldots j_s} \; \mathbf{e}_{i_1}\mathbf{e}_{i_2}\ldots\mathbf{e}_{i_r} \; \mathbf{e}^{j_1}\mathbf{e}^{j_2}\ldots\mathbf{e}^{j_s}$$

Now how do the mixed components change when the basis is changed? Take for example the $\binom{3}{2}$ tensor

$$\mathsf{T} = \bar{t}^{ijk}{}_{\mu\nu}\;\bar{\mathbf{e}}_i\bar{\mathbf{e}}_j\bar{\mathbf{e}}_k\bar{\mathbf{e}}^\mu\bar{\mathbf{e}}^\nu = t^{ijk}{}_{\mu\nu}\;\mathbf{e}_i\mathbf{e}_j\mathbf{e}_k\mathbf{e}^\mu\mathbf{e}^\nu$$

$$= t^{ijk}{}_{\mu\nu}\; b^r{}_i\bar{\mathbf{e}}_r b^s{}_j\bar{\mathbf{e}}_s b^t{}_k\bar{\mathbf{e}}_t\; a^\mu{}_\alpha\,\bar{\mathbf{e}}^\alpha\; a^\nu{}_\beta\,\bar{\mathbf{e}}^\beta$$

Hence

$$\bar{t}^{rst}{}_{\alpha\beta} = b^r{}_i\, b^s{}_j b^t{}_k a^\mu{}_\alpha\, a^\nu{}_\beta t^{ijk}{}_{\mu\nu} \tag{8.16}$$

or in the alternate notation

$$\bar{t}^{rst}{}_{\alpha\beta} = \frac{\partial\bar{x}^r}{\partial x^i}\frac{\partial\bar{x}^s}{\partial x^j}\frac{\partial\bar{x}^t}{\partial x^k}\frac{\partial x^\mu}{\partial\bar{x}^\alpha}\frac{\partial x^\nu}{\partial\bar{x}^\beta}\; t^{ijk}{}_{\mu\nu} \tag{8.17}$$

In words: the covariant (lower) indices each transform according to the covariant law:

$$\bar{v}_i = a^j{}_i v_j = \frac{\partial x^j}{\partial\bar{x}^i}\, v_j$$

and the contravariant (upper) indices each transform according to the contravariant law

$$\bar{v}^i = b^i{}_j v^j = \frac{\partial \bar{x}^i}{\partial x^j} v^j$$

Study the structure of Eqs. (8.16) and (8.17). The pattern is straightforward and easy to remember. The importance of these transformation laws stems from the fact that in practice the base vectors and the duals are often dropped and the components themselves regarded as the tensor. The reason for transforming them, as in Eqs. (8.16) and (8.17), is then not so clear, but we simply use these formulas.

If the two basis systems $\{\mathbf{e}_i\}$ and $\{\bar{\mathbf{e}}_i\}$ and their reciprocal systems $\{\mathbf{e}^i\}$, $\{\bar{\mathbf{e}}^i\}$ are all known, then the numbers a^i_j, b^i_j entering into the transformation formulas

$$\bar{\mathbf{e}}_i = a^j_i \mathbf{e}_j, \quad \mathbf{e}_i = b^j_i \bar{\mathbf{e}}_j, \quad \bar{\mathbf{e}}^i = b^i_j \mathbf{e}^j, \quad \mathbf{e}^i = a^i_j \bar{\mathbf{e}}^j$$

can be evaluated from the following formulas

$$a^j_i = \bar{\mathbf{e}}_i \cdot \mathbf{e}^j, \quad b^j_i = \mathbf{e}_i \cdot \bar{\mathbf{e}}^j$$

which are easily verified from the preceding ones.

In particular, if both basis systems are orthonormal, then $a^j_i = \bar{\mathbf{e}}_i \cdot \mathbf{e}^j = \bar{\mathbf{e}}^i \cdot \mathbf{e}_j = b^i_j$; so that $A^{-1} = B = A^T$ and the matrix is orthogonal. Also a^j_i is simply the cosine of the angle between $\bar{\mathbf{e}}_i$ and $\mathbf{e}_j$ in this case. Similarly b^i_j is the cosine of the angle between these same two vectors when the bases are orthonormal.

CHAPTER NINE

TENSORS AS LINEAR FUNCTIONS

With the inner product we can interpret tensors as linear functions of vectors. A *linear function* of a vector $\mathbf{v}$ is a function satisfying $\mathcal{F}(c\mathbf{u} + d\mathbf{v}) = c\,\mathcal{F}(\mathbf{u}) + d\,\mathcal{F}(\mathbf{v})$ for all vectors $\mathbf{u}, \mathbf{v}$ and all scalars c, d. If $\mathcal{F}$ has real values then it is called a linear *functional*. For example a fixed vector $\mathbf{v}_0$ can be regarded as a linear functional by the formula

$$\mathcal{F}_{\mathbf{v}_0}(\mathbf{v}) = \mathbf{v}_0 \cdot \mathbf{v} \tag{9.1}$$

Thus with each vector $\mathbf{v}_0$ we can associate a linear functional defined by Eq. (9.1), since $\mathbf{v}_0 \cdot (c\mathbf{u} + d\mathbf{v}) = c\,\mathbf{v}_0 \cdot \mathbf{u} + d\mathbf{v}_0 \cdot \mathbf{v} = c\mathcal{F}_{\mathbf{v}_0}(\mathbf{u}) + d\,\mathcal{F}_{\mathbf{v}_0}(\mathbf{v})$.

It is not hard to show that there is a one to one correspondence between the original vectors $\mathbf{v}$, and the linear functionals on them, $\mathcal{F}_{\mathbf{v}}$. If $\mathbf{v}$ is given, define $\mathcal{F}_{\mathbf{v}}$ by

$$\mathcal{F}_{\mathbf{v}}(\ ) = \mathbf{v} \cdot (\ )$$

Conversely, if $\mathcal{F}$ is given, define $\mathbf{v}$ by

$$\mathbf{v} = \mathcal{F}(\mathbf{e}_i)\mathbf{e}^i$$

This correspondence preserves the algebraic operations in the sense that $F_{\mathbf{v}+\mathbf{u}} = F_{\mathbf{v}} + F_{\mathbf{u}}$; $F_{c\mathbf{v}} = cF_{\mathbf{v}}$

Remarks. In certain circumstances we do not wish to assume the existence of an inner product. It is nevertheless possible to carry through an analysis similar to that developed here. We introduce the *space of linear functionals* (this is a vector space of dimension n called the *dual* of the given space). A *dual basis* is then introduced in the space of linear functionals. These are functionals $f^i(\mathbf{v})$ such that $f^i(\mathbf{e}_j) = \delta^i_j$. These then play the role of our reciprocal basis $\mathbf{e}^i$, except that the two kinds of vectors are kept separate. Under these circumstances it is necessary to distinguish between contravariant (the original) *vectors* and covariant *vectors* (linear functionals).* This situation will not concern us here.

If a linear function is vector-valued then it is a linear mapping as discussed in Chapter 6. For example a dyadic $\mathbf{T} = t^{ij}\mathbf{e}_i\mathbf{e}_j$ can be regarded as a linear mapping $\mathcal{T}(\mathbf{v})$ by the formula

$$\mathcal{T}(\mathbf{v}) = \mathbf{T}\cdot\mathbf{v} = t^{ij}\mathbf{e}_i\mathbf{e}_j\cdot\mathbf{v} = t^{ij}\mathbf{e}_i(\mathbf{e}_j\cdot\mathbf{v}) \tag{9.2}$$

which is a vector. [We can also form $\mathbf{v}\cdot\mathbf{T}$ in a similar way. In general $\mathbf{v}\cdot\mathbf{T} \neq \mathbf{T}\cdot\mathbf{v}$ unless (t^{ij}) is a symmetric matrix. If we take $\mathbf{v} = v^i\mathbf{e}_i$ then from Eq. (9.2)

$$\mathbf{w} = \mathcal{T}(\mathbf{v}) = t^{ij}\mathbf{e}_i\mathbf{e}_j\cdot v^r\mathbf{e}_r = t^{ij}g_{jr}v^r\mathbf{e}_i = w^i\mathbf{e}_i \tag{9.3}$$

thus

$$w^i = t^{ir}g_{rj}v^j = t^i_j\, v^j$$

The matrix representing the mapping $\mathcal{T}$ is (t^i_j) .

In Eq. (9.2) we computed the vector $\mathbf{T}\cdot\mathbf{v}$ in terms of the original basis $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$.

If we compute it with respect to a different basis $\{\bar{\mathbf{e}}_1, \bar{\mathbf{e}}_2, \ldots, \bar{\mathbf{e}}_n\}$ we get

$$\bar{t}^{ij}\bar{\mathbf{e}}_i\bar{\mathbf{e}}_j\cdot(\bar{v}^k\bar{\mathbf{e}}_k) = \bar{t}^{ij}\bar{\mathbf{e}}_i\bar{g}_{jr}\bar{\mathbf{e}}^r\cdot(\bar{v}^k\bar{\mathbf{e}}_k) = \bar{t}^i_r\bar{\mathbf{e}}_i\,\delta^r_k\bar{v}^k = \bar{t}^i_j\bar{\mathbf{e}}_i\bar{v}^j$$

$$= (b^i_k\, a^r_j\, t^k_r)(a^\alpha_i\,\mathbf{e}_\alpha)(b^j_\beta v^\beta) = \delta^\alpha_k\,\delta^r_\beta\, t^k_r\,\mathbf{e}_\alpha\, v^\beta$$

$$= \delta^\alpha_k\, t^k_r\mathbf{e}_\alpha v^r = t^k_r\mathbf{e}_k v^r = t^{ij}g_{jr}\mathbf{e}_i v^r$$

---

*Similarly a tensor $t^{i_1 i_2 \ldots i_r}_{j_1 j_2 \ldots j_s}\, \mathbf{e}_{i_1}\mathbf{e}_{i_1}\ldots\mathbf{e}_{i_r}\, f^{j_1}f^{j_2}\ldots f^{j_s}$ is then itself r-times contravariant and s-times covariant.

which is the same as Eq. (9.3). Thus the mapping $\mathbf{w} = \mathbf{T}\cdot\mathbf{v}$ is invariant. This might also be seen directly from the fact that the new matrix $(\bar{t}^i_j)$ will be given by $\bar{t}^i_j = b^i_r\, a^k_j\, t^r_k$. That is, letting $T = (t^i_j)$, $\bar{T} = (\bar{t}^i_j)$, $\bar{T} = BTB^{-1}$. But by Eq. (6.5) this represents the same transformation with respect to the new basis.

By dotting another vector into $\mathbf{T}\cdot\mathbf{v}$ we would obtain a scalar $\mathbf{u}\cdot\mathbf{T}\cdot\mathbf{v} = \mathbf{T}\cdot\mathbf{v}\cdot\mathbf{u}$. Thus a second order tensor can be regarded as a linear functional of two vectors: $f(\mathbf{u}, \mathbf{v})$, where

$$f(\mathbf{u}, c\mathbf{v} + d\mathbf{w}) = cf(\mathbf{u}, \mathbf{v}) + df(\mathbf{u}, \mathbf{w}),$$
$$f(c\mathbf{u} + d\mathbf{v}, \mathbf{w}) = cf(\mathbf{u}, \mathbf{w}) + df(\mathbf{v}, \mathbf{w})$$

Thus if $\mathbf{T} = t^{ij}\mathbf{e}_i\mathbf{e}_j$, $\mathbf{u} = u^i\mathbf{e}_i$, $\mathbf{v} = v^i\mathbf{e}_i$ then

$$\mathbf{u}\cdot\mathbf{T}\cdot\mathbf{v} = u^r(\mathbf{e}_r\cdot t^{ij}\mathbf{e}_i)(\mathbf{e}_j\cdot v^k\mathbf{e}_k) = u^r g_{ri} t^{ij} g_{jk} v^k \qquad (9.4)$$
$$= u_i t^{ij} v_j$$

In particular $\mathbf{e}^i\cdot\mathbf{T}\cdot\mathbf{e}^j = t^{ij}$, $\mathbf{e}_i\cdot\mathbf{T}\cdot\mathbf{e}_j = t_{ij}$ and $\mathbf{e}^i.\mathbf{T}\cdot\mathbf{e}_j = t^i_j$.

The metric tensor $\mathbf{I} = g_{ij}\mathbf{e}^i\mathbf{e}^j$ has the property that

$$\mathbf{v}\cdot\mathbf{I}\cdot\mathbf{v} = v^r\mathbf{e}_r\cdot g_{ij}\mathbf{e}^i\mathbf{e}^j\cdot v^k\mathbf{e}_k = v^r\delta^i_r g_{ij}\delta^j_k v^k = v^i g_{ij} v^j = v^i v_i = \|\mathbf{v}\|^2,$$

as in Eq. (7.6). Note the $\mathbf{I}$ has the alternate forms

$$\mathbf{I} = g_{ij}\mathbf{e}^i\mathbf{e}^j = \mathbf{e}_j\mathbf{e}^j = \mathbf{e}_j\mathbf{e}_k g^{kj}$$

The reason for using the letter $\mathbf{I}$ for the metric tensor is that, when regarded as a linear mapping, it is the identity:

$$\mathbf{I}\cdot\mathbf{v} = \mathbf{e}_j\mathbf{e}^j\cdot v^k\mathbf{e}_k = \mathbf{e}_j\delta^j_k v^k = \mathbf{e}_j v^j = \mathbf{v}$$

Given an arbitrary tensor of order m we can, by dotting vectors $\mathbf{v}_{(1)}, \mathbf{v}_{(2)}, \ldots, \mathbf{v}_{(m)}$ with it, get a scalar. Thus a tensor can be regarded as a linear functional of m vectors.

Given a fourth order tensor $\mathbf{T} = t^{ij}_{kr}\,\mathbf{e}_i\mathbf{e}_j\mathbf{e}^k\mathbf{e}^r$ we can form a linear mapping of second order tensors into second order tensors by the formula

$$\mathbf{W} = \mathbf{T}{:}\mathbf{U} = t^{ij}_{kr}\,\mathbf{e}_i\mathbf{e}_j\mathbf{e}^k\mathbf{e}^r : (u^{\alpha\beta}\,\mathbf{e}_\alpha\mathbf{e}_\beta) = t^{ij}_{kr}\,u^{\alpha\beta}\mathbf{e}_i\mathbf{e}_j\mathbf{e}^k(\mathbf{e}^r\cdot\mathbf{e}_\alpha)\cdot\mathbf{e}_\beta$$
$$= t^{ij}_{kr}\,u^{\alpha\beta}\mathbf{e}_i\mathbf{e}_j(\mathbf{e}^r\cdot\mathbf{e}_\alpha)(\mathbf{e}^k\cdot\mathbf{e}_\beta) = t^{ij}_{kr}\,u^{\alpha\beta}\mathbf{e}_i\mathbf{e}_j\delta^r_\alpha\delta^k_\beta = t^{ij}_{kr}\,u^{rk}\mathbf{e}_i\mathbf{e}_j$$

It is not hard to show that this linear mapping of second-order tensors ($\mathbf{U}$) into second-order tensors ($\mathbf{W}$) is independent of the basis chosen. Conversely we can show that any linear mapping of second-order tensors into second-order tensors is (can be represented by) a fourth-order tensor. This is done by considering the effect of the mapping on a basis for the vector space of second-order tensors, i.e. all those tensors of the form $\mathbf{e}_i\mathbf{e}_j$.

Similarly an r^{th} order tensor can be regarded as a linear mapping of s^{th} order tensors into $(r - s)^{\text{th}}$ order tensors and conversely. More generally by putting in fewer dots other linear mappings can be formed. In particular, if only one dot is used, then a $(r + s - 2)^{\text{nd}}$ order tensor is formed. (Compare Chapter 13 below.)

CHAPTER TEN

CONTRACTION

With an inner product in our space we can form dot products in a variety of ways. For example, given the dyad **uv** we might form the scalar $\mathbf{u}\cdot\mathbf{v}$. This is a linear function of each of the arguments **u, v**. It is called the *contraction* of the tensor. Similarly, given the dyadic, e.g., $\mathsf{T} = t^{ij}\mathbf{e}_i\mathbf{e}_j$, we can form the scalar

$$t^{ij}\mathbf{e}_i\cdot\mathbf{e}_j = t^{ij}g_{ij}$$

In terms of a different basis, $\mathsf{T} = \bar{t}^{ij}\bar{\mathbf{e}}_i\bar{\mathbf{e}}_j$ and the contraction would be given by $\bar{t}^{ij}\bar{g}_{ij}$. But from Eq. (7.7), $\bar{g}_{ij} = a_i^r a_j^s g_{rs}$ and, from Eq. (5.7) $\bar{t}^{ij} = b_r^i b_s^j \; t^{rs}$, so that the scalar obtained by contracting in the new coordinate system is

$$\bar{t}^{ij}\bar{g}_{ij} = b_r^i b_s^j t^{rs} \; a_i^k a_j^t g_{kt} = \delta_r^k \delta_s^t t^{rs} \; g_{kt} = t^{ks} \; g_{ks}$$

which is the same as in the original coordinate system.

Thus the contraction of a second-order tensor is not dependent on the coordinate system. The result is the same scalar, regardless of the coordinate system used.

Exercise: Show that the triadic $t^{ijk}\mathbf{e}_i\mathbf{e}_j\mathbf{e}_k$ can be contracted to a vector by the formula $t^{ijk}\mathbf{e}_i \cdot \mathbf{e}_j\mathbf{e}_k$ and that the result is independent of the basis used; (i.e., show that $t^{ijk}g_{ij}\mathbf{e}_k = \bar{t}^{ijk}\bar{g}_{ij}\bar{\mathbf{e}}_k$).

Generalize this result for any order tensor. Notice that the result depends on where the dot is placed.

Let us look again at the contraction of the tensor $\mathbf{T} = t^{ij}\mathbf{e}_i\mathbf{e}_j$. Before contracting let us replace $\mathbf{e}_i$ by its expression in terms of the dual basis $\{\mathbf{e}^1, \mathbf{e}^2, \ldots, \mathbf{e}^n\}$, $\mathbf{e}_i = g_{ij}\mathbf{e}^j$. Thus $\mathbf{T}$ can be written in the mixed form $\mathbf{T} = t^{ij}g_{ik}\mathbf{e}^k\mathbf{e}_j = t^j_k\mathbf{e}^k\mathbf{e}_j$. Now contracting gives

$$t^j_k\mathbf{e}^k \cdot \mathbf{e}_j = t^j_k\delta^k_j = t^j_j$$

Thus, as far as the components are concerned, contraction consists in equating an upper index to a lower index and summing. This is due to the fact that an upper and lower index pair correspond to terms $t^{---i--}_{---j--}\mathbf{e}_i\mathbf{e}^j$, and when we form the contraction we get

$$t^{---i--}_{---j--}\mathbf{e}_i \cdot \mathbf{e}^j = t^{---i--}_{---j--}\delta^j_i = t^{---i--}_{---i--}$$

Consider again the mapping of vectors into vectors by a second-order tensor, given by the formula

$$\mathbf{w} = \mathbf{T} \cdot \mathbf{v}$$

The vector $\mathbf{T} \cdot \mathbf{v}$ can be regarded as the tensor product $\mathbf{Tv}$ (third order) followed by contraction $\mathbf{T} \cdot \mathbf{v}$ to produce a first-order tensor, or vector.

Similarly the linear mapping of second-order tensors into second order tensors by a fourth-order tensor, as discussed above

$$\mathbf{W} = \mathbf{T} : \mathbf{U}$$

can be regarded as the tensor product $\mathbf{TU}$ (sixth order) followed by two contractions.

CHAPTER ELEVEN

EXAMPLE: MOMENT OF MOMENTUM OF A RIGID BODY

Consider a collection of mass points m_i, rigidly connected together and rotating about an axis through a point Q in the body with angular velocity ω as in Figure 2. Let $\mathbf{R}_i$ denote the vector

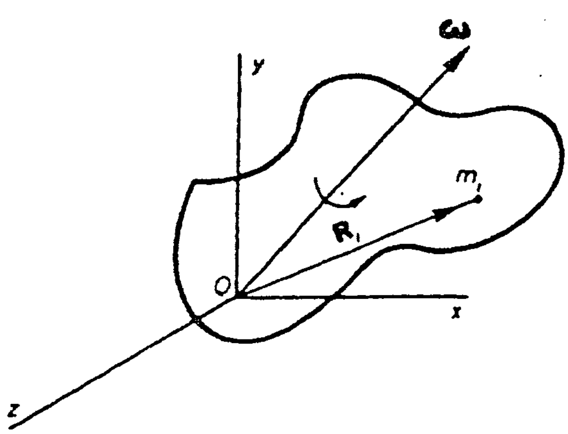


Fig. 2. Mass Rotating About an Axis.

from Q to m_i. The velocity of m_i with respect to Q is then given by $\dot{\mathbf{R}}_i = \boldsymbol{\omega} \times \mathbf{R}_i$. The moment of momentum about Q of the body is given by

$$\mathbf{H}_Q = \sum_i \mathbf{R}_i \times (m_i \dot{\mathbf{R}}_i) = \sum_i m_i \mathbf{R}_i \times (\boldsymbol{\omega} \times \mathbf{R}_i) = \sum_i m_i [(\mathbf{R}_i \cdot \mathbf{R}_i)\boldsymbol{\omega} - (\mathbf{R}_i \cdot \boldsymbol{\omega})\mathbf{R}_i]$$

$$= \sum_i m_i (r_i^2 \boldsymbol{\omega} - \mathbf{R}_i \mathbf{R}_i \cdot \boldsymbol{\omega}) = \sum_i m_i (r_i^2 \mathbf{I} - \mathbf{R}_i \mathbf{R}_i) \cdot \boldsymbol{\omega}$$

where we have used the formula $\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = (\mathbf{A} \cdot \mathbf{C})\mathbf{B} - (\mathbf{A} \cdot \mathbf{B})\mathbf{C}$ and $\mathbf{I}$ denotes the identity tensor $\mathbf{I} \cdot \boldsymbol{\omega} = \boldsymbol{\omega}$. Thus we see that

$$\mathbf{H}_Q = \mathcal{J} \cdot \boldsymbol{\omega}$$

where $\mathcal{J}$ is the inertia tensor, given in rectangular coordinates by

$$\mathcal{J} = \begin{pmatrix} \Sigma m_i (y_i^2 + z_i^2)\mathbf{ii} - \Sigma m_i x_i y_i \mathbf{ij} - \Sigma m_i x_i z_i \mathbf{ik} \\ -\Sigma m_i y_i x_i \mathbf{ji} + \Sigma m_i (x_i^2 + z_i^2)\mathbf{jj} - \Sigma m_i y_i z_i \mathbf{jk} \\ -\Sigma m_i z_i x_i \mathbf{ki} - \Sigma m_i z_i y_i \mathbf{kj} + \Sigma m_i (x_i^2 + y_i^2)\mathbf{kk} \end{pmatrix}$$

Note that $y_i^2 + z_i^2$ is the square of the distance of m_i from the x axis. Letting $I_{xx} = \Sigma m_i (y_i^2 + z_i^2)$, $I_{xy} = \Sigma m_i x_i y_i$ and so forth we can write

$$\mathcal{J} = \begin{pmatrix} I_{xx}\mathbf{ii} - I_{xy}\mathbf{ij} - I_{xz}\mathbf{ik} \\ -I_{yx}\mathbf{ji} + I_{yy}\mathbf{jj} - I_{yz}\mathbf{jk} \\ -I_{zx}\mathbf{ki} - I_{zy}\mathbf{kj} + I_{zz}\mathbf{kk} \end{pmatrix}$$

In the case of a continuous distribution of mass we can write

$$I_{xx} = \int (y^2 + z^2)\,dm, \quad I_{xy} = \int xy\,dm$$

and so forth. Here I_{xx} is called the *moment of inertia* about the x axis and I_{xy} is called a *product of inertia*. To bring the notation into line with what we have done previously let $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ be replaced by $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ and let $I^{11} = I_{xx}$, $I^{12} = -I_{xy}$ and so forth. Then

$$\mathcal{J} = I^{ij}\mathbf{e}_i\mathbf{e}_j$$

The formulas for the change of the moments and products of inertia (when a different coordinate system is used) then follow directly from the general rules

$$\bar{I}^{ij} = b_r^i b_s^j I^{rs}$$

where $\mathbf{e}_r = b_r^s \bar{\mathbf{e}}_s$, so that

$$b_r^s = \mathbf{e}_r \cdot \bar{\mathbf{e}}^s$$

If $\{\bar{\mathbf{e}}_i\}$ is a normal orthogonal system as the $\{\mathbf{e}_i\}$ is in this case, then b_r^s is simply the cosine of the angle between $\mathbf{e}_r$ and $\bar{\mathbf{e}}^s$. Clearly $I^{ij} = I^{ji}$ and the tensor is called *symmetric*.

The advantage of taking base vectors fixed in the body is that the components I^{ij} are constant in time, although the base vectors move.

The theorem that a symmetric matrix can be reduced to diagonal form by an orthogonal transformation, $O\mathcal{I}O^{-1}$, means here that there exists a set of orthogonal *principal axes*; i.e., axes such that the products of inertia are all zero.

Similarly it can be shown that the kinetic energy of rotation is given by $\mathrm{KE} = (1/2)\boldsymbol{\omega} \cdot \mathcal{I} \cdot \boldsymbol{\omega}$.

The situation depicted is not as special as it might seem, since the general instantaneous motion of a rigid body can be described by the motion of one point fixed in the body, plus a rotation about an axis through that point.

CHAPTER TWELVE

STATE OF STRESS AT A POINT

Consider a stressed continuum as in Figure 3(a). Let P be a specific point in that medium. Take a small circular disk with center at P (as in Figure 3(b)) with normal $\mathbf{n}$, radius r and thickness $\epsilon \to 0$. The force exerted by the rest of the medium

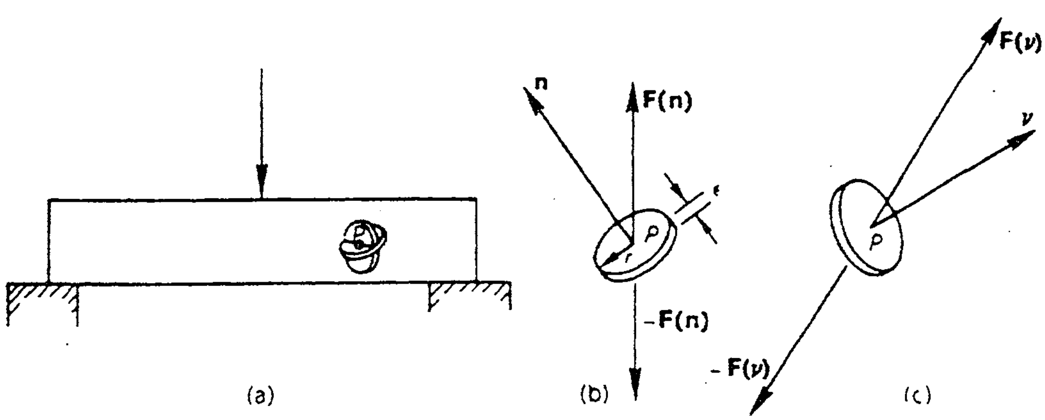


Fig. 3. Stress at a Point

on the positive face is denoted by F(n). Since the mass $\pi r^2 \rho\epsilon$ tends to zero as $\epsilon \to 0$, the force F(n) is balanced by a force on the negative face -F(n). The average stress on the positive face is $\frac{F(n)}{\pi r^2}$ and the limit of this as $r \to 0$ is $\tau(n)$, the stress at P on a plane whose normal is n. Of course if a direction different from n were chosen at the outset, such as ν in Figure 3(c), the stress at the point would be $\tau(\nu) = \lim_{r\to 0} \frac{F(\nu)}{\pi r^2}$. Here $\tau(\nu)$ represents the stress at the point on a plane whose normal is ν. The component of $\tau(n)$ in the direction of n is called a *tension* (compression for $\tau(n) < 0$) or *normal* stress. The component in the plane is a *shear* stress.

Thus the state of stress at the point P is described in terms of two vectors: for each given unit vector n there corresponds the stress vector $\tau(n)$ acting on the plane normal to n. In other words $\tau(n)$ is a function of n. Let us find a standard form for this relationship.

Consider the forces acting on a small cube of side ϵ with origin at P. Figure 4 illustrates the notation we shall be using and shows the resolution of forces acting on the face $x = \epsilon$, where $\epsilon \to 0$. Similar notation applies to the other faces. Note that τ_{xy} represents the y component of the stress acting on the $x = \epsilon$ face; the corresponding stress on the $x = 0$ face is approximately $-\tau_{xy}$ for ϵ very small. The total stress on the $x = \epsilon$

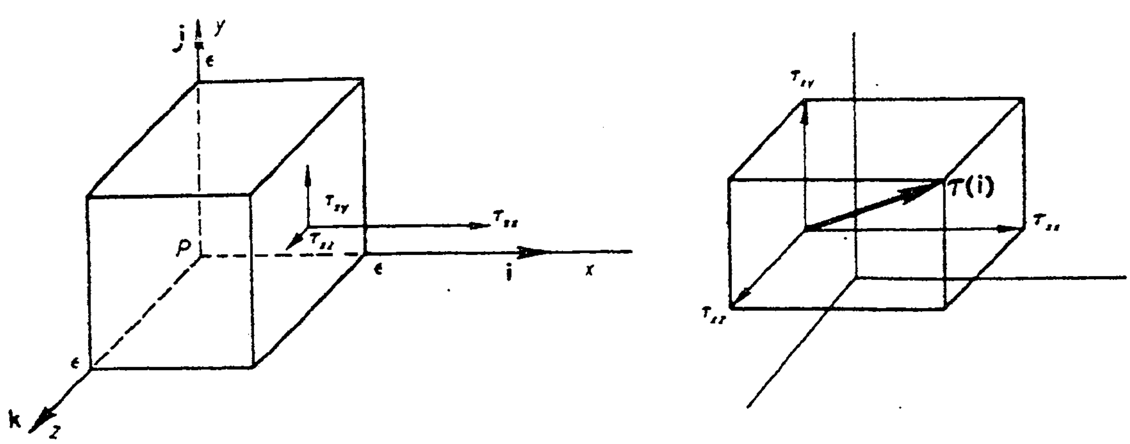


Fig. 4. Stress on Plane $x = 0-$

face is then $\boldsymbol{\tau}(\mathbf{i}) = \tau_{xx}\mathbf{i} + \tau_{xy}\mathbf{j} + \tau_{xz}\mathbf{k}$, where the τ_{ij} are evaluated on the $x = \epsilon$ plane. The stress on the $x = 0$ face is $\boldsymbol{\tau}(-\mathbf{i}) = -[\tau_{xx}\mathbf{i} + \tau_{xy}\mathbf{j} + \tau_{xz}\mathbf{k}]$, where the τ_{ij} are evaluated on the $x = 0$ face. As $\epsilon \to 0$ the corresponding values of τ_{ij} tend to equality. Similarly we may obtain the stresses on the remaining four faces of the cube. By computing the torque about the y axis we find that, in the absence of body torques, the condition for equilibrium is $\tau_{xz} = \tau_{zx}$. Similarly for the other shearing stresses.

Now let $\mathbf{n}$ be a unit normal direction and consider the free body of the tetrahedron in Fig. 5.

From Newton's law, $\mathbf{F} = m\mathbf{a}$, we get:

$$A\boldsymbol{\tau}(\mathbf{n}) - \{A_x(\tau_{xx}\mathbf{i} + \tau_{xy}\mathbf{j} + \tau_{xz}\mathbf{k}) + A_y(\tau_{yx}\mathbf{i} + \tau_{yy}\mathbf{j} + \tau_{yz}\mathbf{k}) + A_z(\tau_{zx}\mathbf{i} + \tau_{zy}\mathbf{j} + \tau_{zz}\mathbf{k})\} + \mathbf{f}\rho V = \rho V\mathbf{a}$$

where A_x, A_y, A_z are the areas of the tetrahedron on the $x = 0$, $y = 0$, $z = 0$ planes; A is the area of the slant face whose normal is $\mathbf{n}$; $\mathbf{f}$ is the force per unit mass due to any body forces that may be acting; ρ is the mass density, and V is the volume of the tetrahedron. Since $A_x = A\,\mathbf{n}\cdot\mathbf{i}$, $A_y = A\mathbf{n}\cdot\mathbf{j}$, $A_z = A\,\mathbf{n}\cdot\mathbf{k}$, we get

$$\boldsymbol{\tau}(\mathbf{n}) = \mathbf{n}\cdot[\mathbf{i}(\tau_{xx}\mathbf{i} + \tau_{xy}\mathbf{j} + \tau_{xz}\mathbf{k}) + \mathbf{j}(\tau_{yx}\mathbf{i} + \tau_{yy}\mathbf{j} + \tau_{yz}\mathbf{k}) + \mathbf{k}(\tau_{zx}\mathbf{i} + \tau_{zy}\mathbf{j} + \tau_{zz}\mathbf{k})] - \frac{\mathbf{f}\rho h}{3} + \frac{\rho h}{3}\mathbf{a}$$

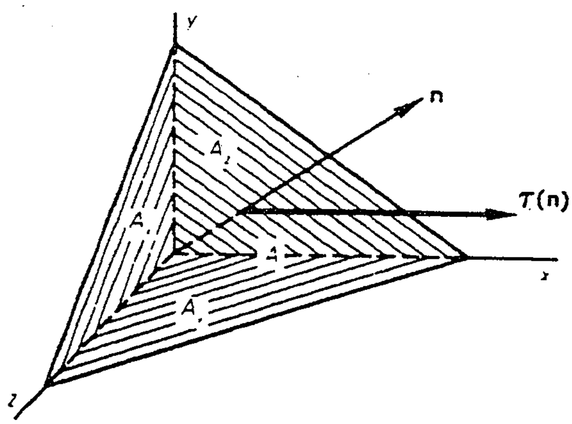


Fig. 5. Tetrahedral Free Body.

where we have used the fact that $V = \frac{hA}{3}$, where h is the distance from the origin P to the slant face. Letting $h \to 0$ we see that we can write

$$\boldsymbol{\tau}(\mathbf{n}) = \mathbf{n} \cdot \boldsymbol{\phi}$$

where $\boldsymbol{\phi}$ is the stress tensor

$$\boldsymbol{\phi} = \begin{pmatrix} \tau_{xx}\mathbf{ii} + \tau_{xy}\mathbf{ij} + \tau_{xz}\mathbf{ik} \\ \tau_{yx}\mathbf{ji} + \tau_{yy}\mathbf{jj} + \tau_{yz}\mathbf{jk} \\ \tau_{zx}\mathbf{ki} + \tau_{zy}\mathbf{kj} + \tau_{zz}\mathbf{kk} \end{pmatrix}$$

Thus the stress tensor has the property that a unit vector $\mathbf{n}$ dotted into it results in the vector $\boldsymbol{\tau}(\mathbf{n})$, which is the stress vector on the plane whose normal is $\mathbf{n}$. We shall assume that there are no body torques and the medium is in equilibrium so that $\boldsymbol{\phi}$ is symmetric and $\boldsymbol{\tau}(\mathbf{n}) = \mathbf{n} \cdot \boldsymbol{\phi} = \boldsymbol{\phi} \cdot \mathbf{n}$,

To express these facts in terms of our general notation, let $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ be the orthonormal basis $\mathbf{i}, \mathbf{j}, \mathbf{k}$. Then, using our standard notation, we have

$$\boldsymbol{\phi} = \tau^{ij}\mathbf{e}_i\mathbf{e}_j = \tau_{ij}\mathbf{e}^i\mathbf{e}^j$$

where $\tau^{ij} = \tau_{ij}$, and, since our basis is orthonormal, $\mathbf{e}^i = \mathbf{e}_i$.

The theorem that a symmetric matrix can be diagonalized by an orthogonal transformation means here that there exists an orthogonal set of *principal directions*. The planes perpendicular to these have no shear on them; only normal stresses. In terms of this coordinate system, the stress components $\bar{\tau}_{ij} = \sigma_i\delta_{ij}$ (no sum on i). From the invariants given in Eq. (6.8) we see that

$$\tau_{xx} + \tau_{yy} + \tau_{zz} = \sigma_1 + \sigma_2 + \sigma_3$$

$$\tau_{xx}\tau_{yy} - \tau_{xy}^2 + \tau_{xx}\tau_{zz} - \tau_{xz}^2 + \tau_{yy}\tau_{zz} - \tau_{yz}^2 = \sigma_2\sigma_3 + \sigma_1\sigma_3 + \sigma_1\sigma_2$$

$$\begin{vmatrix} \tau_{xx} & \tau_{xy} & \tau_{xz} \\ \tau_{yx} & \tau_{yy} & \tau_{yz} \\ \tau_{zx} & \tau_{zy} & \tau_{zz} \end{vmatrix} = \sigma_1\sigma_2\sigma_3 .$$

The stresses $\sigma_1, \sigma_2, \sigma_3$ are called *principal stresses*.

For a fluid at rest there are no shearing stresses, so that

$$\boldsymbol{\tau}(\mathbf{i}) = -p_1\mathbf{i}, \boldsymbol{\tau}(\mathbf{j}) = -p_2\mathbf{j}, \boldsymbol{\tau}(\mathbf{k}) = -p_3\mathbf{k}.$$

Also

$$\tau(\mathbf{n}) = [\tau(\mathbf{i})\mathbf{i} + \tau(\mathbf{j})\mathbf{j} + \tau(\mathbf{k})\mathbf{k}] \cdot \mathbf{n} = -(p_1\mathbf{ii} + p_2\mathbf{jj} + p_3\mathbf{kk}) \cdot \mathbf{n} = -p\mathbf{n}$$

Dotting $\mathbf{i}$ on each side we get $p_1 \mathbf{i} \cdot \mathbf{n} = p\mathbf{i} \cdot \mathbf{n}$. Similarly it follows that $p_1 = p_2 = p_3 = p$. Therefore, for a fluid at rest, the stress tensor is of the form $-p\mathbf{e}_i\mathbf{e}^i$ and the pressure is the same in all directions. Such a stress situation in general is called *hydrostatic pressure.*

If we transform to a new basis $\bar{\mathbf{e}}_i = a_i^j \mathbf{e}_j$, the components of the stress tensor are given by

$$\bar{\tau}_{ij} = a_i^r a_j^s \tau_{rs} \qquad (12.1)$$

$$\bar{\tau}^{ij} = b_r^i b_s^j \tau^{rs}$$

Thus, although $\tau_{ij} = \tau^{ij}$ (because our original basis was orthonormal) we see, in general, that $\bar{\tau}_{ij} \neq \bar{\tau}^{ij}$, unless the matrix a_j^i is orthogonal, so that $B = A^{-1} = A^T$, and the new basis is also orthonormal. The relation $\tau(\mathbf{n}) = \mathbf{n} \cdot \boldsymbol{\phi}$ is not dependent on the basis, however, and so remains true in any coordinate system. From this we can see how to interpret the $\bar{\tau}^{ij}$. Take $\bar{n}^i = \dfrac{\bar{e}^i}{||\bar{e}^i||}$; (no sum on i). Then

$$\tau(\bar{n}^i) = \frac{\bar{e}^i}{||\bar{e}^i||} \cdot \bar{\tau}^{jk}\bar{e}_j\bar{e}_k = \frac{\bar{\tau}^{jk}}{||\bar{e}^i||}\delta_j^i\bar{e}_k = \frac{\bar{\tau}^{ik}}{||\bar{e}^i||}e_k .$$

Now dot $\bar{e}^j$ into each side to obtain

$$\bar{\tau}^{ij} = ||\bar{e}^i||\tau(\bar{n}^i)\cdot\bar{e}^j = ||\bar{e}^i||\;||\bar{e}^j||\tau(\bar{n}^i)\cdot\bar{n}^j .$$

Since $||\bar{e}^i|| = \sqrt{\bar{g}^{ii}}$ we see that $\bar{\tau}^{ij}$ is equal to $\sqrt{\bar{g}^{ii}}\sqrt{\bar{g}^{jj}}$ times the physical component, in the $\bar{n}^j$ direction, of the stress on the plane whose normal is in the $\bar{n}^i$ direction.

Similarly

$$\bar{\tau}_j^i = \bar{\mathbf{e}}_j \cdot \tau(\bar{\mathbf{e}}^i), \text{ and } \bar{\tau}_{ij} = \tau(\bar{\mathbf{e}}_i) \cdot \bar{\mathbf{e}}_j$$

In general if $\boldsymbol{\nu}$ and $\mathbf{n}$ are unit vectors then $\boldsymbol{\nu} \cdot \boldsymbol{\phi} \cdot \mathbf{n}$ is the component in the $\boldsymbol{\nu}$ direction of the stress on the plane normal to $\mathbf{n}$.

Exercise. Show that if the covariant tensor components are symmetric ($\tau_{rs} = \tau_{sr}$) in one coordinate system, then they are symmetric in every coordinate system. [Hint: $\bar{\tau}_{rs} = a^i_r a^j_s \tau_{ij}$; or in matrix notation $\bar{\tau} = A^T \tau A$. Hence $\bar{\tau}^T = A^T \tau^T A = \bar{\tau}$.] On the other hand if the mixed components are symmetric ($\tau^r_s = \tau^s_r$) in one coordinate system, it does not follow that these are symmetric in another coordinate system. [Hint: $\bar{\tau}^r_s = a^j_s b^r_i \tau^i_j$ or, in matrix notation, $\bar{\tau} = B\tau A$. Thus $\bar{\tau}^T = A^T \tau B^T$, which does equal $\bar{\tau}$ if $A^T = A^{-1}$, (i.e., for orthogonal transformations) but not in general.]

CHAPTER THIRTEEN

SOME OTHER POINT-TENSORS

Before going on to the study of *fields* we will consider some examples of *point-tensors* which arise in the study of crystals. The reader interested in more such examples will find them in J. P. Nye, *Physical Properties of Crystals* (Oxford University Press, 1957) or W. P. Mason, *Physical Acoustics and the Properties of Solids* (D. Van Nostrand Co., Inc., 1958). The following examples are from Appendix C of Nye's book, and we follow his notation, with some modifications.

1. *Heat Capacity* C. This is a scalar relating two scalars: $\Delta S = (C/T)\Delta T$.

   A small increase in *temperature* ΔT produces a change in *entropy* ΔS. Here ΔS is a linear function of ΔT. The relation between the two is given in terms of the scalar C/T. Here C is a zero-order tensor and the orders in the stated relationship are [0= 0-0].

2. *Pyroelectricity* (p_i). This is a vector relating a vector to a scalar.

$$\Delta P_i = p_i \Delta T$$

A small change in *temperature* ΔT produces a change in the *polarization* vector ΔP_i, which is a linear function of ΔT. The relation is given in terms of the *pyroelectric vector* $\{p_i\}$. The orders involved in the stated relationship are [1= 1-0].

The same pyroelectric vector is also involved in the electrocaloric effect given by $\Delta S = p_i E^i$, where $\{E^i\}$ is the *electric field vector*. Here the order of the tensors involved in the equation is [0= 1-1].

3. *Permeability* (μ_{ij}). Here we have (as in the stress tensor and the inertia tensor treated in Chapters 11 and 12) a second-rank tensor relating two vectors [1=2-1]. The equations $B_i = \mu_{ij} H^j$ give the *magnetic induction* or *flux density* vector **B** as a linear function of the *magnetic field intensity* or *field strength* vector **H**. The *permeability* tensor (μ_{ij}) defines the linear function; in our general notation the relation would be written $\mathbf{B} = \boldsymbol{\mu} \cdot \mathbf{H}$ and the function described as $\boldsymbol{\mu} \cdot (\ )$. As another example: In the present notation the inertia tensor as discussed in Chapter 11 relates the angular momentum vector **H** to the angular velocity vector $\boldsymbol{\omega}$ through the equation $H_i = I_{ij} \omega^j$. Similarly the stress tensor of Chapter 12 relates the stress vector on a plane, to the normal vector of the plane $\tau_i = \tau_{ij} n^j$. Similarly the strain tensor (ϵ_{ij}) relates the strain vector of a direction, to a unit vector in that direction. In each of these cases we have a second-order tensor relating two vectors, with the scheme [1=2-1].

4. The *thermal expansion* tensor (α_{ij}) is a second-order tensor. In the equation $\epsilon_{ij} = \alpha_{ij} \Delta T$ it relates the second-order strain tensor to the scalar change in temperature ΔT. The orders of the tensors involved in the equation take the form [2=2-0]. The same tensor α_{ij} is involved in the *piezocaloric* effect $\Delta S = \alpha_{ij} \tau^{ij}$ relating the change in entropy ΔS to the stress tensor (τ^{ij}). Here the orders are of the form [0=2-2]. The *thermal pressure* tensor f^{ij} is similarly involved in two relationships one giving thermal pressure $\tau^{ij} = -f^{ij} \Delta T$ [2=2-0], and the other giving the heat of deformation $\Delta S = f^{ij} \epsilon_{ij}$ [0=2-2].

5. In the direct *piezoelectric effect* $P_i = d_{ijk} \tau^{jk}$ we have the form [1=3-2] in which the polarization vector (P_i) is expressed as a linear function of the stress tensor (τ^{ij}) through the third-order tensor of *piezoelectric moduli* (d_{ijk}). The

same third-order tensor (d_{ijk}) is involved in the converse piezoelectric effect $\epsilon_{jk} = d_{ijk} E^i$ relating the strain tensor (ϵ_{jk}) to the electric field strength vector (E^i). Here the orders of the tensors take the form [2=3-1].

6. As an example of a fourth-order tensor we cite the elastic stiffness tensor c_{ij}^{rs} which relates the stress tensor to the strain tensor:

$$\tau_{ij} = c_{ij}^{rs} \epsilon_{rs}$$

Here the form is [2=4-2]. There are $3^4 = 81$ quantities c_{ij}^{rs}, but many of these can be shown to be equal. One of the objectives of the study of crystal structures is to use the symmetries of the crystal structure to determine the minimum number of essential constants involved in these tensor relationships, and then find ways of measuring them experimentally.

CHAPTER FOURTEEN

CROSS PRODUCTS

We first introduce the *alternating symbol*

$$\epsilon^{i_1 i_2 \ldots i_n} = \epsilon_{i_1 i_2 \ldots i_n}$$

which has the value +1 if $(i_1, i_2, \ldots i_n)$ is an even permutuation of the integers $(1, 2, \ldots, n)$; it has the value -1 if $(i_1, i_2, \ldots, i_n)$ is an odd permutuation of the integers $(1, 2, \ldots, n)$. In all other cases it has the value 0.

Let $F = (f^i{}_j)$ be an $n \times n$ matrix. Then by the definition* of a determinant we have

$$\det(F) = \epsilon^{i_1 i_2 \ldots i_n} f^1_{i_1} f^2_{i_2} \ldots f^n_{i_n} \tag{14.1}$$

where $\det(F)$ denotes the determinant of F.

*The classical definition of a determinant is that it is a sum of products, each product consisting of n elements of the matrix such that each row and each column is represented; the sign associated with each such term is (+) or (-) as, after the elements are arranged in order of increasing rows, the resulting column indices form an even or odd permutation of $(1, 2, \ldots, n)$.

This can be generalized easily to

$$\epsilon^{i_1 i_2 \ldots i_n} f^{j_1}_{i_1} f^{j_2}_{i_2} \ldots f^{j_n}_{i_n} = \epsilon^{j_1 \ldots j_n} \det(F) \qquad (14.2)$$

The student may, on occasion, have tried to prove the fundamental relation that the determinant of a product is the product of the determinants; i.e., if F and G are $n \times n$ matrices then

$$\det(FG) = \det(F) \cdot \det(G) \qquad (14.3)$$

Those who have tried it know what a complicated business it can be. As an illustration of the beauty of the notation just introduced, we shall prove it here as follows.

$$\det(F) \cdot \det G = [\epsilon^{i_1 i_2 \ldots i_n} f^1_{i_1} f^2_{i_2} \ldots f^n_{i_n}] \cdot \det(G)$$

$$= f^1_{i_1} f^2_{i_2} \ldots f^n_{i_n} \cdot \epsilon^{i_1 i_2 \ldots i_n} \det(G)$$

$$= f^1_{i_1} f^2_{i_2} \ldots f^n_{i_n} g^{i_1}_{j_1} g^{i_2}_{j_2} \ldots g^{i_n}_{j_n} \epsilon^{j_1 j_2 \ldots j_n}$$

$$= \epsilon^{j_1 j_2 \ldots j_n} h^1_{j_1} h^2_{j_2} \ldots h^n_{j_n} = \det H$$

where $h^i_j = f^i_s g^s_j$; i.e. $H = FG$. Q.E.D.

So far we have not defined a cross product for abstract vectors. We have assumed that the student knew about the cross product of two physical (3-dimensional) vectors in order to provide illustrative material in Chapter 8. Let us examine this case first.

The usual definition of the direction of $\mathbf{A} \times \mathbf{B}$ is that if $\mathbf{A}$ and $\mathbf{B}$ have the same origin and if the fingers of the right hand are curled from the tip of $\mathbf{A}$ to the tip of $\mathbf{B}$ then the thumb will point in the direction of $\mathbf{A} \times \mathbf{B}$. This definition is too anthropomorphic for use in an age when communication with creatures from outer space seems imminent. Let us try to define this without using hands. Since the cross product is additive and homogeneous it suffices to define it for the basis elements. For convenience let the usual orthonormal basis ($\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$) be denoted by ($\mathbf{i}_1$, $\mathbf{i}_2$, $\mathbf{i}_3$) = ($\mathbf{i}^1$, $\mathbf{i}^2$, $\mathbf{i}^3$). If we define the cross product by the formula

$$\mathbf{i}_\mu \times \mathbf{i}_\nu = \epsilon_{\mu\nu\sigma} \, \mathbf{i}^\sigma \qquad (14.4)$$

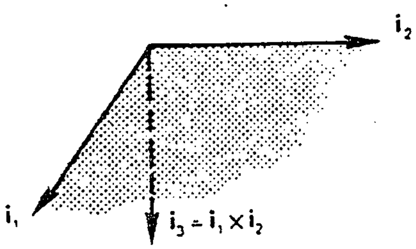
then we get the usual cross product if the coordinate system is right-handed. A colleague using the same formula but a left-handed coordinate system would get the oppositely directed vector. (Figure 6.)

Our next objective is to define the cross product in an *abstract* 3-dimensional vector space. Of course we cannot use our hands, and the above illustration shows us that we can anticipate difficulties, if we have left handed colleagues.

Let $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ be a basis, not necessarily orthogonal, and let $(\mathbf{e}^1, \mathbf{e}^2, \mathbf{e}^3)$ be the reciprocal basis. We want to find the formula for $\mathbf{e}_i \times \mathbf{e}_j$ in terms of objects already defined. Let $(\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3) = (\mathbf{i}^1, \mathbf{i}^2, \mathbf{i}^3)$ be an orthonormal basis such as $(\mathbf{i}, \mathbf{j}, \mathbf{k})$ mentioned above.* Since $(\mathbf{e}^1, \mathbf{e}^2, \mathbf{e}^3)$ is a basis we have

$$\mathbf{e}_i \times \mathbf{e}_j = c_k \mathbf{e}^k \tag{14.5}$$

and dotting $\mathbf{e}_k$ into each side we have $c_k = \mathbf{e}_i \times \mathbf{e}_j \cdot \mathbf{e}_k$. We



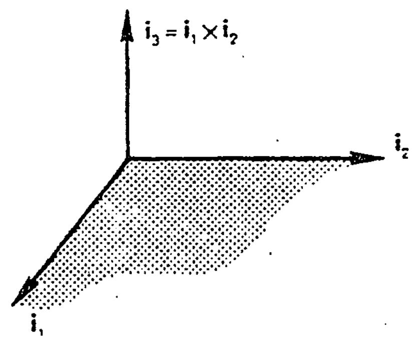


Fig. 6.

*The actual construction of an orthonormal basis can be accomplished by the "Gram-Schmidt Orthogonalization Procedure": take

$$\mathbf{v}_1 = \mathbf{e}_1 \qquad \mathbf{i}_1 = \mathbf{v}_1/\|\mathbf{v}_1\|$$

$$\mathbf{v}_2 = \mathbf{e}_2 - (\mathbf{e}_2 \cdot \mathbf{i}_1)\mathbf{i}_1 \qquad \mathbf{i}_2 = \mathbf{v}_2/\|\mathbf{v}_2\|$$

$$\mathbf{v}_3 = \mathbf{e}_3 - (\mathbf{e}_3 \cdot \mathbf{i}_1)\mathbf{i}_1 - (\mathbf{e}_3 \cdot \mathbf{i}_2)\mathbf{i}_2 \qquad \mathbf{i}_3 = \mathbf{v}_3/\|\mathbf{v}_3\|$$

are not finished yet because this formula for c_k involves the cross product. Let $\mathbf{e}_i = E_i^\alpha \mathbf{i}_\alpha$. Using Eq. (14.4) for the cross product of the orthonormal basis we get

$$c_k = \mathbf{e}_i \times \mathbf{e}_j \cdot \mathbf{e}_k = E_i^\alpha \mathbf{i}_\alpha \times E_j^\beta \mathbf{i}_\beta \cdot E_k^\gamma \mathbf{i}_\gamma = E_i^\alpha E_j^\beta E_k^\gamma \epsilon_{\alpha\beta\mu} \mathbf{i}^\mu \cdot \mathbf{i}_\gamma$$

$$= E_i^\alpha E_j^\beta E_k^\gamma \epsilon_{\alpha\beta\mu} \delta_\gamma^\mu = E_i^\alpha E_j^\beta E_k^\gamma \epsilon_{\alpha\beta\gamma} = \epsilon_{ijk} \det(E)$$

where E is the matrix (E_i^α). Therefore, from Eq. (14.5), we have $\mathbf{e}_i \times \mathbf{e}_j = \epsilon_{ijk} \det(E)\, \mathbf{e}^k$.

Now $g_{ij} = \mathbf{e}_i \cdot \mathbf{e}_j = E_i^\alpha \mathbf{i}_\alpha \cdot E_j^\beta \mathbf{i}_\beta = E_i^\alpha E_j^\beta \delta_{\alpha\beta} = E_i^\alpha E_j^\alpha$

Thus we have the matrix relation $g = E^T E$. Therefore $\det(g) = [\det(E)]^2$. Hence

$$\mathbf{e}_i \times \mathbf{e}_j = \pm \epsilon_{ijk}\, e^k \sqrt{\det(g)}$$

Now in an abstract vector space we do not have the preferred system $(\mathbf{i}, \mathbf{j}, \mathbf{k})$ to start with. We start instead with *some* chosen basis $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$. Then, in terms of this basis we define

$$\mathbf{e}_i \times \mathbf{e}_j = \epsilon_{ijk}\, \mathbf{e}^k \sqrt{\det(g)} \qquad (14.6)$$

If our neighbor has his own basis $(\bar{\mathbf{e}}_1, \bar{\mathbf{e}}_2, \bar{\mathbf{e}}_3)$ and uses the same recipe (14.6), he will define the cross product by

$$\bar{\mathbf{e}}_i \times \bar{\mathbf{e}}_j = \epsilon_{ijk}\, \bar{\mathbf{e}}^k \sqrt{\det(\bar{g})} \qquad (14.7)$$

Is this what we would get if we used our formula (14.6) and applied it to our neighbor's vectors in Eq. (14.7)? In terms of our system we then wish to verify Eq. (14.7). For Eq. (14.7) to be true we must have:

$$(a_i^r \mathbf{e}_r) \times (a_j^s \mathbf{e}_s) = \epsilon_{ijk}\, b_t^k\, \mathbf{e}^t \sqrt{\det(a_i^r a_j^s g_{rs})}$$

$$a_i^r a_j^s \epsilon_{rst}\, \mathbf{e}^t \sqrt{\det(g)} = \epsilon_{ijk}\, b_t^k \mathbf{e}^t \sqrt{\det(g)} \sqrt{[\det(A)]^2}$$

$$a_i^r a_j^s \epsilon_{rst} = \epsilon_{ijk}\, b_t^k \sqrt{[\det(A)]^2}$$

Multiplying each side by a_r^t, (and summing on t), we get

$$\epsilon_{ijr} \det(A) = \epsilon_{ijk} \delta^k_r \, |\det(A)|$$

i.e.

$$\det(A) = |\det(A)|$$

In other words our neighbor gets the same result as we do, provided the determinant of the transformation is positive; i.e., det $(A) > 0$. Otherwise the sense of the coordinate system is reversed and his formula for cross product does not lead to the same result as ours.

In an n-dimensional space with basis $(\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ we might define the cross product to be the $(n-2)^{\text{nd}}$ -order tensor

$$\mathbf{e}_{i_1} \times \mathbf{e}_{i_2} = \epsilon_{i_1 i_2 \ldots i_n} \mathbf{e}^{i_3} \mathbf{e}^{i_4} \ldots \mathbf{e}^{i_n} \sqrt{\det(g)} \qquad (14.8)$$

Our friend, using the same recipe, would write

$$\bar{\mathbf{e}}_{i_1} \times \bar{\mathbf{e}}_{i_2} = \epsilon_{i_1 i_2 \ldots i_n} \bar{\mathbf{e}}^{i_3} \ldots \bar{\mathbf{e}}^{i_n} \sqrt{\det(\bar{g})} \qquad (14.9)$$

Now we want to check his formula (14.9) using our formula (14.8). Formula (14.9) will check, provided we have the following sequence of equalities:

$$a^{j_1}_{i_1} \mathbf{e}_{j_1} \times a^{j_2}_{i_2} \mathbf{e}_{j_2} = \epsilon_{i_1 i_2 \ldots i_n} b^{i_3}_{j_3} \mathbf{e}^{j_3} b^{i_4}_{j_4} \mathbf{e}^{j_4} \ldots b^{i_n}_{j_n} \mathbf{e}^{j_n} |\det(A)| \sqrt{\det(g)}$$

$$\epsilon_{j_1 j_2 j_3 \ldots j_n} a^{j_1}_{i_1} a^{j_2}_{i_2} \mathbf{e}^{j_3} \mathbf{e}^{j_4} \ldots \mathbf{e}^{j_n}$$

$$= \epsilon_{i_1 i_2 \ldots i_n} b^{i_3}_{j_3} b^{i_4}_{j_4} \ldots b^{i_n}_{j_n} \mathbf{e}^{j_3} \mathbf{e}^{j_4} \ldots \mathbf{e}^{j_n} |\det(A)|$$

$$\epsilon_{j_1 j_2 \ldots j_n} a^{j_1}_{i_1} a^{j_2}_{i_2} = \epsilon_{i_1 i_2 \ldots i_n} b^{i_3}_{j_3} b^{i_4}_{j_4} \ldots b^{i_n}_{j_n} |\det(A)|$$

multiplying by $a^{j_3}_{r_3} a^{j_4}_{r_4} \ldots a^{j_n}_{r_n}$ (and summing) we get

$$\epsilon_{i_1 i_2 r_3 r_4 \ldots r_n} \det(A) = \epsilon_{i_1 i_2 \ldots i_n} \delta^{i_3}_{r_3} \delta^{i_4}_{r_4} \ldots \delta^{i_n}_{r_n} |\det(A)|$$

$$\det(A) = |\det(A)|$$

That is, our neighbor gets the same result provided his transformation from the preferred coordinate system (ours) has a positive determinant, i.e., $\det(A) > 0$.

The cross product can be defined more elegantly in terms of the "wedge product" [See G. Berman "The Wedge Product." American Mathematical Monthly, **68** (1961) pp. 112-119]. We shall not go into this here, nor into the "exterior algebra" [see H. K. Nickerson, D. C. Spencer and N. E. Steenrod, *Advanced Calculus* (D. Van Nostrand Co., Inc., Princeton, N. J., 1959)], which furnishes the natural generalization of these concepts.

CHAPTER FIFTEEN

CURVILINEAR COORDINATES

Consider the case of curvilinear coordinates $x^i = x^i(\xi^1, \xi^2, \xi^3)$, as in Fig. 7.

Here **R** is a vector from a fixed origin 0 to an arbitrary point

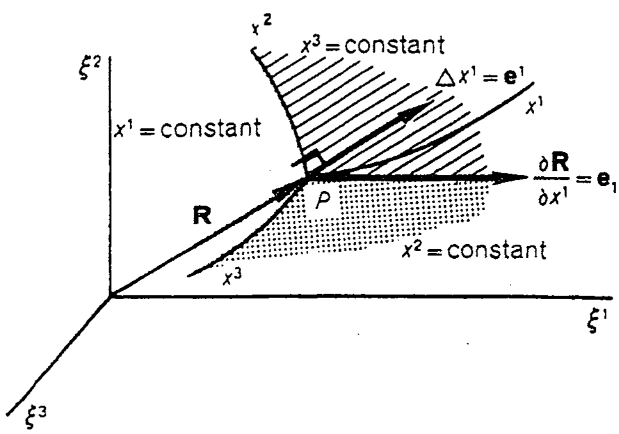


Fig. 7. Curvilinear Coordinate System.

P. Now $\dfrac{\partial \mathbf{R}}{\partial x^i}$ is tangent to the x^i coordinate curve. Also ∇x^i is normal to the surface: x^i = constant. Thus $\nabla x^i \cdot \dfrac{\partial \mathbf{R}}{\partial x^j} = 0$ if $i \neq j$. Furthermore $\nabla x^i \cdot \dfrac{\partial \mathbf{R}}{\partial x^i} = 1$, since $dx^i = \nabla x^i \cdot d\mathbf{R}$. [Another way to see all this is to let $\mathbf{R}(t)$ represent a path through P. Then $\nabla x^i \cdot \dot{\mathbf{R}} = \dfrac{\partial x^i}{\partial \xi^k} \dot{\xi}^k = \dot{x}^i$. Hence $\nabla x^i \cdot d\mathbf{R} = dx^i$.]* Thus

$$\nabla x^i \cdot \frac{\partial \mathbf{R}}{\partial x^j} = \frac{\partial x^i}{\partial x^j} = \delta^i_j$$

Therefore if we take as a basis, at the point P, $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3,\}$. where

$$\mathbf{e}_i = \frac{\partial \mathbf{R}}{\partial x^i}$$

then the reciprocal basis is given by $\mathbf{e}^i = \nabla x^i$. We shall use $\mathbf{e}_i = \dfrac{\partial \mathbf{R}}{\partial x^i}$ as our base vectors at the point P.

If $\mathbf{R}(t)$ is a path through P, then the square of the speed at the point P is given by

$$\left(\frac{ds}{dt}\right)^2 = v^2 = \dot{\mathbf{R}} \cdot \dot{\mathbf{R}} = \frac{\partial \mathbf{R}}{\partial x^i} \dot{x}^i \cdot \frac{\partial \mathbf{R}}{\partial x^j} \dot{x}^j = \dot{x}^i \dot{x}^j \mathbf{e}_i \cdot \mathbf{e}_j = g_{ij} \dot{x}^i \dot{x}^j$$

This also may be written

$$(ds)^2 = g_{ij}\, dx^i dx^j \tag{15.1}$$

Consider a new coordinate system $\bar{x}^i = \bar{x}^i(\xi^1, \xi^2, \xi^3)$. We shall consider all transformations to be one-to-one, so that we might also write $\bar{x}^i = \bar{x}^i(x^1, x^2, x^3)$ and $x^i = x^i(\bar{x}^1, \bar{x}^2, \bar{x}^3)$.

The point P is located by $\mathbf{R}(x^1, x^2, x^3) = \bar{\mathbf{R}}(\bar{x}^1, \bar{x}^2, \bar{x}^3)$. Again take $\bar{\mathbf{e}}_i = \dfrac{\partial \bar{\mathbf{R}}}{\partial \bar{x}^i}$, $\bar{\mathbf{e}}^i = \nabla \bar{x}^i$. Then, as in Chapter 2, $\bar{\mathbf{e}}_i = a^j_i \mathbf{e}_j$. Let us find the a^j_i in terms of the quantities present here.

$$\bar{\mathbf{e}}_i = \frac{\partial \bar{\mathbf{R}}}{\partial \bar{x}^i} = \frac{\partial \mathbf{R}}{\partial x^j} \frac{\partial x^j}{\partial \bar{x}^i} = \frac{\partial x^j}{\partial \bar{x}^i} \mathbf{e}_j \tag{15.2}$$

*Another derivation is as follows: $\dfrac{\partial x^i}{\partial x^j} = \dfrac{\partial x^i}{\partial \xi^k} \dfrac{\partial \xi^k}{\partial x^j} = \nabla x^i \cdot \dfrac{\partial \mathbf{R}}{\partial x^j}$. But $\dfrac{\partial x^i}{\partial x^j} = \delta^i_j$.

Thus $a_i^j = \frac{\partial x^j}{\partial \bar{x}^i}$ as in Eq. (5.9), although the x's of Chapter 5 are a very special case of the x's being used here. Similarly to Eq. (5.10) we have

$$b_j^i = \frac{\partial \bar{x}^i}{\partial x^j}$$

since the matrices $\frac{\partial x^i}{\partial \bar{x}^j}$ and $\frac{\partial \bar{x}^i}{\partial x^j}$ are inverses of each other

$$\frac{\partial \bar{x}^i}{\partial x^j} \frac{\partial x^j}{\partial \bar{x}^k} = \frac{\partial \bar{x}^i}{\partial \bar{x}^k} = \delta_k^i$$

Thus solving Eq. (15.2) for $\mathbf{e}_j$ we get

$$\mathbf{e}_j = \frac{\partial \bar{x}^i}{\partial x^j} \bar{\mathbf{e}}_i \tag{15.3}$$

Notice how Eq. (15.3) may be obtained from Eq. (15.2) by "high school algebra." Similarly

$$\bar{\mathbf{e}}^i = \frac{\partial \bar{x}^i}{\partial x^j} \mathbf{e}^j \tag{15.4}$$

since

$$\nabla \bar{x}^i = \frac{\partial \bar{x}^i}{\partial x^j} \nabla x^j$$

An alternate proof may be obtained from Eq. (15.2) as in Chapter 5. The laws of transformation of components of tensors now follow as in Eqs. (5.11) to (5.13) and so on, except that the coefficients

$$a_j^i = \frac{\partial x^i}{\partial \bar{x}^j}, \quad b_j^i = \frac{\partial \bar{x}^i}{\partial x^j}$$

are now variable as we move about in the field. If $\mathbf{R}(t)$ describes a curve, we can compute the velocity vector

$$\dot{\mathbf{R}} = \frac{\partial \mathbf{R}}{\partial x^i} \dot{x}^i = \dot{x}^i \mathbf{e}_i = v^i \mathbf{e}_i$$

Thus the velocity vector has contravariant components $\dot{x}^i$, in terms of the given basis.

The acceleration is slightly more complicated, since *as we move along the path, the base vectors also change.* Thus

$$a = \ddot{\mathbf{R}} = \dot{v}^i \mathbf{e}_i + v^i \dot{\mathbf{e}}_i = \ddot{x}^i \mathbf{e}_i + v^i \frac{\partial \mathbf{e}_i}{\partial x^j} \dot{x}^j = \ddot{x}^i \mathbf{e}_i + v^i v^j \frac{\partial \mathbf{e}_i}{\partial x^j}$$

The quantities $\dfrac{\partial \mathbf{e}_i}{\partial x^j} = \dfrac{\partial^2 \mathbf{R}}{\partial x^i \partial x^j}$ will be discussed in the next chapter.

Suppose next we have a scalar field $\phi(\mathbf{R})$. We can define the vector field $\nabla\phi$ by the equation

$$\nabla\phi = \frac{\partial \phi}{\partial x^i} \nabla x^i = \frac{\partial \phi}{\partial x^i} \mathbf{e}^i \tag{15.5}$$

The only trouble with this definition is that it is given in terms of a particular coordinate system. Suppose we used the same rule Eq. (15.5), but computed in a new coordinate system. We would get

$$\nabla\phi = \frac{\partial \phi}{\partial \bar{x}^i} \bar{\mathbf{e}}^i = \frac{\partial \phi}{\partial x^j} \frac{\partial x^j}{\partial \bar{x}^i} \bar{\mathbf{e}}^i = \frac{\partial \phi}{\partial x^i} \mathbf{e}^i$$

which is the same as before. Thus, even though Eq. (15.5) defines $\nabla\phi$ in terms of a particular coordinate system, the resulting vector field comes out the same, no matter which coordinate system is used. It would have been more elegant to define $\nabla\phi$ in a manner which does not depend on any coordinate system. This may be done as follows: Take any curve $\mathbf{R}(t)$ through P. Then, as we move along the curve, the scalar ϕ is a function of time; its derivative with respect to time is a linear function of the velocity $\dot{\mathbf{R}}$. Thus $\dot{\phi} = \mathbf{A} \cdot \dot{\mathbf{R}}$ for a suitable vector $\mathbf{A}$. This vector $\mathbf{A}$ is defined to be $\nabla\phi$. This definition does not depend on a coordinate system. In any coordinate system we obtain at once

$$\dot{\phi} = \frac{\partial \phi}{\partial x^i} \dot{x}^i = \frac{\partial \phi}{\partial x^i} \mathbf{e}^i \cdot \dot{x}^j \mathbf{e}_j = \frac{\partial \phi}{\partial x^i} \mathbf{e}^i \cdot \dot{\mathbf{R}}. \quad \text{Hence} \quad \nabla\phi = \frac{\partial \phi}{\partial x^i} \mathbf{e}^i$$

While this procedure may seem artificial, it has the advantage of being invariant. We may use the same method to now define the *gradient* of a vector field $\mathbf{u}$. Again $\dot{\mathbf{u}}$ is a linear (vector valued) function of the velocity $\dot{\mathbf{R}}$ so $\nabla\mathbf{u}$ is a second-order tensor $\mathbf{A}$ such that $\dot{\mathbf{u}} = \mathbf{A} \cdot \dot{\mathbf{R}}$. In any coordinate system we have

$$\dot{\mathbf{u}} = \frac{\partial \mathbf{u}}{\partial x^i} \dot{x}^i = \frac{\partial \mathbf{u}}{\partial x^i} \mathbf{e}^i \cdot \dot{x}^j \mathbf{e}_j = \frac{\partial \mathbf{u}}{\partial x^i} \mathbf{e}^i \cdot \dot{\mathbf{R}}$$

Hence

$$\nabla\mathbf{u} = \frac{\partial \mathbf{u}}{\partial x^i} \nabla x^i$$

a formula which is often used as the definition of $\nabla \mathbf{u}$. However, as a definition it involves a particular coordinate system. Thus we must verify that if we perform the same operation in a new coordinate system, we arrive at the same second-order tensor.

Written out in more detail we have

$$\nabla \mathbf{u} = \nabla(u^i \mathbf{e}_i) = \left(\frac{\partial u^i}{\partial x^j}\, \mathbf{e}_i + u^i \frac{\partial \mathbf{e}_i}{\partial x^j}\right) \nabla x^j$$

The quantities $\dfrac{\partial \mathbf{e}_i}{\partial x^j}$ have appeared again. They will be dealt with in the next chapter.

The *divergence* of a vector field $\mathbf{u}$ may be defined in terms of a particular coordinate system by the formula

$$\nabla \cdot \mathbf{u} = \frac{\partial \mathbf{u}}{\partial x^i} \cdot \nabla x^i \tag{15.6}$$

The same form computed in a new coordinate system would give

$$\nabla \cdot \mathbf{u} = \frac{\partial \mathbf{u}}{\partial \bar{x}^i} \cdot \nabla \bar{x}^i = \frac{\partial \mathbf{u}}{\partial x^j} \cdot \frac{\partial x^j}{\partial \bar{x}^i}\, \bar{\mathbf{e}}^i = \frac{\partial \mathbf{u}}{\partial x^j} \cdot \mathbf{e}^j$$

so that the same result is obtained.

If we write out Eq. (15.6) we get

$$\nabla \cdot \mathbf{u} = \frac{\partial (u^j \mathbf{e}_j)}{\partial x^i} \cdot \mathbf{e}^i = \frac{\partial u^j}{\partial x^i}\, \delta^i_j + u^j \frac{\partial \mathbf{e}_j}{\partial x^i} \cdot \mathbf{e}^i = \frac{\partial u^i}{\partial x^i} + u^j \frac{\partial \mathbf{e}_j}{\partial x^i} \cdot \mathbf{e}^i$$

Again the quantities $\dfrac{\partial \mathbf{e}_j}{\partial x^i}$ have appeared.

CHAPTER SIXTEEN

COVARIANT DERIVATIVES

As we move the point P, not only do the coordinates of the point, $\{x^i\}$, change, but so do the basis vectors $\{\mathbf{e}_i\}$ and also the reciprocal basis vectors $\{\mathbf{e}^i\}$. Thus $\dfrac{\partial \mathbf{e}_i}{\partial x^j} = \dfrac{\partial^2 \mathbf{R}}{\partial x^j \partial x^i}$ is a vector which we now expand in terms of the basis $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$

$$\frac{\partial \mathbf{e}_i}{\partial x^j} = \left\{ {k \atop ij} \right\} \mathbf{e}_k \tag{16.1}$$

The components $\left\{ {k \atop ij} \right\}$ are called the *Christoffel symbols of the second kind*. Note that all three of the quantities in Eq. (16.1), viz. $\dfrac{\partial \mathbf{e}_i}{\partial x^j}$, $\left\{ {k \atop ij} \right\}$, $\mathbf{e}_k$ depend on the particular coordinate system being used. Solving Eq. (16.1) for $\left\{ {k \atop ij} \right\}$ we get

$$\left\{ {k \atop ij} \right\} = \frac{\partial \mathbf{e}_i}{\partial x^j} \cdot \mathbf{e}^k = \frac{\partial^2 \mathbf{R}}{\partial x^j \partial x^i} \cdot \nabla x^k \tag{16.2}$$

Suppose we have a vector field $\mathbf{v} = \mathbf{v}(x^1, x^2, x^3) = v^i \mathbf{e}_i$. We now take at each point the partial derivative of the vector field with respect to x^j. The result is a new vector field:

$$\frac{\partial \mathbf{v}}{\partial x^j} = \frac{\partial}{\partial x^j}(v^i \mathbf{e}_i) = \frac{\partial v^i}{\partial x^j}\mathbf{e}_i + v^i \frac{\partial \mathbf{e}_i}{\partial x^j} = \left[\frac{\partial v^k}{\partial x^j} + v^i \begin{Bmatrix} k \\ ij \end{Bmatrix}\right]\mathbf{e}_k$$

Hence

$$\frac{\partial \mathbf{v}}{\partial x^j} = v^k{}_{,j}\mathbf{e}_k \tag{16.3}$$

where the contravariant components of this new field are given by

$$v^k{}_{,j} = \frac{\partial v^k}{\partial x^j} + \begin{Bmatrix} k \\ ij \end{Bmatrix} v^i \tag{16.4}$$

and are called the *covariant derivatives.*

Let us find the formula for $\frac{\partial \mathbf{e}^i}{\partial x^j}$. Since $\mathbf{e}^i \cdot \mathbf{e}_k = \delta^i_k$ we have

$$0 = \frac{\partial \mathbf{e}^i}{\partial x^j} \cdot \mathbf{e}_k + \mathbf{e}^i \cdot \frac{\partial \mathbf{e}_k}{\partial x^j} = \frac{\partial \mathbf{e}^i}{\partial x^j} \cdot \mathbf{e}_k + \mathbf{e}^i \cdot \begin{Bmatrix} r \\ kj \end{Bmatrix} \mathbf{e}_r$$

or

$$\frac{\partial \mathbf{e}^i}{\partial x^j} \cdot \mathbf{e}_k = -\begin{Bmatrix} r \\ kj \end{Bmatrix} \mathbf{e}^i \cdot \mathbf{e}_r = -\begin{Bmatrix} i \\ kj \end{Bmatrix} \tag{16.5}$$

Hence

$$\frac{\partial \mathbf{e}^i}{\partial x^j} = -\begin{Bmatrix} i \\ kj \end{Bmatrix} \mathbf{e}^k$$

Therefore we can compute $\frac{\partial \mathbf{v}}{\partial x^j}$ in the form

$$\frac{\partial \mathbf{v}}{\partial x^j} = \frac{\partial}{\partial x^j}(v_i \mathbf{e}^i) = \frac{\partial v_i}{\partial x^j}\mathbf{e}^i + v_i \frac{\partial \mathbf{e}^i}{\partial x^j} = \left[\frac{\partial v_k}{\partial x^j} - \begin{Bmatrix} i \\ kj \end{Bmatrix} v_i\right]\mathbf{e}^k = v_{k,j}\mathbf{e}^k$$

Thus

$$\frac{\partial \mathbf{v}}{\partial x^j} = v_{k,j}\mathbf{e}^k \text{, where} \tag{16.6}$$

$$v_{k,j} = \frac{\partial v_k}{\partial x^j} - \begin{Bmatrix} i \\ kj \end{Bmatrix} v_i \tag{16.7}$$

gives the formula for the covariant derivative of the covariant components.

Similarly the covariant derivative of mixed components is arrived at as follows. Let $\mathsf{T} = t^i_j\, \mathbf{e}_i \mathbf{e}^j$. Then

$$\nabla \mathsf{T} = \frac{\partial \mathsf{T}}{\partial x^k} \nabla x^k = \frac{\partial \mathsf{T}}{\partial x^k} \mathbf{e}^k$$

where

$$\frac{\partial \mathsf{T}}{\partial x^k} = \frac{\partial t^i{}_j}{\partial x^k} \mathbf{e}_i \mathbf{e}^j + t^i{}_j \frac{\partial \mathbf{e}_i}{\partial x^k} \mathbf{e}^j + t^i{}_j \mathbf{e}_i \frac{\partial \mathbf{e}^j}{\partial x^k}$$

$$= \frac{\partial t^i{}_j}{\partial x^k} \mathbf{e}_i \mathbf{e}^j + t^i{}_j \left\{ {r \atop ik} \right\} \mathbf{e}_r \mathbf{e}^j - t^i{}_j \mathbf{e}_i \left\{ {j \atop kr} \right\} \mathbf{e}^r$$

$$= \left[\frac{\partial t^i{}_j}{\partial x^k} + t^s_j \left\{ {i \atop sk} \right\} - t^i_s \left\{ {s \atop kj} \right\} \right] \mathbf{e}_i \mathbf{e}^j$$

Hence if $\mathsf{T} = t^i_j\, \mathbf{e}_i \mathbf{e}^j$, then $\dfrac{\partial \mathsf{T}}{\partial x^k} = t^i_{j,k} \mathbf{e}_i \mathbf{e}^j$, where

$$t^i_{j,k} = \frac{\partial t^i_j}{\partial x^k} + \left\{ {i \atop sk} \right\} t^s_j - \left\{ {s \atop kj} \right\} t^i_s \tag{16.8}$$

Equation (16.8) gives the formula for the covariant derivative of components of the $\binom{1}{1}$ type. The extension to components of arbitrary type is completely similar.

Exercise 1. From the definition $g_{ij} = \dfrac{\partial \mathbf{R}}{\partial x^i} \cdot \dfrac{\partial \mathbf{R}}{\partial x^j}$ verify the formula

---

$$\left\{ {i \atop jk} \right\} = \frac{1}{2} g^{ir} \left(\frac{\partial g_{rj}}{\partial x^k} + \frac{\partial g_{rk}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^r} \right)$$

The notation Γ^i_{jk} is often used for $\left\{ {i \atop jk} \right\}$. Since these quantities do not follow the $\binom{1}{2}$ transformation law, the use of a special symbol such as { } emphasizes the fact that they are not components of a tensor.

Similarly, verify that the Christoffel symbols of the first kind

$$[ij,k] = \gamma_{ijk} = g_{rk} \left\{ {r \atop ij} \right\}$$

are equal to

$$\frac{1}{2} \left(\frac{\partial g_{kj}}{\partial x^i} + \frac{\partial g_{ki}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^k} \right)$$

Exercise 2. Show that the Christoffel symbols $\left\{ {i \atop jk} \right\}$ follow the transformation law

$$\overline{\left\{ {i \atop jk} \right\}} = \frac{\partial \bar{x}^i}{\partial x^r} \frac{\partial x^s}{\partial \bar{x}^k} \frac{\partial x^t}{\partial \bar{x}^j} \left\{ {r \atop st} \right\} + \frac{\partial \bar{x}^i}{\partial x^r} \frac{\partial^2 x^r}{\partial \bar{x}^k \partial \bar{x}^j}$$

which, in view of the second term, shows that $\left\{ {i \atop jk} \right\}$ are not the form of the components of a tensor with respect to an arbitrary coordinate system.

Similarly find the transformation law of the functions $\frac{\partial v^i}{\partial x^j}$ and observe that they are not tensor components.

By combining the above results show that the transformation law of the quantities $v^i{}_{,j} = \frac{\partial v^i}{\partial x^j} + \left\{ {i \atop jk} \right\} v^k$ satisfy the law of $\binom{1}{1}$ components.

Hint: on differentiating the identity $\frac{\partial \bar{x}^i}{\partial x^j} \frac{\partial x^j}{\partial \bar{x}^k} = \delta^i_k$ with respect to x^s we get the identity $\frac{\partial^2 \bar{x}^i}{\partial x^s \partial x^j} \frac{\partial x^j}{\partial \bar{x}^k} = - \frac{\partial \bar{x}^i}{\partial x^j} \frac{\partial^2 x^j}{\partial \bar{x}^r \partial \bar{x}^k} \frac{\partial \bar{x}^r}{\partial x^s}$.

Exercise 3. Verify the formula

$$\left\{ {i \atop ji} \right\} = \frac{\partial}{\partial x^j} \left(\log \sqrt{\det(g)} \right)$$

Using Eqns. (16.1)-(16.4), we can write some of the results of Ch. 15 more compactly:

For the acceleration vector $\ddot{\mathbf{R}}$ we have $\ddot{\mathbf{R}} = \ddot{x}^i \mathbf{e}_i + \left\{ {k \atop ij} \right\} v^i v^j \mathbf{e}_k$, or alternatively

$$\ddot{\mathbf{R}} = \frac{d\mathbf{v}}{dt} = \frac{\partial \mathbf{v}}{\partial x^i} \dot{x}^i = v^j{}_{,i} v^i \mathbf{e}_j = \left[\frac{\partial v^j}{\partial x^i} + \left\{ {j \atop ik} \right\} v^k \right] v^i \mathbf{e}_j$$

Similarly for the time derivative of a vector field $\mathbf{u} = u^i \mathbf{e}_i$ (the "intrinsic derivative") as we move along a path we have

$$\dot{\mathbf{u}} = \frac{\partial \mathbf{u}}{\partial x^i} \dot{x}^i = u^k{}_{,i} \mathbf{e}_k \dot{x}^i = u^k{}_{,i} \dot{x}^i \mathbf{e}_k$$

or alternatively, $\dot{\mathbf{u}} = \nabla \mathbf{u} \cdot \dot{\mathbf{R}}$, since for the gradient of a vector field we get

$$\nabla \mathbf{u} = \left(\frac{\partial u^k}{\partial x^j} \mathbf{e}_k + u^i \left\{ {k \atop ij} \right\} \mathbf{e}_k \right) \mathbf{e}^j = u^k{}_{,j} \, \mathbf{e}_k \mathbf{e}^j$$

(this is an easier solution to Exercise 2 above.)

Thus the formula for the gradient of a vector field may be written simply as

$$\nabla \mathbf{u} = \frac{\partial \mathbf{u}}{\partial x^i} \nabla x^i$$

For the divergence we get

$$\nabla \cdot \mathbf{u} = \frac{\partial u^i}{\partial x^i} + u^j \left\{ {k \atop ji} \right\} \mathbf{e}_k \cdot \mathbf{e}^i = \frac{\partial u^i}{\partial x^i} + u^j \left\{ {i \atop ji} \right\} \text{ so that}$$

$$\nabla \cdot \mathbf{u} = u^i{}_{,\,i} \text{ , which is the contraction of } \nabla \mathbf{u}.$$

For the Laplacian of a scalar field we get

$$\nabla^2 \phi = \nabla \cdot \nabla \phi = \nabla \cdot \left(\frac{\partial \phi}{\partial x^i} \mathbf{e}^i \right) = \nabla \cdot \left(\frac{\partial \phi}{\partial x^i} g^{ij} \mathbf{e}_j \right) = \left(\frac{\partial \phi}{\partial x^i} g^{ij} \right)_{,j}$$

This last expression can be reduced to the form

$$\nabla^2 \phi = g^{ij} \left(\frac{\partial \phi}{\partial x^i} \right)_{,\,j}$$

since the metric components g^{ij} "behave like constants under covariant differentiation." To see this, we have for any vector field $\mathbf{v}$

$$\frac{\partial \mathbf{v}}{\partial x^k} = v^i{}_{,k} \mathbf{e}_i = v_{j,k} \mathbf{e}^j$$

Dotting $\mathbf{e}^r$ into each side of the last equation, we get

$$v^r{}_{,k} = g^{jr} v_{j,k}$$

i.e.

$$(g^{rj} v_j)_{,k} = g^{rj} v_{j,k}$$

Similarly

$$(g_{rj} v^j)_{,k} = g_{rj} v^j{}_{,k}$$

The natural generalization of these concepts occurs in the theory of exterior differential forms. We shall not go into this here, but the interested reader is referred to the book of Nickerson, Spencer and Steenrod cited earlier.

FURTHER READING

The following books illustrate the applications of tensors to physical problems or elaborate and generalize the theory.

Bauer, E. *Champs de Vecteurs et de Tenseurs* (Paris: Masson, 1955.)
Bergmann, P. G. *Introduction to the Theory of Relativity* (New York: Prentice-Hall, Inc., 1942.)
Bergmann, P. G. *Basic Theories of Physics* New York: Prentice-Hall Inc., 1949.)
Bewley, L. V. *Tensor Analysis of Electric Circuits and Machines* (New York: Ronald Press, 1961.)
Birkhoff, G. D. *Relativity and Modern Physics* (Cambridge: Harvard University Press, 1923.)
Brand, L. *Vector and Tensor Analysis* (New York: John Wiley & Sons, Inc., 1947.)
Brillouin, L. *Les Tenseurs en Mécanique et en Elasticité* (Paris: Masson, 1938.)
Buck, R. C. *Advanced Calculus* (New York: McGraw-Hill Book Co., Inc., 1956.)
Cartan, E. J. *Leçons sur le géometrie des espaces de Riemann* 2nd ed. (Paris: Gauthier-Villars, 1946.)
Coburn, N. *Vector and Tensor Analysis* (New York: The MacMillan Co., 1955.)
Coe, C. J. *Theoretical Mechanics* (New York: The MacMillan Co., 1938.)
Corben, H. C. and Stehle, P. *Classical Mechanics* (New York: John Wiley & Sons, Inc., 1950.)

Craig, H. V. *Vector and Tensor Analysis* (New York: McGraw-Hill Book Co., Inc., 1943.)

Gans, R. *Vector Analysis with Applications to Physics* (London & Glasgow: Blackie & Sons Ltd., 1932.)

Geiringer, H. *Geometrical Foundations of Mechanics* (Providence, R. I.: Brown University Press, 1942.)

Gibbs, J. W. *Vector Analysis* (New Haven: Yale University Press, 1931. First Edition, 1901.)

Gibbs, W. J. *Tensors in Electrical Machine Theory* (London: Chapman & Hall, 1952.)

Goldstein, H. *Classical Mechanics* (Cambridge, Mass.: Addison-Wesley, 1950.)

Green, A. E. and Zerna, W. *Theoretical Elasticity* (Oxford: Clarendon, 1954.)

Haas, A. *Introduction to Theoretical Physics* (New York: D. Van Nostrand, Co., 1925.)

Hay, G. E. *Vector and Tensor Analysis* (New York: Dover Publications, Inc., 1953.)

Heaviside, O. *Electromagnetic Theory* (New York: Dover Publications, Inc., 1950.)

Kron, G. *Short Course in Tensor Analysis for Electrical Engineers* (New York: John Wiley & Sons, Inc., 1942.)

Lass, H. *Vector and Tensor Analysis* (New York: McGraw-Hill Book Co., Inc., 1950.)

Levi-Civita, T. *The Absolute Differential Calculus* (London & Glasgow: Blackie and Son Ltd., 1927.)

Lichnerowicz, A. *Algèbre et Analyse Linéaires* (Paris: Masson, 1947.)

Lichnerowicz, A. *Éléments de Calcul Tensoriel* (Paris, 1950.)

McConnell, A. J. *Applications of the Absolute Differential Calculus* (London & Glasgow: Blackie & Son Ltd., 1931.)

Mason, W. P. *Physical Acoustics and the Properties of Solids* (D. Van Nostrand Co., 1958.)

Michal, A. D. *Matrix and Tensor Calculus* (New York: John Wiley & Sons Inc., 1947.)

Milne, E. A. *Vectorial Mechanics* (New York: Interscience Publishers, 1948.)

Milne-Thomson, L. M. *Theoretical Hydrodynamics* 2nd ed., (New York: The MacMillan Co., 1950.)

Munroe, M. E. "Bringing Calculus Up-to-Date," American Mathematical Monthly, **65** (1958) p. 81-90.

Murnaghan, F. D. *Vector Analysis and the Theory of Relativity* (Baltimore: Johns Hopkins University Press, 1922.)

Nickerson, H. K., Spencer, D. C., Steenrod, N. E. *Advanced Calculus* (Princeton: D. Van Nostrand Co., Inc., 1959.)

Nye, J. P. *Physical Properties of Crystals: Their Representation by Tensors and Matrices.* (Oxford: Clarendon Press, 1957.)

Page, L. *Introduction to Theoretical Physics* 3rd ed. (New York: D. Van Nostrand Co., Inc., 1952.)

Pai, S. *Viscous Flow Theory II-Turbulent Flow* (Princeton, N. J.: D. Van Nostrand, 1957.) Chapter XI.

Rutherford, D. E. *Vector Methods; Applied to Differential Geometry, Mechanics and Potential Theory* 8th ed. (Edinburgh & London: Oliver & Boyd, 1954.)

Slebodzinski, W. *Formes Exterieures et Leurs Applications* Vol. 1, 1954 (Monografje Matematyczne v. 31)

Sokolnikoff, I. S. *Tensor Analysis: Theory and Applications* (New York: John Wiley & Sons, Inc., 1951.)

Sommerfeld, A. *Mechanics of Deformable Bodies* (New York: Academic Press, 1950.)

Spain, B. *Tensor Calculus* 2nd ed. (Edinburgh: Oliver & Boyd, 1956.)

Spiegel, M. R. *Theory and Problems of Vector Analysis and an Introduction to Tensor Analysis* (New York: Schaum Publishing Co., 1959.)

Synge, J. L. and Schild, A. *Tensor Calculus* (Toronto: University of Toronto Press, 1949.)

Thomas, T. Y. *Concepts from Tensor Analysis and Differential Geometry* (New York: Academic Press, 1961.)

Weatherburn, C. E. *Advanced Vector Analysis with Applications to Mathematical Physics* (London: Bell, 1928.)

Weatherburn, C. E. *Elementary Vector Analysis with Applications to Geometry and Physics* (London: Bell, 1931.)

Whitney, H. *Geometric Integration Theory* (Princeton: Princeton University Press, 1957.)

Willmore, T. J. *An Introduction to Differential Geometry* (Oxford: Clarendon Press, 1959.)

Wills, A. P. *Vector and Tensor Analysis* (New York: Prentice-Hall Inc., 1931.)

Zamansky, M. *Introduction à L'Algèbre et L'Analyse Modernes* (Paris: Dunod, 1958.)

Additions (February, 1978)

Aris, Rutherford. Vectors, Tensors, and the Basic Equations of Fluid Mechanics. Pentice-Hall Englewood Cliffs, NJ 1962.

Chorlton, Frank. Vector and Tensor Methods. Ellis Horwood, Limited. Chichester, England. 1976.

Lovelock, David; and Rund, Hanno. Tensors, Differential Forms, and Variational Principles. John Wiley & Sons, New York 1975.

Segel, Lee A. Mathematics Applied to Continuum Mechanics. Macmillan, New York 1977.

INDEX