Undergraduate Texts in Mathematics

James G. Simmonds

A Brief on Tensor Analysis

With 28 Illustrations

Springer-Verlag
New York Heidelberg Berlin

James G. Simmonds
Department of Applied Mathematics
and Computer Science
Thornton Hall
University of Virginia
Charlottesville, VA 22901
U.S.A.

AMS Classification: 15-01, 15A72

With 28 illustrations.

Library of Congress Cataloging in Publication Data

Simmonds, James G.
A brief on tensor analysis.
(Undergraduate texts in mathematics)
Includes index.
1. Calculus of tensors. I. Title. II. Series.
QA433.S535 515′.63 82-702
AACR2

Softcover reprint of the hardcover 1st edition 1982

9 8 7 6 5 4 3 2 1

ISBN-13: 978-1-4684-0143-1 e-ISBN-13: 978-1-4684-0141-7
DOI: 10.1007/978-1-4684-0141-7

To my father,
My first and greatest teacher

Contents

Preface

When I was an undergraduate, working as a co-op student at North American Aviation, I tried to learn something about tensors. In the Aeronautical Engineering Department at MIT, I had just finished an introductory course in classical mechanics that so impressed me that to this day I cannot watch a plane in flight—especially in a turn—without imaging it bristling with vectors. Near the end of the course the professor showed that, if an airplane is treated as a rigid body, there arises a mysterious collection of rather simple-looking integrals called the components of the moment of inertia tensor. Tensor—what power those two syllables seemed to resonate. I had heard the word once before, in an aside by a graduate instructor to the *cognoscenti* in the front row of a course in strength of materials. "What the book calls stress is actually a tensor. . . ."

With my interest twice piqued and with time off from fighting the brushfires of a demanding curriculum, I was ready for my first serious effort at self-instruction. In Los Angeles, after several tries, I found a store with a book on tensor analysis. In my mind I had rehearsed the scene in which a graduate student or professor, spying me there, would shout, "You're an undergraduate. What are you doing looking at a book on tensors?" But luck was mine: the book had a plain brown dust jacket. Alone in my room, I turned immediately to the definition of a tensor: "A 2nd order tensor is a collection of n^2 objects that transform according to the rule . . ." and thence followed an inscrutable collection of superscripts, subscripts, overbars, and partial derivatives. A pedagogical disaster! Where was the connection with those beautiful, simple, boldfaced symbols, those arrows that I could visualize so well?

I was not to find out until after graduate school. But it is my hope that, with this book, you, as an undergraduate, may sail beyond that bar on which I once floundered. You will find that I take nearly three chapters to prepare you for the shock of the tensor transformation formulas. I don't try to hide them—they're the only equations in the book that are boxed. But long before, about halfway through Chapter 1, I tell you what a 2nd order tensor *really* is—a linear operator that sends vectors into vectors. If you apply the stress tensor to the unit normal to a plane through a point in a body, then out comes the stress vector, the force/area acting across the plane at that point. (That the stress vector is linear in the unit normal, i.e., that a stress tensor even exists, is a gift of nature; nonlinearity is more often the rule.) The subsequent "débauche des indices" that follows this tidy definition of a 2nd order tensor is the result of exposing the gears of a machine for grinding out the workings of a tensor. Abolish the machine and there is no hope of producing numerical results except in the simplest of cases.

This book falls into halves: Algebra and Calculus. The first half of the first half (Chapter 1) emphasizes concepts. Here, I have made a special effort to relate the mathematical and physical notions of a vector. I acknowledge my debt to Hoffman's intriguing little book, *About Vectors* (Dover, 1975). (But there are points where we differ—I disagree with his contention that vectors cannot represent finite rotations.) Chapter 2 deals mostly with the index apparatus necessary to represent and manipulate vectors and tensors in general bases. Chapter 3, through the vehicle of Newton's law of motion, introduces moving frames and the Christoffel symbols. To help keep the basic kinematic ideas and their tensor generalizations in mind simultaneously, I list a number of equations in dual form, a device that I have found successful in the classroom. The last chapter starts with a homely example of the gradient and builds to the covariant derivative. Throughout this chapter there are applications to continuum mechanics. Although the basic equations (excluding electricity and magnetism) were known by the 1850's, it was only under the spur of general relativity that tensor analysis began to diffuse into this older field. (In my own specialty, shell theory, tensor analysis did not appear until the early 1940's, in the Soviet literature, even though the underlying theory of surfaces and their tensor description had been central to the understanding of general relativity.)

I have provided no systematic lists of grad, div, curl, etc. in various coordinate systems. Such useful information can be found in Magnus, Oberhettinger, and Soni, *Formulas and Theorems for the Special Functions of Mathematical Physics,* 3rd enlarged edition, Chapter XII, Springer-Verlag 1966; or in Gradshteyn and Ryzhik, *Tables of Integrals, Series and Products,* 4th edition, corrected and enlarged, Academic Press, 1980.

It is a happy thought that much of the drudgery involved in expanding equations and verifying solutions in specific coordinate systems can now be done by computers, programmed to do symbol manipulation. The interested reader should consult "Computer Symbolic Math in Physics Eduation," by D. R. Stoutemyer, *Am. J. Phys.*, vol. 49 (1981), pp. 85–88, or "A Review of Algebraic Computing in General Relativity," by R. A. d'Inverno, Chapter 16 of *General Relativity and Gravitation,* vol. 1, ed. A. Held, Plenum Press, N.Y. and London, 1980.

I am pleased to acknowledge the help of three friends: Mark Duva, a former student, who, in his gracious but profound way, let me get away with nothing in class; Bruce Chartres, who let me filter much of this book through his fine mind; and Ernst Soudek, who, though not a native speaker, tuned the final manuscript with his keen ear for English.

Finally, my thanks to Carolyn Duprey and Ruth Nissley, who typed the original manuscript, and then with patience and good humor, retyped what must have seemed to be hundreds of petty changes.

JAMES G. SIMMONDS

CHAPTER I

Introduction: Vectors and Tensors

The magic of this theory will hardly fail to impose itself on anybody who has truly understood it; it represents a genuine triumph of the method of absolute differential calculus, founded by Gauss, Riemann, Christoffel, Ricci and Levi–Civita.[1]

This little book is about tensor analysis, as Einstein's philosophers' stone, the absolute differential calculus, is called nowadays. I have written it, though, with an eye not toward general relativity, but to *continuum mechanics,* a more modest theory that attempts to predict the gross behavior of "the masses of matter we see and use from day to day: air, water, earth, flesh, wood, stone, steel, concrete, glass, rubber, ..."[2]

Continuum mechanics is a limiting case of general relativity; yet it is best treated on its own merits. Viewed thus, there is a fundamental difference at the foundations of the two theories. The geometry of continuum mechanics is that of *three-dimensional Euclidean space* (E_3 for short) and *the real line,* R. The geometry of general relativity is that of a *four-dimensional Riemannian manifold.* (A sphere is a two-dimensional Riemannian manifold.) To those who will settle for nothing less than a complete understanding of general relativity (and who, therefore, will want to consult *Gravitation,* by Misner, Thorne, and Wheeler), take heart. From the tools that we shall fashion comes the gear to scale that pinnacle. And to those content to cultivate the

[1] Albert Einstein, "Contribution to the Theory of General Relativity", 1915; as quoted and translated by C. Lanczos in *The Einstein Decade,* p. 213.

[2] Truesdell and Noll, *The Non-Linear Field Theories of Mechanics,* p. 1. Two outstanding introductory texts on continuum mechanics are *A First Course in Rational Continuum Mechanics* by Truesdell and *Continuum Mechanics* by Chadwick.

garden of continuum mechanics, let me say that, embedded within it, are intrinsically curved two-dimensional continua, called shells, that in dwarf form exhibit nearly all of the mathematical foliage found in full-flowered general relativity.

In attempting to give mathematical form to the laws of mechanics, we face a dichotomy. On the one hand, if physical events and entities are to be quantified, then a (reference) *frame* and a *coordinate system* within that frame must be introduced.[3] On the other hand, as a frame and coordinates are mere scaffolding, it should be possible to express the laws of physics in frame- and coordinate-free form, i.e. in *invariant* form. Indeed this is the great program of general relativity.

In continuum mechanics, however, there are exceptional frames called *inertial*; Newton's Law of motion for a particle—force equals mass times acceleration—holds only in such frames.[4] A basic concern of continuum mechanics is therefore how laws such as Newton's change from one frame to another.[5] Save for the last exercise in the book, we shall not analyze changes of frame. Rather we shall study how, within a *fixed* frame, the mathematical *representation* of a physical object or law changes when one coordinate system (say Cartesian) is replaced by another (say spherical).

In what follows, I have assumed that you remember some of the plane and solid geometry that you once learned and that you have seen a bit of vector algebra and calculus. For conciseness, I have omitted a number of details and examples that you can find in texts devoted to vectors. At the same time I have emphasized several points, especially those concerning the physical meaning of vector addition and component representation, that are *not* found in most conventional texts. The exercises at the end of each chapter are intended to amplify and to supplement material in the text.

[3] A frame is a mathematical representation of a physical apparatus which assigns to each event e in the physical world $\mathcal{W}$ a unique *place* (i.e. point) in E_3 and a unique *instant* on the real line R. I like to imagine an idealized, all-seeing stereographic movie camera mounted on 3 rigid, mutually perpendicular rods. The rods have knife edges that intersect at a point and one of the rods carries a scratch to fix a unit of length. The 3 knife edges (indefinitely prolonged) are represented by a right-handed Cartesian reference frame $0xyz$ in E_3, and one instant (arbitrarily chosen) is taken as the origin of the real line. The exposed film is a physical realization of a *framing* (to use the terminology of Truesdell, *op. cit.*), i.e. a map f from $\mathcal{W}$ to $E_3 \times R$.

A coordinate system in a frame assigns to each place a unique triple of real numbers (u, v, w) called spatial coordinates and to each instant a unique number t call the time.

[4] Inertial frames are also special, but in a different way, in general relativity where frames *are* coordinate systems! (and physics *is* geometry). An inertial frame may be introduced in general relativity in the same way as a two-dimensional Cartesian coordinate system may be introduced in an arbitrarily small neighborhood of a point on a sphere.

[5] To change frames means, for example, to film the world with a *copy* of our super camera. If the cameras are in relative motion, then the two exposed films f and f_* will map the same event e into different places P and P_* in E_3 and into different instants T and T_* on R. Of course, the two sets of knife edges are represented by the same frame $0xyz$ and the cameras run at the same rate. This *change of frame* is a special type of time-dependent map of $E_3 \times R$ into itself that preserves the distance and elapsed time between two events. When the elapsed time is zero, this transformation has the same form as *a rigid body motion*. See Exercise *4.19* and *Truesdell, op. cit.*

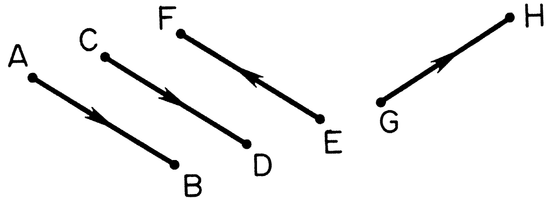


Figure 1.1

Three-Dimensional Euclidean Space

Three-dimensional Euclidean space, E_3, may be characterized by a set of axioms that express relationships among primitive, undefined quantities called points, lines, etc.[6] These relationships so closely correspond to the results of ordinary measurements of distance in the physical world that, until the appearance of general relativity, it was thought that Euclidean geometry was *the* kinematic model of the universe.

Directed Line Segments

Directed line segments, or *arrows,* are of fundamental importance in Euclidean geometry. Logically, an arrow is an ordered pair of points, (A,B). A is called the *tail* of the arrow and B the *head.* It is customary to represent such an arrow typographically as $\overline{AB}$, and pictorially as a line segment from A to B with an arrow head at B. (To avoid crowding, the arrow head may be moved towards the center of the segment). Assigning a length to an arrow or multiplying it by a real number (holding the tail fixed) are precisely defined operations in E_3.

Two arrows are said to be *equivalent* if one can be brought into coincidence with the other by a parallel translation.[7] In Fig. 1.1, $\overline{AB}$ and $\overline{CD}$ are equivalent, but neither $\overline{AB}$ and $\overline{EF}$ nor $\overline{AB}$ and $\overline{GH}$ are.

The set of *all* arrows equivalent to a given arrow is called the (geometric) *vector* of that arrow and is usually denoted by a symbol such as $\mathbf{v}$. A vector is an example of an *equivalence class* and, by convention, a vector is represented by any one of its arrows.

Equivalence classes are more familiar (and more useful) than you may realize. Suppose that we wish to carry out, on a computer, exact arithmetic operations on rational numbers. Then, for example, $2/3$ must be read in as the

[6] This was Hilbert's program: reduce geometry to a branch of logic. No pictures allowed! See, for example, the discussion at the end of Eisenhart's *Analytic Geometry*. Our approach, however, shall be informal and visual.

[7] A definition that makes no sense on a sphere. Why?

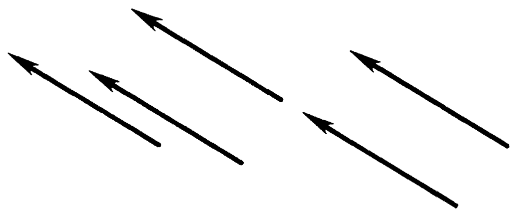


Figure 1.2

ordered pair of integers (2,3). We test for the equivalence of two ordered pairs of integers (a,b) and (c,d) stored within the computer by checking to see if $ad = bc$. In doing so, we are tacitly using the definition of a rational number a/b as the equivalence class of all ordered pairs of integers (c,d) such that $ad = bc$.

In practice, it is expedient (and rarely causes problems) to confound a "number", such as two-thirds, with its various representations e.g., 2/3, 4/6, etc. Likewise, we shall be using the term "vector" when we mean one of its arrows (and vice versa), relying on context for the proper interpretation. Thus in Fig. 1.2 we call any one of the equivalent arrows "the vector $\mathbf{v}$."

The length of a vector $\mathbf{v}$ is denoted by $|\mathbf{v}|$ and defined to be the length of any one of its arrows. The zero vector, $\mathbf{0}$, is the unique vector having zero length. We call the unit vector

$$\bar{\mathbf{v}} = \mathbf{v}/|\mathbf{v}|, \qquad \mathbf{v} \neq \mathbf{0}, \tag{1.1}$$

the direction of $\mathbf{v}$; $\mathbf{0}$ has no direction.

We may choose, *arbitrarily*, a point 0 in E_3 and call it the *origin*. The vector $\mathbf{x}$ (of the arrow) from 0 to a point P is called the *position* of P. We shall sometimes write $P(\mathbf{x})$ as shorthand for "the point with position $\mathbf{x}$."

Addition of Two Vectors

Addition of two vectors $\mathbf{u}$ and $\mathbf{v}$ may be defined in two equivalent ways.[8]

A. The Head-to-Tail-Rule (Fig. 1.3a)

Take any arrow representing $\mathbf{u}$, say $\overline{AB}$. For this choice there is a unique arrow $\overline{BC}$ representing $\mathbf{v}$; $\mathbf{u} + \mathbf{v}$ is defined to be the vector of the arrow $\overline{AC}$. This definition is convenient if one wishes to add a string of vectors (Exercise 1.1), but commutativity is not obvious. For reasons of symmetry it may be preferable to use the following.

[8] The equivalence and uniqueness of the two definitions can be proved from the postulates of Euclidean geometry.

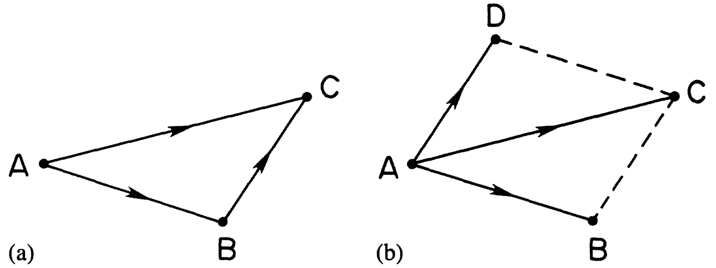


Figure 1.3

B. The Parallelogram Rule (Fig. 1.3b)

Let **u** and **v** be represented by any two arrows having coincident tails, say $\overline{AB}$ and $\overline{AD}$. Then $\mathbf{u} + \mathbf{v}$ is the vector of the arrow $\overline{AC}$, where C is the vertex opposite A of the parallelogram having $\overline{AB}$ and $\overline{AD}$ as co-terminal edges.

Adding three or more vectors in this way is a bit awkward graphically, but $\mathbf{u} + \mathbf{v} + \mathbf{w}$ has the following neat interpretation in three-dimensional space. Let **u**, **v**, and **w** be represented by arrows $\overline{AB}$, $\overline{AC}$, and $\overline{AD}$. Then $\mathbf{u} + \mathbf{v} + \mathbf{w}$ is the vector of the arrow, with tail at A, lying along the diagonal of the parallelopiped having $\overline{AB}$, $\overline{AC}$, and $\overline{AD}$ as co-terminal edges. See Exercise 1.2.

Multiplication of a Vector v by a Scalar α

Multiplication of a vector **v** by a scalar α is defined in an obvious way: if $\overline{AB}$ is an arrow representing **v**, then $\alpha\mathbf{v}$ is the vector of the arrow $\alpha\overline{AB}$, (Recall that in Euclidean geometry, we do multiplication with the aid of similar triangles).

The set of all geometric vectors, together with the operations of addition and multiplication by a scalar, form a *linear vector space*. Other familiar examples of linear vector spaces are the set of all polynomials of degree n, the set of all solutions of a linear homogeneous differential equation of order n, and the set of all $m \times n$ matrices together with appropriate definitions of addition and multiplication by a scalar.

Things That Vectors May Represent

Many physical and kinematic objects—the displacement from one point to another, a force acting on a particle, the finite rotation of a rigid body about

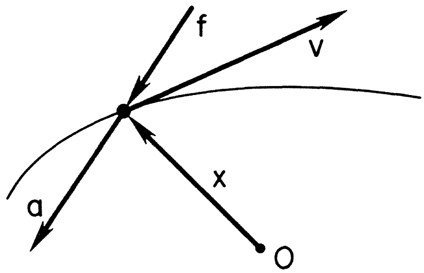


Figure 1.4

an axis—have direction and magnitude. We may represent these attributes by vectors.[9] In doing so, we must keep in mind two fundamental points.

A. Different Types of Objects are Represented by Vectors That Belong to Different Vector Spaces

Otherwise we could add, for example, a force to a displacement. *Nevertheless,* for conciseness and clarity, we often place different types of vectors in the same picture, as in Fig. 1.4 that shows the position **x**, the velocity **v**, the acceleration **a**, and the force **f** acting on a cannon ball flying through the air.

B. Vector Addition May or May Not Reflect an Attribute of the Objects Represented

For displacements, forces, or velocities, there are obvious physical analogues of vector addition; for successive finite rotations of a rigid body about a fixed point, there is not. We shall say more about vector addition later.[10]

Cartesian Coordinates

Thanks to Descartes, we may characterize three-dimensional Euclidean space in algebraic terms as follows. Through the origin 0 draw three mutually perpendicular ($\perp$) but otherwise arbitrarily chosen lines. Call one of these the

[9] A phrase such as "the vector **f** representing the attributes of a force" is a precise mouthful. For palatability we may pare it to "the vector **f** representing a force" or even to "a force **f**." When the context is clear, the tidbit "**f**" may suffice.

[10] Physically, there are differences in how displacements, forces and velocities add: displacements follow the head-to-tail-rule, forces, at a common point, the parallelogram rule, while velocities "add like vectors" only because of the postulates of continuum mechanics concerning moving frames. Velocities do not add like vectors in relativity theory.

x-axis and on it place a point $I \neq 0$. The ray (or half-line) from 0 containing I is called *the positive x-axis*. $\overline{OI}$ is called *the unit arrow along the x-axis* and we denote its vector by $\mathbf{e}_x$. Choose one of the remaining lines through 0, call it the y-axis, and place on it a point J such that the length of $\overline{OJ}$ is equal to that of $\overline{OI}$. $\overline{OJ}$ is called the y-unit arrow and we denote its vector by $\mathbf{e}_y$. The remaining line through 0 is called the z-axis and, by arbitrarily adopting *the right-hand rule,* we may place a unique point K on the z-axis[11] such that the length of $\overline{OK}$ is equal to that of $\overline{OI}$. $\overline{OK}$ is the z-unit arrow and $\mathbf{e}_z$ denotes its vector.

Any point P may be represented by an *ordered triple* of real numbers (x,y,z), called the *Cartesian coordinates of P*. The first number, or x-coordinate, is the *directed distance* from the yz-plane to P. Thus x is positive if P lies on the same side of the yz-plane as does I and negative if it lies on the opposite side; x is zero if P lies in the yz-plane. The second and third coordinates, y and z, are defined in an analogous way. To indicate that a point P has coordinates (x,y,z) we sometimes write $P(x,y,z)$.

When a vector $\mathbf{v}$ is represented by the arrow whose tail is the origin 0, then the coordinates of the head of this arrow, say (v_x,v_y,v_z), are called the *Cartesian components* of $\mathbf{v}$. We indicate this relation by writing $\mathbf{v} \sim (v_x,v_y,v_z)$. Thus, in particular,

$$\mathbf{e}_x \sim (1,0,0),\ \mathbf{e}_y \sim (0,1,0),\ \mathbf{e}_z \sim (0,0,1) \tag{1.2}$$

$$\mathbf{x} \sim (x,y,z). \tag{1.3}$$

With a way of assigning Cartesian components (v_x,v_y,v_z) to a vector $\mathbf{v}$, and vice-versa, we may easily deduce the following relations.

i) $$|\mathbf{v}| \equiv \sqrt{v_x^2 + v_y^2 + v_z^2}, \qquad (1.4)^{12}$$

by the Pythagorean theorem.

ii) If α is a real number, then

$$\alpha\mathbf{v} \sim (\alpha v_x,\ \alpha v_y,\ \alpha v_z). \tag{1.5}$$

iii) If $\mathbf{w} \sim (w_x,w_y,w_z)$, then

$$\mathbf{v} \pm \mathbf{w} \sim (v_x \pm w_x,\ v_y \pm w_y,\ v_z \pm w_z). \tag{1.6}$$

iv) $$\mathbf{v} = \mathbf{w} \Leftrightarrow v_x = w_x,\ v_y = w_y,\ v_z = w_z. \tag{1.7}$$

These relations allow us to represent and to manipulate vectors on a computer.

[11] This means that if we curl the fingers of our right hand from $\overline{OI}$ to $\overline{OJ}$, then our thumb will point in the direction of $\overline{OK}$.

[12] Here we are deducing the algebraic properties of E_3 from its geometric ones. To go the other way, which is easier in many respects, we *define* E_3 to be the set of all ordered triples of real numbers (x,y,z) such that the distance between any two points (x_1,y_1,z_1) and (x_2,y_1,z_2) is given by

$$\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}.$$

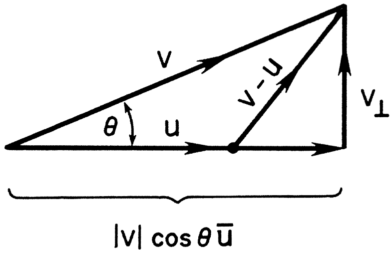


Figure 1.5

The Dot Product

The dot product of two vectors **u** and **v**, denoted by $\mathbf{u}\cdot\mathbf{v}$, arises in many different physical and geometric contexts.[13] Some authors define the dot product by the formula

$$\mathbf{u}\cdot\mathbf{v} = |\mathbf{u}||\mathbf{v}| \cos \theta, \qquad 0 \leq \theta \leq \pi, \tag{1.8}$$

where "θ is the angle between **u** and **v**." This definition seems simple and reasonable. But wait. If all we know is how to compute the length of a vector, how do we compute θ? We can't unless we turn (1.8) into the definition of θ. Then, we have no choice but to come up with a definition for $\mathbf{u}\cdot\mathbf{v}$ that is independent of θ. Here goes.

If $|\mathbf{u}||\mathbf{v}| = 0$, take $\mathbf{u}\cdot\mathbf{v} = 0$. Otherwise, consider Fig. 1.5, which may be constructed using strictly geometric methods. The double-headed arrow indicates that θ is always nonnegative, regardless of the relative orientation of **u** and **v**. As shown, **v** may be expressed as the sum of a vector $|\mathbf{v}| \cos \theta\, \bar{\mathbf{u}}$, parallel to **u**, and a vector $\mathbf{v}_\perp$, $\perp$ to **u**. Take first the larger of the two right triangles in the sketch and then the smaller to obtain, by the Pythagorean theorem,

$$|\mathbf{v}|^2 = |\mathbf{v}|^2 \cos^2 \theta + |\mathbf{v}_\perp|^2, \tag{1.9}$$

$$\begin{aligned} |\mathbf{v} - \mathbf{u}|^2 &= (|\mathbf{v}| \cos \theta - |\mathbf{u}|)^2 + |\mathbf{v}_\perp|^2 \\ &= |\mathbf{v}|^2 \cos^2 \theta - 2|\mathbf{u}||\mathbf{v}| \cos \theta + |\mathbf{u}|^2 + |\mathbf{v}_\perp|^2 \\ &= -2|\mathbf{u}||\mathbf{v}| \cos \theta + |\mathbf{u}|^2 + |\mathbf{v}|^2, \end{aligned} \tag{1.10}$$ [14]

by (1.9). Comparing (1.8) with (1.10), we are led to the definition

$$\mathbf{u}\cdot\mathbf{v} \equiv \tfrac{1}{2}(|\mathbf{u}|^2 + |\mathbf{v}|^2 - |\mathbf{v} - \mathbf{u}|^2). \tag{1.11}$$

Note that if $|\mathbf{u}||\mathbf{v}| = 0$, then $\mathbf{u}\cdot\mathbf{v} = 0$, in agreement with our earlier definition. If $\mathbf{u} = \mathbf{v}$,

$$\mathbf{v}\cdot\mathbf{v} = |\mathbf{v}|^2. \tag{1.12}$$

[13] There are certain technicalities to be considered when we try to interpret the dot product physically. See the discussion at the end of the chapter.

[14] This result is usually called the *law of cosines*.

Two vectors whose dot product is zero are said to be *orthogonal* ($\perp$).

To evaluate $\mathbf{u}\cdot\mathbf{v}$ on a computer, we need a component representation. This is easy to find. If $\mathbf{u} \sim (u_x,u_y,u_z)$ and $\mathbf{v} \sim (v_x,v_y,v_z)$, then, from (1.4) and (1.11),

$$\mathbf{u}\cdot\mathbf{v} = u_x v_x + u_y v_y + u_z v_z. \tag{1.13}$$

Suppose that, without changing $\mathbf{u}$ or $\mathbf{v}$, we were to introduce another set of Cartesian axes. The components of $\mathbf{u}$ and $\mathbf{v}$ relative to the new axes would, in general, be different, but would the right side of (1.13) change? The individual terms, being the products of components, would, of course, but their sum would not. Why? Because (1.11) defines $\mathbf{u}\cdot\mathbf{v}$ in terms of lengths and lengths, in Euclidean geometry, may be defined without reference to coordinate axes. We say, therefore, that the dot product is a *geometric invariant.*

Another key property of the dot product is that it is *distributive with respect to addition,* i.e.

$$\mathbf{u}\cdot(\mathbf{v} + \mathbf{w}) = \mathbf{u}\cdot\mathbf{v} + \mathbf{u}\cdot\mathbf{w}, \qquad \forall \mathbf{u},\ \mathbf{v},\ \mathbf{w}. \tag{1.14}$$

To prove (1.14) let $\mathbf{w} \sim (w_x,w_y,w_z)$. Then, by (1.13),

$$\begin{aligned}\mathbf{u}\cdot(\mathbf{v} + \mathbf{w}) &= u_x(v_x + w_x) + u_y(v_y + w_y) + u_z(v_z + w_z)\\ &= u_x v_x + u_y v_y + u_z v_z + u_x w_x + u_y w_y + u_z w_z.\\ &= \mathbf{u}\cdot\mathbf{v} + \mathbf{u}\cdot\mathbf{w}. \qquad \square\end{aligned}$$

This proof illustrates an important point: often, geometric facts are proved most easily using Cartesian coordinates. Yet, as (1.14) is coordinate-free, it must follow directly from (1.11). This is Exercise 1.7. Always try to interpret formulas both algebraically and geometrically, for one of these viewpoints, though awkward for proof, may lead to new insights or suggest generalizations.

PROBLEM 1.1.

If $\mathbf{u} \sim (1,2,3)$ and $\mathbf{v} \sim (-3,1,-2)$, compute $\mathbf{u}\cdot\mathbf{v}$ and the enclosed angle.

SOLUTION.

From (1.13)

$$\mathbf{u}\cdot\mathbf{v} = (1)(-3) + (2)(1) + (3)(-2) = -7,$$

and from (1.8)

$$\theta = \cos^{-1}\frac{\mathbf{u}\cdot\mathbf{v}}{|\mathbf{u}|\,|\mathbf{v}|}, \qquad 0 \le \theta \le \pi.$$

Now,

$$|\mathbf{u}| = \sqrt{1^2 + 2^2 + 3^2} = \sqrt{14}$$
$$|\mathbf{v}| = \sqrt{(-3)^2 + 1^2 + (-2)^2} = \sqrt{14}.$$

Hence

$$\theta = \cos^{-1}(-7/14) = 120° = 2\pi/3 \text{ radians}.$$

Cartesian Base Vectors

Given the Cartesian components (v_x, v_y, v_z) of *any* vector **v**, (1.6) and (1.5) allow us to set

$$\begin{aligned}(v_x, v_y, v_z) &= (v_x, 0, 0) + (0, v_y, 0) + (0, 0, v_z) \\ &= v_x(1,0,0) + v_y(0,1,0) + v_z(0,0,1).\end{aligned} \tag{1.15}$$

Recalling (1.2) and (1.6), we infer that **v** has the *unique* representation

$$\mathbf{v} = v_x\mathbf{e}_x + v_y\mathbf{e}_y + v_z\mathbf{e}_z. \tag{1.16}$$

The set $\{\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\}$ is called the *standard Cartesian basis*, and its elements, the *Cartesian base vectors*. The Cartesian components of **v** may be referred to, alternatively, as the *components of* **v** *relative to the basis* $\{\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\}$.

The Interpretation of Vector Addition

It is essential to emphasize the distinction between vector addition, as a reflection of a physical or kinematic attribute, and vector addition, as it is used on the right side of (1.16). Suppose, for example, that **v** represents the rotation of a rigid body about a fixed point. Let the direction and magnitude of **v** represent, respectively, the axis and angle of rotation, using the right hand rule to insure a unique correspondence. If this rotation is followed by another, represented by a vector **u**, then it is a fact from kinematics that these two successive rotations are equivalent to a single rotation that we may represent by a vector **w**. However, in general, $\mathbf{v} + \mathbf{u} \neq \mathbf{w}$, i.e. *successive finite rotations do not add like vectors.*[15] Nevertheless, the additions that appear on the right hand side of (1.16) *do* make sense if we *do not* interpret the summands *individually* as rotations. Let me try to illustrate this point with a simple analogy. Suppose that there are 20 students and 30 chairs in a classroom. For statistical purposes we may speak of "⅔ of a student per chair" even though, physically, there is no such thing as ⅔ of a student. The students and chairs are represented by whole numbers and what we did to get a statistic was to perform a *mathematical operation* on these whole numbers.

Vector addition *does* mirror a physical attribute when, for example, **v** in (1.16) represents a force. It is an experimental fact that "forces add like

[15] Rotations may also be represented by matrices. If the two successive rotations are represented by the matrices V and U, then they are equivalent to a single rotation represented by the matrix $W = UV$. **u**, **v**, or **w** may be identified, respectively, with the single real eigenvector of U, V, or W.

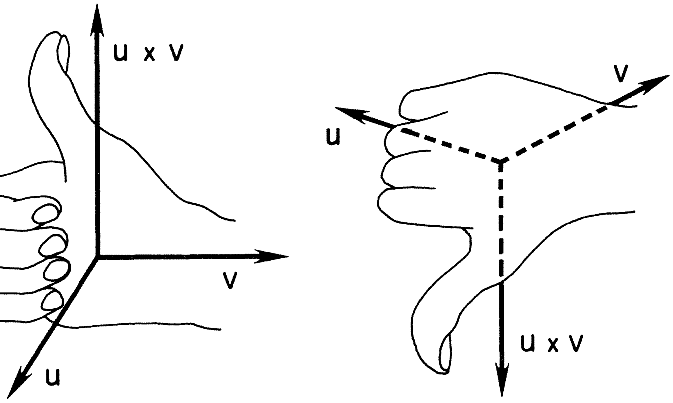


Figure 1.6

vectors." See Exercise 1.3. In this case the mathematical decomposition of a vector into components happens to have a direct physical interpretation. The analogy here might be that of a dining hall in which there are 20 one-gallon bottles of milk and 30 pitchers. Not only can we speak of "⅔ gallons of milk per pitcher," we can also, physically, put ⅔ gallons of milk into each pitcher. Here the mathematical operation of division of whole numbers has a physical counterpart.

The Cross Product

The cross product arises in mechanics when we want to compute the torque of a force about a point, in electromagnetics when we want to compute the force on a charge moving in a magnetic field, in geometry when we want to compute the volume of a parallelopiped or tetrahedron, and in many other physical and geometric situations.[16]

The cross product of two vectors **u** and **v** is denoted by $\mathbf{u} \times \mathbf{v}$ and *defined to be the right oriented area of a parallelogram having* **u** *and* **v** *as co-terminal edges*. That is,

$$|\mathbf{u} \times \mathbf{v}| = |\mathbf{u}||\mathbf{v}| \sin \theta, \qquad 0 \leq \theta \leq \pi \tag{1.17}$$

where the direction of $\mathbf{u} \times \mathbf{v}$ is that of the thumb on the right hand when the fingers are curled from **u** to **v**. Two possibilities are indicated in Fig. 1.6.

From the definition of the cross product,

$$\mathbf{u} \times \mathbf{v} = -\mathbf{v} \times \mathbf{u}. \tag{1.18}$$

Further useful properties of the cross product follow from a geometrical inter-

[16] The remarks made earlier concerning physical interpretations of the dot product apply to the cross product as well. See the discussion at the end of this chapter.

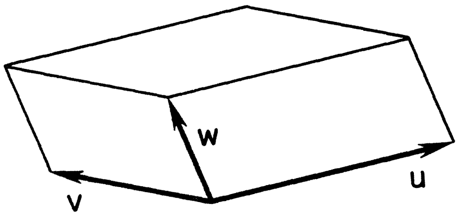


Figure 1.7

pretation of the *scalar triple product* $(\mathbf{u} \times \mathbf{v})\cdot\mathbf{w}$. Consider the parallelopiped having $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ as co-terminal edges and let vol $(\mathbf{u},\mathbf{v},\mathbf{w})$ denote its volume. By regarding this volume as the limit of the sum of *right* parallelopipeds of vanishingly small altitude and identical bases, each base parallelogram having $\mathbf{u}$ and $\mathbf{v}$ as co-terminal edges, we have

$$\text{vol}\,(\mathbf{u},\mathbf{v},\mathbf{w}) = |(\mathbf{u} \times \mathbf{v})\cdot\mathbf{w}|. \qquad (1.19)$$ [17]

If $(\mathbf{u},\mathbf{v},\mathbf{w})$, in that order, form a right handed system, as shown in Fig. 1.7, then, by symmetry,

$$\text{vol}\,(\mathbf{u},\mathbf{v},\mathbf{w}) = (\mathbf{u} \times \mathbf{v})\cdot\mathbf{w} = (\mathbf{w} \times \mathbf{u})\cdot\mathbf{v} = (\mathbf{v} \times \mathbf{w})\cdot\mathbf{u} \geq 0, \qquad (1.20)$$

i.e. $(\mathbf{u} \times \mathbf{v})\cdot\mathbf{w}$ *is invariant under cyclic permutation.*

With the aid of (1.20), we shall verify that *the cross product is distributive with respect to addition*:

$$\mathbf{u} \times (\mathbf{v} + \mathbf{w}) = \mathbf{u} \times \mathbf{v} + \mathbf{u} \times \mathbf{w}. \qquad (1.21)$$

To do so we use the fact that *two vectors are equal if and only if their dot products with all vectors* $\mathbf{c}$ *are equal.*[18] Thus,

$$\begin{aligned} [\mathbf{u} \times (\mathbf{v} + \mathbf{w})]\cdot\mathbf{c} &= (\mathbf{c} \times \mathbf{u})\cdot(\mathbf{v} + \mathbf{w}), \text{ by (1.20)} \\ &= (\mathbf{c} \times \mathbf{u})\cdot\mathbf{v} + (\mathbf{c} \times \mathbf{u})\cdot\mathbf{w}, \text{ by (1.14)} \\ &= (\mathbf{u} \times \mathbf{v})\cdot\mathbf{c} + (\mathbf{u} \times \mathbf{w})\cdot\mathbf{c}, \text{ by (1.20)} \\ &= (\mathbf{u} \times \mathbf{v} + \mathbf{u} \times \mathbf{w})\cdot\mathbf{c}, \text{ by (1.14).} \end{aligned}$$

□

PROBLEM 1.2.

Prove the *vector triple product* identity

$$(\mathbf{u} \times \mathbf{v}) \times \mathbf{w} \equiv (\mathbf{u}\cdot\mathbf{w})\mathbf{v} - (\mathbf{v}\cdot\mathbf{w})\mathbf{u} \qquad (1.22)$$

without introducing components.

[17] We hardly need calculus to get this. Imagine a deck of identical cards, each in the shape of a parallelogram. The volume of the deck is its thickness times the area of the face of a card, and this volume is obviously unchanged if the deck is sheared into the shape of a parallelopiped.

[18] Why? Well, if $\mathbf{u} = \mathbf{v}$, then, certainly $\mathbf{u}\cdot\mathbf{c} = \mathbf{v}\cdot\mathbf{c}$ for all $\mathbf{c}$. On the other hand, if $\mathbf{u}\cdot\mathbf{c} = \mathbf{v}\cdot\mathbf{c}$ for *all* $\mathbf{c}$ we can, in particular, take $\mathbf{c} = \mathbf{u} - \mathbf{v}$, which implies that $\mathbf{u}\cdot\mathbf{c} - \mathbf{v}\cdot\mathbf{c} = (\mathbf{u} - \mathbf{v})\cdot\mathbf{c} = (\mathbf{u} - \mathbf{v})\cdot(\mathbf{u} - \mathbf{v}) = |\mathbf{u} - \mathbf{v}|^2 = 0$. This in turn implies that $\mathbf{u} - \mathbf{v} = \mathbf{0}$, i.e., $\mathbf{u} = \mathbf{v}$.

SOLUTION.

The vector $\mathbf{u} \times \mathbf{v}$ is $\perp$ to the plane of $\mathbf{u}$ and $\mathbf{v}$. Therefore $(\mathbf{u} \times \mathbf{v}) \times \mathbf{w}$, which is $\perp$ to $\mathbf{u} \times \mathbf{v}$, must lie in the plane of $\mathbf{u}$ and $\mathbf{v}$. That is,

$$(\mathbf{u} \times \mathbf{v}) \times \mathbf{w} = A\mathbf{u} + B\mathbf{v}, \tag{*}$$

where A and B are unknown scalar functions of $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$. As the left side of (*) is $\perp$ to $\mathbf{w}$, it follows that

$$0 = \mathrm{A}\mathbf{u}\cdot\mathbf{w} + \mathrm{B}\mathbf{v}\cdot\mathbf{w},$$

which implies

$$\mathrm{A} = -\mathrm{C}(\mathbf{v}\cdot\mathbf{w}), \qquad \mathrm{B} = \mathrm{C}(\mathbf{u}\cdot\mathbf{w}),$$

where C is an unknown scalar function of $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$. Hence

$$(\mathbf{u} \times \mathbf{v}) \times \mathbf{w} \equiv \mathrm{C}(\mathbf{u},\mathbf{v},\mathbf{w})[(\mathbf{u}\cdot\mathbf{w})\mathbf{v} - (\mathbf{v}\cdot\mathbf{w})\mathbf{u}]. \tag{**}$$

In the special case $\mathbf{u} = \mathbf{w}$, (**) reduces to

$$(\mathbf{u} \times \mathbf{v}) \times \mathbf{u} = \mathrm{C}(\mathbf{u},\mathbf{v},\mathbf{u})[|\mathbf{u}|^2\mathbf{v} - (\mathbf{v}\cdot\mathbf{u})\mathbf{u}]. \tag{***}$$

Take the dot product of both sides with $\mathbf{v}$. Permuting the resulting scalar triple product on the left, we have

$$(\mathbf{u} \times \mathbf{v})\cdot(\mathbf{u} \times \mathbf{v}) = \mathrm{C}(\mathbf{u},\mathbf{v},\mathbf{u})[|\mathbf{u}|^2|\mathbf{v}|^2 - (\mathbf{u}\cdot\mathbf{v})^2]. \tag{****}$$

But

$$\begin{aligned}(\mathbf{u} \times \mathbf{v})\cdot(\mathbf{u} \times \mathbf{v}) &= |\mathbf{u} \times \mathbf{v}|^2 \\ &= |\mathbf{u}|^2|\mathbf{v}|^2 \sin^2\theta, \text{ by } (1.17) \\ &= |\mathbf{u}|^2|\mathbf{v}|^2(1 - \cos^2\theta) \\ &= |\mathbf{u}|^2|\mathbf{v}|^2 - (\mathbf{u}\cdot\mathbf{v})^2, \text{ by } (1.8).\end{aligned}$$

Comparing the last line of this expression with the right side of (****) we have that $\mathrm{C}(\mathbf{u},\mathbf{v},\mathbf{u}) = 1$.

Now take the dot product of both sides of (**) with $\mathbf{u}$. Permuting the resulting scalar triple product on the left, we have

$$[\mathbf{u} \times (\mathbf{u} \times \mathbf{v})]\cdot\mathbf{w} = \mathrm{C}(\mathbf{u},\mathbf{v},\mathbf{w})[(\mathbf{u}\cdot\mathbf{w})(\mathbf{v}\cdot\mathbf{u}) - (\mathbf{v}\cdot\mathbf{w})|\mathbf{u}|^2].$$

But (***) with $\mathrm{C}(\mathbf{u},\mathbf{v},\mathbf{u}) = 1$ implies that

$$[\mathbf{u} \times (\mathbf{u} \times \mathbf{v})]\cdot\mathbf{w} = (\mathbf{v}\cdot\mathbf{u})(\mathbf{u}\cdot\mathbf{w}) - |\mathbf{u}|^2\mathbf{v}\cdot\mathbf{w}.$$

Hence, $\mathrm{C}(\mathbf{u},\mathbf{v},\mathbf{w}) = 1$. □

Another useful interpretation of the scalar triple product is in terms of projected areas. See Exercise 1.15.

How can we find the Cartesian components of $\mathbf{u} \times \mathbf{v}$? This is not as easy as the analogous problem for $\mathbf{u}\cdot\mathbf{v}$. For one reason, $\mathbf{u} \times \mathbf{v}$ is a vector rather than a scalar, and for another, $\mathbf{u} \times \mathbf{v}$ is a beast that lives in three dimensions only! (There is an algebra created by Grassmann that gives meaning to objects such as $\mathbf{u} \times \mathbf{v}$ or $\mathbf{u} \times \mathbf{v} \times \mathbf{w}$ in higher dimensional Euclidean spaces where they are called wedge products.) The straightforward (but tedious) way is to note that repeated application of the distributive law (1.21) allows us to set

$$\begin{aligned}\mathbf{u} \times \mathbf{v} &= (u_x\mathbf{e}_x + u_y\mathbf{e}_y + u_z\mathbf{e}_z) \times (v_x\mathbf{e}_x + v_y\mathbf{e}_y + v_z\mathbf{e}_z) \\ &= u_xv_x\mathbf{e}_x \times \mathbf{e}_x + u_xv_y\mathbf{e}_x \times \mathbf{e}_y + u_xv_z\mathbf{e}_x \times \mathbf{e}_z \\ &\quad + u_yv_x\mathbf{e}_y \times \mathbf{e}_x + u_yv_y\mathbf{e}_y \times \mathbf{e}_y + u_yv_z\mathbf{e}_y \times \mathbf{e}_z \\ &\quad + u_zv_x\mathbf{e}_z \times \mathbf{e}_x + u_zv_y\mathbf{e}_z \times \mathbf{e}_y + u_zv_z\mathbf{e}_z \times \mathbf{e}_z. \end{aligned} \tag{1.23}$$

(Phew!) But $\mathbf{e}_x \times \mathbf{e}_u = \mathbf{0}$, $\mathbf{e}_x \times \mathbf{e}_y = \mathbf{e}_z$, $\mathbf{e}_x \times \mathbf{e}_z = -\mathbf{e}_y$, etc. (Draw a sketch!) Thus (1.23) reduces to

$$\mathbf{u} \times \mathbf{v} = \mathbf{e}_x(u_yv_z - u_zv_y) + \mathbf{e}_y(u_zv_x - u_xv_z) + \mathbf{e}_z(u_xv_y - u_yv_x). \tag{1.24}$$

An easy way to remember (1.24) is to write it as

$$\mathbf{u} \times \mathbf{v} = \begin{vmatrix} \mathbf{e}_x & \mathbf{e}_y & \mathbf{e}_z \\ u_x & u_y & u_z \\ v_x & v_y & v_z \end{vmatrix}, \tag{1.25}$$

where the *formal* determinant is to be expanded about its first row.

From (1.24), (1.25), and the property of a determinant that interchanging rows and column leaves the value unchanged and interchanging two rows (or columns) changes the sign, we get the useful formulas

$$(\mathbf{u} \times \mathbf{v})\cdot\mathbf{w} = \begin{vmatrix} w_x & w_y & w_z \\ u_x & u_y & u_z \\ v_x & v_y & v_z \end{vmatrix} = \begin{vmatrix} u_x & u_y & u_z \\ v_x & v_y & v_z \\ w_x & w_y & w_z \end{vmatrix} = \begin{vmatrix} u_x & v_x & w_x \\ u_y & v_y & w_y \\ u_z & v_z & w_z \end{vmatrix}. \tag{1.26}$$

PROBLEM 1.3.

Determine a unit vector $\mathbf{e}$ mutually $\perp$ to the vectors $\mathbf{a} \sim (1,-2,3)$ and $\mathbf{b} \sim (-1,0,1)$ using and without using the cross product.

SOLUTION.

In terms of the cross product, the vector we seek is

$$\mathbf{e} = \pm\frac{\mathbf{a} \times \mathbf{b}}{|\mathbf{a} \times \mathbf{b}|}.$$

From (1.25),

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{e}_x & \mathbf{e}_y & \mathbf{e}_z \\ 1 & -2 & 3 \\ -1 & 0 & 1 \end{vmatrix} \sim (-2,-4,-2).$$

Hence $|\mathbf{a} \times \mathbf{b}| = 2\sqrt{6}$, so that $\mathbf{e} \sim \pm(1/\sqrt{6}, 2/\sqrt{6}, 1/\sqrt{6})$.

To compute $\mathbf{e}$ without using the cross product, let $\mathbf{e} \sim (a,b,c)$. Then $\mathbf{e}\cdot\mathbf{a} = a - 2b + 3c = 0$ and $\mathbf{e}\cdot\mathbf{b} = -a + c = 0$, which implies that $b = 2a$ and $c = a$. Hence $\mathbf{e} \sim a(1,2,1)$. To make $|\mathbf{e}| = 1$, set, $a = \pm 1/\sqrt{6}$.

Alternate Interpretation of the Dot and Cross Product. Tensors

In physics we learn that a constant force **F** acting through a displacement **D** does **F**·**D** units of work. Unlike displacement, force is never measured directly, only its effects (extension of a spring, change in the resistance of a strain gauge, etc.). Also, force has different physical units than displacement. This suggests that forces—more precisely, the vectors that represent forces—lie in a different space than do displacements. If forces are measured in, say, pounds and displacements in feet, then the unit x-arrows $\overline{OI}$ that we set up arbitrarily in each space represent, respectively, one pound of force and one foot of distance. Think now: how would you graphically compute **F**·**D**? You might, first, bring the positive x, y, and z-axes of each space into coincidence and next, since the $\overline{OI}$'s were chosen arbitrarily, stretch or shrink one until it coincided with the other. Finally, the length of the orthogonal projection of **F** onto **D** would be multiplied graphically by $|\mathbf{D}|$ to obtain a number with the units of foot pounds.

All of this imaginary pushing and shoving was to give **F**·**D** a geometric meaning so that one could do graphical computations (an ancient art little practiced nowadays.) Isn't there a simpler way of looking at things that, while not affecting numerical computation, better mirrors different types of physical objects? Yes there is: a force **F** may be thought of, mathematically, as a representation of a *linear functional* that sends any vector **D** (a displacement) into a real number (called the work of **F** through **D**).

Analogous but more elaborate considerations hold for the cross product. For example, suppose that a force **F** acts at a point P with position **x**. Representing the torque about O as $\mathbf{x} \times \mathbf{F}$ leads to the notion of linear operators that send vectors into vectors. Such operators are called *2nd order tensors*. The name tensor comes from elasticity theory where in a loaded elastic body the stress tensor acting on a unit vector normal to a plane through a point delivers the tension (i.e., the force per unit area) acting across the plane at the point. See Exercise 1.20. Other important 2nd order tensors include the inertia tensor in rigid body dynamics, the strain tensor in elasticity and the momentum-flux tensor in fluid dynamics.

The simplest, nontrivial example of a 2nd order tensor that I can think of is the following. Let the *projection of a vector* **v** *on a vector* **u** be denoted and defined by

$$\text{Proj}_{\mathbf{u}}\mathbf{v} \equiv (\mathbf{v}\cdot\overline{\mathbf{u}})\overline{\mathbf{u}}. \tag{1.27}$$

The geometric meaning of $\text{Proj}_{\mathbf{u}}\mathbf{v}$ is shown in Fig. 1.8. The left side of (1.27) may be interpreted as the action of the *operator* $\text{Proj}_{\mathbf{u}}$ on the vector **v**. The right side of (1.27) tells us that $\text{Proj}_{\mathbf{u}}$ sends **v** into a vector of magnitude $\mathbf{v}\cdot\overline{\mathbf{u}}$ in the direction of **u**. To qualify for the title tensor, $\text{Proj}_{\mathbf{u}}$ must be linear. But if β and γ are arbitrary scalars and **w** is an arbitrary vector, then, by the properties of the dot product,

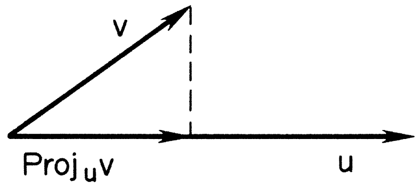


Figure 1.8

$$\begin{aligned}\mathrm{Proj}_{\mathbf{u}}(\beta\mathbf{v} + \gamma\mathbf{w}) &\equiv [(\beta\mathbf{v} + \gamma\mathbf{w})\cdot\bar{\mathbf{u}}]\bar{\mathbf{u}} \\ &= (\beta\mathbf{v}\cdot\bar{\mathbf{u}} + \gamma\mathbf{w}\cdot\bar{\mathbf{u}})\bar{\mathbf{u}} \\ &= \beta(\mathbf{v}\cdot\bar{\mathbf{u}})\bar{\mathbf{u}} + \gamma(\mathbf{w}\cdot\bar{\mathbf{u}})\bar{\mathbf{u}} \\ &= \beta\mathrm{Proj}_{\mathbf{u}}\mathbf{v} + \gamma\mathrm{Proj}_{\mathbf{u}}\mathbf{w}, \end{aligned} \qquad (1.28)$$

i.e., $\mathrm{Proj}_{\mathbf{u}}$ *is* linear and therefore is a tensor. Now work Exercise 1.5 to make sure that you can apply this tensor in a concrete situation.

The form of the right side of (1.27) suggests the following generalization. The *direct product* **uv** of two vectors **u** and **v** is a tensor that sends any vector **w** into a new vector according to the rule

$$\mathbf{uv}(\mathbf{w}) = \mathbf{u}(\mathbf{v}\cdot\mathbf{w}). \qquad (1.29)$$ [19]

Thus, in particular,

$$\mathrm{Proj}_{\mathbf{u}} = \bar{\mathbf{u}}\bar{\mathbf{u}}. \qquad (1.30)$$

Tensors such as $\mathrm{Proj}_{\mathbf{u}}$, that can be represented as direct products, are called *dyads*. As we shall see, *any* 2nd order tensor can be represented as a linear combination of dyads.

Definitions

To say that we are given a 2nd order tensor **T** means that we are told **T**'s action on (i.e., where it sends) any vector **v**[20]. Thus *two 2nd order tensors* **S** *and* **T** *are said to be equal if their action on all vectors* **v** *is the same*. More formally,

$$\mathbf{S} = \mathbf{T} \Leftrightarrow \mathbf{Sv} = \mathbf{Tv}, \qquad \forall\mathbf{v}, \qquad (1.31)$$

or equivalently,

[19] Many authors denote the direct product **uv** by $\mathbf{u} \otimes \mathbf{v}$.

[20] The action of **T** on **v** will be denoted by **Tv**, **T**(**v**), or **T**·**v**, as convenient. To a fastidious mathematician, the description of **T** is not complete without mention of its *domain*, the set of vectors on which **T** acts, and its *range*, the space into which **T** sends these vectors. We shall assume that the domain and range of a 2nd order tensor are obvious from the context, though often they are different; e.g., the domain of the moment of inertia tensor is the space of angular velocities, but its range is the space of rotational momenta (see Exercise 4.22).

$$\mathbf{S} = \mathbf{T} \Leftrightarrow \mathbf{u}\cdot\mathbf{S}\mathbf{v} = \mathbf{u}\cdot\mathbf{T}\mathbf{v}, \qquad \forall \mathbf{u},\mathbf{v}. \tag{1.32}$$

The *zero* and *identity (unit) tensors* are denoted and defined, respectively, by $\mathbf{0}\mathbf{v} = \mathbf{0}$, $\forall \mathbf{v}$, and $\mathbf{I}\mathbf{v} = \mathbf{v}$, $\forall \mathbf{v}$.

The *transpose* of a 2nd order tensor $\mathbf{T}$ is defined as that unique 2nd order tensor $\mathbf{T}^T$ such that

$$\mathbf{u}\cdot\mathbf{T}\mathbf{v} = \mathbf{v}\cdot\mathbf{T}^T\mathbf{u}, \qquad \forall \mathbf{u},\mathbf{v}. \tag{1.33}$$

A 2nd order tensor $\mathbf{T}$ is said to be

(a) *symmetric* if $\mathbf{T} = \mathbf{T}^T$.
(b) *skew* (or antisymmetric) if $\mathbf{T} = -\mathbf{T}^T$. (1.34)
(c) *singular* if there exists a $\mathbf{v} \neq \mathbf{0}$ such that $\mathbf{T}\mathbf{v} = \mathbf{0}$.

An arbitrary tensor $\mathbf{T}$ may always be decomposed into the sum of a symmetric and skew tensor as follows:

$$\mathbf{T} = \tfrac{1}{2}(\mathbf{T} + \mathbf{T}^T) + \tfrac{1}{2}(\mathbf{T} - \mathbf{T}^T), \tag{1.35}$$

$\mathbf{T} + \mathbf{T}^T = (\mathbf{T} + \mathbf{T}^T)^T$ being symmetric and $(\mathbf{T} - \mathbf{T}^T) = -(\mathbf{T} - \mathbf{T}^T)^T$ being skew.

PROBLEM 1.4.

If $\mathbf{v} \sim (v_x, v_y, v_z)$ and $\mathbf{T}\mathbf{v} \sim (-2v_x + 3v_z, -v_z, v_x + 2v_y)$, determine the Cartesian components of $\mathbf{T}^T\mathbf{v}$.

SOLUTION.

Let $\mathbf{T}^T\mathbf{v} \sim (a,b,c)$ and $\mathbf{u} \sim (\alpha,\beta,\gamma)$. By definition, $\mathbf{u}\cdot\mathbf{T}^T\mathbf{v} = \mathbf{v}\cdot\mathbf{T}\mathbf{u}$ for all vectors $\mathbf{u}$ and $\mathbf{v}$. That is,

$$\begin{aligned}\alpha a + \beta b + \gamma c &= v_x(-2\alpha + 3\gamma) + v_y(-\gamma) + v_z(\alpha + 2\beta)\\ &= \alpha(-2v_x + v_z) + \beta(2v_z) + \gamma(3v_x - v_y). \end{aligned} \tag{*}$$

As $\mathbf{u}$ is arbitrary, the coefficients of α, β, and γ on both sides of (*) must match. Hence

$$\mathbf{T}^T\mathbf{v} \sim (-2v_x + v_z,\ 2v_z,\ 3v_x - v_y).$$

The Cartesian Components of a Second Order Tensor

The Cartesian components of a second order tensor $\mathbf{T}$ fall out almost automatically when we apply $\mathbf{T}$ to any vector $\mathbf{v}$ expressed in terms of its Cartesian components. Thus

$$\begin{aligned}\mathbf{T}\mathbf{v} &= \mathbf{T}(v_x\mathbf{e}_x + v_y\mathbf{e}_y + v_z\mathbf{e}_z)\\ &= v_x\mathbf{T}\mathbf{e}_x + v_y\mathbf{T}\mathbf{e}_y + v_z\mathbf{T}\mathbf{e}_z, \text{ by the linearity of } \mathbf{T}. \end{aligned} \tag{1.36}$$

But $\mathbf{T}\mathbf{e}_x$, $\mathbf{T}\mathbf{e}_y$, and $\mathbf{T}\mathbf{e}_z$ are vectors and therefore may be expressed in terms of their Cartesian components, which we label as follows:

$$\mathbf{T}\mathbf{e}_x = T_{xx}\mathbf{e}_x + T_{yx}\mathbf{e}_y + T_{zx}\mathbf{e}_z \tag{1.37}$$

$$\mathbf{T}\mathbf{e}_y = T_{xy}\mathbf{e}_x + T_{yy}\mathbf{e}_y + T_{zy}\mathbf{e}_z \tag{1.38}$$

$$\mathbf{T}\mathbf{e}_z = T_{xz}\mathbf{e}_x + T_{yz}\mathbf{e}_y + T_{zz}\mathbf{e}_z. \tag{1.39}$$

The 9 coefficients T_{xx}, T_{xy}, . . . , T_{zz} are called the *Cartesian components of* **T**. We indicate this by writing $\mathbf{T} \sim T$, where T^T is the matrix of coefficients appearing in (1.37) to (1.39).

Here's how to keep the subscripts straight: From (1.37) to (1.39), $T_{xx} = \mathbf{e}_x \cdot \mathbf{T}\mathbf{e}_x$, $T_{xy} = \mathbf{e}_x \cdot \mathbf{T}\mathbf{e}_y$, etc.

PROBLEM 1.5.

Determine the Cartesian components of the tensor **T** defined in Problem 1.4.

SOLUTION.

Applying T successively to $\mathbf{e}_x$, $\mathbf{e}_y$, and $\mathbf{e}_z$, we have

$$\begin{aligned}
\mathbf{T}\mathbf{e}_x &\sim (-2,0,1) = -2(1,0,0) && + (0,0,1)\\
\mathbf{T}\mathbf{e}_y &\sim (0,0,2) \;\;= && + 2(0,0,1)\\
\mathbf{T}\mathbf{e}_z &\sim (3,-1,0) = \;\;3(1,0,0) - (0,1,0) &&
\end{aligned}$$

Hence,

$$\mathbf{T} \sim \begin{bmatrix} -2 & 0 & 3 \\ 0 & 0 & -1 \\ 1 & 2 & 0 \end{bmatrix}.$$

PROBLEM 1.6.

If $\mathbf{u} \sim (u_x, u_y, u_z)$, then $\mathbf{u}\times$ may be regarded as a 2nd order tensor whose action on any vector $\mathbf{v} \sim (v_x, v_y, v_z)$ is defined by (1.24). Find the Cartesian components of $\mathbf{u}\times$.

SOLUTION.

Applying $\mathbf{u}\times$, successively, to $\mathbf{e}_x$, $\mathbf{e}_y$, and $\mathbf{e}_z$, we have, according to (1.24),

$$\begin{aligned}
\mathbf{u}\times\mathbf{e}_x &\sim (0,u_z,-u_y) = && u_z(0,1,0) - u_y(0,0,1)\\
\mathbf{u}\times\mathbf{e}_y &\sim (-u_z,0,u_x) = -u_z(1,0,0) && + u_x(0,0,1)\\
\mathbf{u}\times\mathbf{e}_z &\sim (u_y,-u_x,0) = \;\; u_y(1,0,0) - u_x(0,1,0) &&
\end{aligned}$$

Hence

$$\mathbf{u}\times \sim \begin{bmatrix} 0 & -u_z & u_y \\ u_z & 0 & -u_x \\ -u_y & u_x & 0 \end{bmatrix}.$$

The Cartesian Basis for Second Order Tensors

Given the Cartesian components (v_x, v_y, v_z) of a vector $\mathbf{v}$, we recover the vector via the representation $\mathbf{v} = v_x\mathbf{e}_x + v_y\mathbf{e}_y + v_z\mathbf{e}_z$. Given the Cartesian components of a 2nd order tensor $\mathbf{T}$, what is the analogous representation for $\mathbf{T}$? To save space we shall answer, first, in 2-dimensions. From (1.36) to (1.38),

$$\mathbf{T}\mathbf{v} = v_x(T_{xx}\mathbf{e}_x + T_{yx}\mathbf{e}_y) + v_y(T_{xy}\mathbf{e}_x + T_{yy}\mathbf{e}_y). \tag{1.40}$$

But $v_x\mathbf{e}_x = (\mathbf{v}\cdot\mathbf{e}_x)\mathbf{e}_x = \mathbf{e}_x\mathbf{e}_x(\mathbf{v})$, etc. Thus

$$\mathbf{T}\mathbf{v} = (T_{xx}\mathbf{e}_x\mathbf{e}_x + T_{xy}\mathbf{e}_x\mathbf{e}_y + T_{yx}\mathbf{e}_y\mathbf{e}_x + T_{yy}\mathbf{e}_y\mathbf{e}_y)\mathbf{v}, \qquad \forall\mathbf{v}. \tag{1.41}$$

Because $\mathbf{v}$ is arbitrary, we infer from (1.31) that

$$\mathbf{T} = T_{xx}\mathbf{e}_x\mathbf{e}_x + T_{xy}\mathbf{e}_x\mathbf{e}_y + T_{yx}\mathbf{e}_y\mathbf{e}_x + T_{yy}\mathbf{e}_y\mathbf{e}_y. \tag{1.42}$$

Is this representation unique? You betcha. For suppose there is another, say

$$\mathbf{T} = T'_{xx}\mathbf{e}_x\mathbf{e}_x + T'_{xy}\mathbf{e}_x\mathbf{e}_y + T'_{yx}\mathbf{e}_y\mathbf{e}_x + T'_{yy}\mathbf{e}_y\mathbf{e}_y. \tag{1.43}$$

Then (1.42), take away (1.43), implies that

$$[(T_{xx} - T'_{xx})\mathbf{e}_x\mathbf{e}_x + (T_{xy} - T'_{xy})\mathbf{e}_x\mathbf{e}_y + \cdots]\mathbf{v} = \mathbf{0}, \qquad \forall\mathbf{v}, \tag{1.44}$$

which is equivalent to

$$\mathbf{u}\cdot[(T_{xx} - T'_{xx})\mathbf{e}_x\mathbf{e}_x + (T_{xy} - T'_{xy})\mathbf{e}_x\mathbf{e}_y + \cdots]\mathbf{v} = 0, \qquad \forall\mathbf{u},\mathbf{v}. \tag{1.45}$$

But if $\mathbf{u} = \mathbf{v} = \mathbf{e}_x$, (1.45) reduces to $T_{xx} - T'_{xx} = 0$; if $\mathbf{u} = \mathbf{e}_x$, $\mathbf{v} = \mathbf{e}_y$, to $T_{xy} - T'_{xy} = 0$, etc. Thus (1.42) and (1.43) must be identical.

In three dimensions, the conclusion is that, uniquely,

$$\begin{aligned}\mathbf{T} = {} & T_{xx}\mathbf{e}_x\mathbf{e}_x + T_{xy}\mathbf{e}_x\mathbf{e}_y + T_{xz}\mathbf{e}_x\mathbf{e}_z \\ & + T_{yx}\mathbf{e}_y\mathbf{e}_x + T_{yy}\mathbf{e}_y\mathbf{e}_y + T_{yz}\mathbf{e}_y\mathbf{e}_z \\ & + T_{zx}\mathbf{e}_z\mathbf{e}_x + T_{zy}\mathbf{e}_z\mathbf{e}_y + T_{zz}\mathbf{e}_z\mathbf{e}_z.\end{aligned} \tag{1.46}$$

Thus the set $\{\mathbf{e}_x\mathbf{e}_x, \mathbf{e}_x\mathbf{e}_y, \ldots, \mathbf{e}_z\mathbf{e}_z\}$ of the 9 possible direct products of the 3 Cartesian base vectors is a *basis for the set of all 2nd order tensors*. That is, any 2nd order tensor can be represented as a unique linear combination of these 9 dyads.

Exercises

1.1. (a). Using graph paper, draw the five vectors

$$\mathbf{p} \sim (1,1),\ \mathbf{q} \sim (1,2),\ \mathbf{r} \sim (2,1),\ \mathbf{s} \sim (0,-3),\ \mathbf{t} \sim (-1,1)$$

so that each has its tail at the origin.

(b). Using the head-to-tail rule for addition and graph paper, place the tail of $\mathbf{q}$ at the head of $\mathbf{p}$, the tail of $\mathbf{r}$ at the head of $\mathbf{q}$, etc., and indicate the sum of the five vectors as a vector whose tail coincides with that of **p** and whose head coincides with that of **t**.

(c). What are the Cartesian components of the vector **u** such that $\mathbf{p} + \mathbf{q} + \cdots + \mathbf{t} + \mathbf{u} = \mathbf{0}$?

1.2. Draw a careful perspective sketch illustrating that if **u** and **v** and then $(\mathbf{u} + \mathbf{v})$ and **w** are added according to the parallelogram rule, the last sum will be the diagonal of a parallelopiped having **u**,**v**, and **w** as co-terminal edges.

1.3. Describe in words and with a sketch an experiment to demonstrate that forces add like vectors.

1.4. If $\mathbf{u} \sim (2,-3,4)$ and $\mathbf{v} \sim (1,0,1)$, compute $\mathbf{u}\cdot\mathbf{v}$ and the enclosed angle.

1.5. Using the vectors given in Exercise 1.4, compute $\mathrm{Proj}_{\mathbf{u}}\mathbf{v}$ and $\mathrm{Proj}_{\mathbf{v}}\mathbf{u}$.

1.6. Washington, D.C. lies roughly at latitude 39°N and longitude 77°W and Moscow roughly at latitude 56°N and longitude 38°E. Taking the radius of the earth as 4000 miles, compute the great circle distance between these two cities.
Hint: Imagine vectors from the center of the earth to the cities and use the dot product.

1.7. Using (1.11), prove the distributive law (1.14) without introducing Cartesian coordinates.
Hint: Use the fact that if **a** and **b** are co-terminal edges of a parallelogram, then $2|\mathbf{a}|^2 + 2|\mathbf{b}|^2 = |\mathbf{a} + \mathbf{b}|^2 + |\mathbf{a} - \mathbf{b}|^2$. Apply (1.11) to $\mathbf{u}\cdot(\mathbf{v} + \mathbf{w})$ and in the resulting expression set $\mathbf{v} + \mathbf{w} - \mathbf{u} = (\mathbf{v} - \tfrac{1}{2}\mathbf{u}) + (\mathbf{w} - \tfrac{1}{2}\mathbf{u}) \equiv \mathbf{a} + \mathbf{b}$.

1.8. *The Schwarz' inequality*

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

is an example of a simple geometric fact that has a useful analog for functions: if f and g are square integrable on an interval (a,b), then

$$\left|\int_a^b fg\,dx\right| \leq \sqrt{\int_a^b f^2 dx}\,\sqrt{\int_a^b g^2 dx}. \qquad (**)$$

(a). Prove (*) in three ways by using

(i). (1.25) and the fact that

$$|\mathbf{u} \times \mathbf{v}|^2 \geq 0.$$

(ii). the fact that the second degree polynomial

$$p(x) \equiv |\mathbf{u} + x\mathbf{v}|^2$$

(iii). (1.11) and the fact from Euclidean geometry that the length of a side of a triangle is greater than or equal to the difference between the lengths of the other two sides.

(b). Use the analogue of (i) or (ii) for functions to prove (**).

1.9. Let **a** and **b** be given *three-dimensional* vectors and **x** unknown. Without introducing components, show that the unique solution of the linear algebraic equations

$$\mathbf{x} + \mathbf{a} \times \mathbf{x} = \mathbf{b} \qquad (*)$$

is

$$\mathbf{x} = \frac{\mathbf{b} + (\mathbf{a}\cdot\mathbf{b})\mathbf{a} + \mathbf{b} \times \mathbf{a}}{1 + \mathbf{a}\cdot\mathbf{a}}.$$

Hint: Set $\mathbf{x} = A\mathbf{a} + B\mathbf{b} + C\mathbf{a} \times \mathbf{b}$ and solve for A,B,C. To prove uniqueness, note that if $\mathbf{y}$ is another solution of (*), then

$$\mathbf{x} - \mathbf{y} + \mathbf{a} \times (\mathbf{x} - \mathbf{y}) = \mathbf{0}.$$

What can you conclude from this last fact?

1.10. If $\mathbf{a} \sim (-2,1,0)$ and $\mathbf{b} \sim (3,2,1)$,
(a) Compute $\mathbf{a} \times \mathbf{b}$.
(b) find the equation of the plane *spanned* (i.e., determined) by $\mathbf{a}$ and $\mathbf{b}$.

1.11. Given a point $P(x_0,y_0,z_0)$ and a plane

$$\Pi: Ax + By + Cz = D, \qquad (*)$$

derive a formula for the (shortest) distance from P to Π by using
(a). calculus to minimize $(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2$ subject to (*).
(b). vector methods to construct a $\perp$ from P to Π.
Hint: Recall what $\mathbf{N} \sim (A,B,C)$ represents.

1.12. If $\mathbf{u} \sim (1,-1,2)$, $\mathbf{v} \sim (3,2,1)$, and $\mathbf{w} \sim (4,1,7)$, compute
(a). $\mathbf{uv}(\mathbf{w})$ (b). $\mathbf{vu}(\mathbf{w})$ (c). $\mathbf{wv}(\mathbf{u})$

1.13. Show that
(a). $(\mathbf{uv})^T = \mathbf{vu}$
(b). for any vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, $(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} = (\mathbf{ba} - \mathbf{ab})(\mathbf{c})$.

1.14. Using either a calculus or geometric argument (a physical take-apart model is extremely useful here), show that the volume of a tetrahedron with co-terminal edges $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ is $(1/6)|(\mathbf{u} \times \mathbf{v})\cdot\mathbf{w}|$.

1.15. Give a convincing (though not necessarily rigorous) argument to show that $|(\mathbf{u} \times \mathbf{v})\cdot\overline{\mathbf{w}}|$ is the area of the parallelogram with co-terminal edges $\mathbf{u}$ and $\mathbf{v}$ projected onto a plane $\perp$ to $\mathbf{w}$.

1.16. Show that in 3-dimensions,

$$\mathbf{u} = \mathbf{v} \Leftrightarrow \mathbf{u} \times \mathbf{c} = \mathbf{v} \times \mathbf{c}, \ \forall \mathbf{c}.$$

1.17. Let $\mathbf{T}$ be the tensor defined in Problem 1.4.
(a). If $\mathbf{v} \sim (v_x, v_y, v_z)$, fill in the blanks.
$\mathbf{Sv} \equiv 1/2(\mathbf{T} + \mathbf{T}^T)\mathbf{v} \sim (\_\_, \_\_, \_\_)$
$\mathbf{Av} \equiv 1/2(\mathbf{T} - \mathbf{T}^T)\mathbf{v} \sim (\_\_, \_\_, \_\_)$
(b). Determine the matrices of the Cartesian components of $\mathbf{S}$ and $\mathbf{A}$.
(c). Find the Cartesian components of the vector $\boldsymbol{\omega}$ such that $\mathbf{A} = \boldsymbol{\omega} \times$.

1.18. Let $\mathbf{A}$ be an arbitrary, 3-dimensional skew tensor.
(a). By expressing $\mathbf{A}$ in terms of its Cartesian components (and noting that only 3 of these can be assigned arbitrarily since $\mathbf{A} = -\mathbf{A}^T$), find a vector $\boldsymbol{\omega}$ such that

$$\mathbf{Av} = \boldsymbol{\omega} \times \mathbf{v}, \ \forall \mathbf{v}.$$

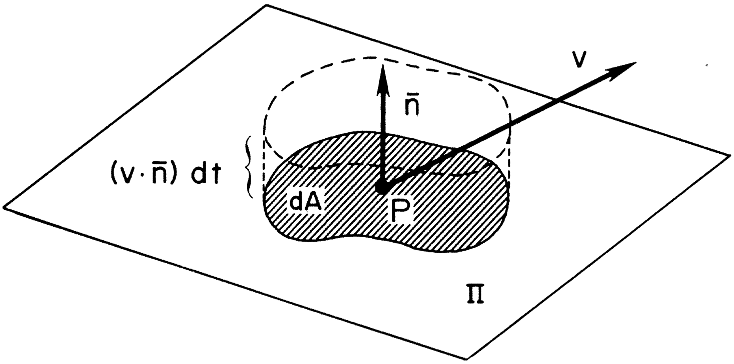


Figure 1.9

(b). Use the results of Exercise 1.16 to show that $\boldsymbol{\omega}$ is unique. $\boldsymbol{\omega}$ is called the *axis* of $\mathbf{A}$ and is important in rigid body dynamics. See Exercise 4.20.
(c). Show that $\mathbf{v}\cdot\mathbf{A}\mathbf{v} = 0$, $\forall\mathbf{v}$, in *any* number of dimensions.
(d). Show that $\mathbf{A}\boldsymbol{\omega} = \mathbf{0}$. Check this result by using the numerical values obtained in Exercise 1.17(c).

1.19. Let ρ and $\mathbf{v}$ denote, respectively, the density and velocity of a fluid at a given point P in space at a given time t. If Π is a plane with normal $\mathbf{n}$ passing through P, then the *momentum flux* across Π at P and t is defined to be $\rho\mathbf{v}(\mathbf{v}\cdot\bar{\mathbf{n}}) = \rho\mathbf{v}\mathbf{v}(\bar{\mathbf{n}})$. As indicated in Fig. 1.9, $\rho\mathbf{v}(\mathbf{v}\cdot\bar{\mathbf{n}})dAdt$ is the momentum at P and t carried across an *oriented* differential element of area $\bar{\mathbf{n}}dA$ in time dt. $\rho\mathbf{v}\mathbf{v}$ is called the *momentum flux tensor* at P and t.
(a). If $\mathbf{v} \sim (v_x, v_y, v_z)$, determine the Cartesian components of $\rho\mathbf{v}\mathbf{v}$.
(b). If $\mathbf{v} \sim (3,-1,2)$ and $\rho = 4$ at a given point and time, determine the momentum flux across the plane with normal $\mathbf{n} \sim (-1,1,3)$.

1.20. As in Fig. 1.10, let $\bar{\mathbf{n}}dA$ denote an oriented differential element of area at a point P and time t in a continuum (e.g. a fluid or solid) and let $\mathbf{t}dA$ denote the force that the material into which $\mathbf{n}$ points exerts across dA. $\mathbf{t}$ is called the *stress* at P and t in the direction $\bar{\mathbf{n}}$; $\mathbf{t}_n \equiv \mathrm{Proj}_{\mathbf{n}}\mathbf{t}$ the *normal* stress, and $\mathbf{t}_s \equiv \mathbf{t} - \mathbf{t}_n$ the *shear* stress. By considering the equations of motion of a tetrahedron of the material of arbitrarily small volume, instantaneously centered at P, it can be shown that $\mathbf{t} = \mathbf{T}\bar{\mathbf{n}}$, where $\mathbf{T} = \mathbf{T}^T$ is the (Cauchy) *stress tensor* at P and t.[21] If

$$\mathbf{T} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & -1 \\ 0 & -1 & 3 \end{bmatrix}$$

and $\mathbf{n} \sim (1,2,-1)$, compute the normal and shear stress.

1.21. The *directions of principal stress* are defined as those unit vectors $\mathbf{x}$ such that $\mathbf{T}\mathbf{x} = \lambda\mathbf{x}$. Note that $\lambda\mathbf{x}$ is just the normal stress and that the shear stress is zero in a principal direction. The determination of all possible values of λ and the associ-

[21] The set of all unit vectors is *not* a linear vector space (Why)? and so is not a suitable domain for $\mathbf{T}$. However, the definition and domain of the stress tensor can be extended in an obvious way, as suggested by Noll: $\mathbf{T0} \equiv \mathbf{0}$, $\mathbf{T}\mathbf{v} \equiv |\mathbf{v}|\mathbf{T}\bar{\mathbf{v}}$, $\forall\mathbf{v} \neq \mathbf{0}$.

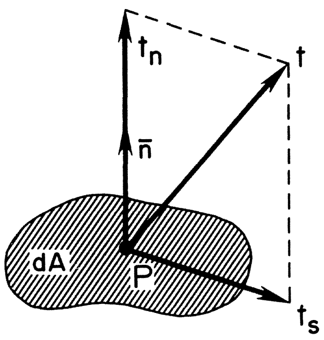


Figure 1.10

ated directions **x** is an *eigenvalue problem*. You have seen such problems in linear algebra. Find the *eigenvalues* λ and the associated *eigenvectors* **x** for the stress tensor **T** of the preceding exercise.

1.22. (a). Determine the matrix of the Cartesian components of the direct products **uv**, **uw**, and **vw**, where **u**, **v**, and **w** are given in Exercise 1.12.
(b). Compute det(**uv**), det(**uw**), det(**vw**).
(c). Show that the determinant of the direct product of *any* two vectors is zero.

1.23. The *trace* of the direct product is denoted and defined by

$$\mathrm{tr}(\mathbf{uv}) = \mathbf{u}\cdot\mathbf{v},\ \mathrm{tr}(\mathbf{uv} + \mathbf{wx}) = \mathrm{tr}(\mathbf{uv}) + \mathrm{tr}(\mathbf{wx}).$$

Use this definition and (1.46) to find an expression for tr**T** in terms of its Cartesian components.

1.24. *The derivative of a second order tensor* **T**, that depends on a single parameter *t*, may be defined by first applying the product rule *formally*:

$$(\mathbf{Tv})^{\cdot} = \mathbf{T}\dot{\mathbf{v}} + \dot{\mathbf{T}}\mathbf{v}. \tag{1.47}$$

But vector differentiation is well-defined. Thus $(\mathbf{Tv})^{\cdot}$ and $\mathbf{T}\dot{\mathbf{v}}$ make sense (provided, of course, that the vectors **Tv** and **v** are differentiable). We therefore *define* $\dot{\mathbf{T}}$ by the rule

$$\dot{\mathbf{T}}\mathbf{v} \equiv (\mathbf{Tv})^{\cdot} - \mathbf{T}\dot{\mathbf{v}},\ \forall\mathbf{v}. \tag{1.48}$$

(a). By definition, **T** sends vectors into vectors. Show that **T** is also linear and hence a 2nd order tensor.
(b). If $\mathbf{v} \sim (v_x, v_y, v_z)$, and $\mathbf{Tv} \sim (-2t^2 v_x + 3t^3 v_z,\ \cos \pi t v_z,\ v_x + 2tv_y + \sqrt{1 + t^2 v_z})$, fill in the blanks: $\mathbf{Tv} \sim$ (________, ________, ________).
(c). If $\mathbf{T} = \mathbf{xy}$, where **x** and **y** are functions of *t*, what is $\dot{\mathbf{T}}$?
(d). Given the Cartesian components of **T**, how do we obtain those of $\dot{\mathbf{T}}$?

1.25. *The composition or dot product of two second order tensors* **S** and **T** is a 2nd order tensor denoted by **S·T** and defined by

$$\mathbf{S}\cdot\mathbf{Tv} = \mathbf{S}(\mathbf{Tv}),\ \forall\mathbf{v}$$

(a). If $\mathbf{S} = \mathbf{wx}$ and $\mathbf{T} = \mathbf{yz}$, what is **S·T**?

(b). If $\mathbf{S}$ and $\mathbf{T}$ are differentiable functions of t, show that

$$(\mathbf{S}\cdot\mathbf{T})^{\cdot} = \mathbf{S}\cdot\dot{\mathbf{T}} + \dot{\mathbf{S}}\cdot\mathbf{T} \tag{1.49}$$

1.26. With reference to Exercise 1.18a, show what

$$\mathbf{A}^T\cdot\mathbf{A} = (\boldsymbol{\omega}\cdot\boldsymbol{\omega})\mathbf{1} - \boldsymbol{\omega}\boldsymbol{\omega}$$

and hence that

$$|\boldsymbol{\omega}|^2 = \tfrac{1}{2}\,\mathrm{tr}\mathbf{A}^T\cdot\mathbf{A}.$$

1.27. *The inverse of a second order tensor.* A function f is said to be one-to-one (1 : 1) if $f(x_1) = f(x_2) \Rightarrow x_1 = x_2$. In this case f has a unique inverse f^{-1} such that $f^{-1}(f(x)) = x$, for all x in the domain of f. Moreover $f(f^{-1}(y)) = y$ for all y in the range of f. If $\mathbf{T}$ is a 2nd order tensor, then $\mathbf{T}\mathbf{v}_1 = \mathbf{T}\mathbf{v}_2$ implies that $\mathbf{T}(\mathbf{v}_1 - \mathbf{v}_2) = \mathbf{0}$, by linearity. If $\mathbf{T}$ is non-singular, this, in turn, implies that $\mathbf{v}_1 = \mathbf{v}_2$. That is, $\mathbf{T}$ has a unique inverse $\mathbf{T}^{-1}$ such that

$$\begin{aligned} \mathbf{T}^{-1}\cdot\mathbf{T}\mathbf{v} &= \mathbf{v},\ \forall\mathbf{v} \\ \mathbf{T}\cdot\mathbf{T}^{-1}\mathbf{w} &= \mathbf{w},\ \forall\mathbf{w}. \end{aligned} \tag{1.50}$$

(a). Show that $\mathbf{T}^{-1}$ is linear (and hence a 2nd order tensor as the notation presumes.)

(b) Show that $(\mathbf{T}^{-1})^T = (\mathbf{T}^T)^{-1}$.

(c). Show that the tensor defined in Problem 1.4 is non-singular.

(d) If $\mathbf{w} \sim (w_x, w_y, w_z)$, and $\mathbf{T}$ is the tensor in Problem 1.4, fill in the blanks: $\mathbf{T}^{-1}\mathbf{w} \sim$ (—, —, —).
Hint: Let $\mathbf{T}^{-1}\mathbf{w} \sim (a,b,c)$ and note that the action of $\mathbf{T}$ on $\mathbf{T}^{-1}\mathbf{w}$ must equal $\mathbf{w}$.

CHAPTER II

General Bases and Tensor Notation

While the laws of mechanics can be written in coordinate-free form, they can be solved, in most cases, only if expressed in component form. This requires that we introduce a basis. Though the standard Cartesian basis is often the simplest, the physics and geometry of a problem, and especially the so-called boundary conditions, may dictate another. For example, if we wished to study the temperature distribution in a body the shape of a parallelopiped, we would choose most likely a basis consisting of vectors lying along three co-terminal edges of the body. An aim of tensor analysis is to embrace arbitrary coordinate systems and their associated bases, yet to produce formulas for computing invariants, such as the dot product, that are as simple as the Cartesian forms.

General Bases

Let $\{\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3\}$ denote any *fixed* set of *noncoplanar* vectors. Then any vector $\mathbf{v}$ may be represented uniquely as

$$\mathbf{v} = v^1\mathbf{g}_1 + v^2\mathbf{g}_2 + v^3\mathbf{g}_3 = \sum_1^3 v^i\mathbf{g}_i. \tag{2.1}$$

Used thus, the set $\{\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3\}$ is called a *basis*, and its elements, *base vectors. The base vectors need not be of unit length nor mutually* $\perp$. A basis is illustrated in Fig. 2.1.

To understand the representation (2.1) geometrically, we first consider the 2-dimensional case. In Fig. 2.2 we have drawn a typical vector $\mathbf{v}$ and a basis $\{\mathbf{g}_1,\mathbf{g}_2\}$, i.e., two fixed, non-zero, non-parallel vectors. Through the head of

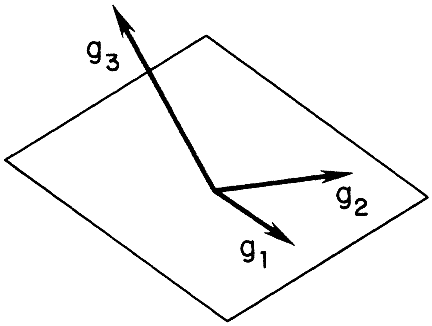


Figure 2.1

v we draw a line parallel to $\mathbf{g}_2$. It will intersect the line along $\mathbf{g}_1$ at some point P. The vector of $\overline{OP}$ is then some unique scalar v^1 times $\mathbf{g}_1$. (Thus v^1 might be negative though we have drawn the figure as if it were positive.) We construct v^2 in a similar way.

In 3-dimensions we are given a vector **v** and a basis $\{\mathbf{g}_1,\mathbf{g}_2,\mathbf{g}_3\}$, all vectors having a common tail O. A plane passed through the head of **v** parallel to the plane of $\mathbf{g}_2$ and $\mathbf{g}_3$ will intersect the line along $\mathbf{g}_1$ in some point P. This determines a unique scalar v^1 such that $v^1\mathbf{g}_1$ is the vector from O to P. In like fashion we determine v^2 and v^3.

The representation (2.1) is effected analytically as in the following.

PROBLEM 2.1.

Given

$$\mathbf{g}_1 \sim (1,-1,2),\ \mathbf{g}_2 \sim (0,1,1),\ \mathbf{g}_3 \sim (-1,-2,1)$$

$$\mathbf{v} \sim (3,3,6),$$

find (v^1,v^2,v^3).

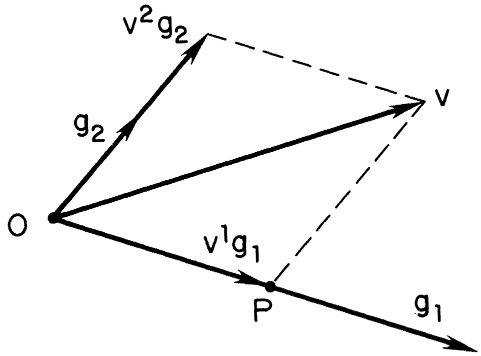


Figure 2.2

SOLUTION.

Equating corresponding Cartesian components on both sides of (2.1), i.e. applying rules (1.5)–(1.7), we get

$$\begin{aligned} 3 &= v^1 \qquad\quad - v^3 \\ 3 &= -v^1 + v^2 - 2v^3 \\ 6 &= 2v^1 + v^2 + v^3. \end{aligned}$$

Solving these simultaneous linear algebraic equations, we have

$$(v^1,v^2,v^3) = (2,3,-1).$$

The Jacobian of a Basis Is Nonzero

Unless you are good at perspective, it may be difficult to see from a sketch whether three vectors form a basis. (In higher dimensions it is hopeless.) Is there a numerical test for a set of vectors to be a basis? Yes, providing we know their Cartesian components. As Problem 2.1 illustrates, the answer hinges on whether a set of n simultaneous linear algebraic equations in n unknowns has a unique solution. It does, as you well know, if and only if the determinant of coefficients does not vanish. Let us rephrase this fact as follows. If $G = [\mathbf{g}_1,\mathbf{g}_2, \ldots]$ denotes the $n \times n$ matrix whose columns are the Cartesian components of $\mathbf{g}_1,\mathbf{g}_2, \ldots$, then $\{\mathbf{g}_1,\mathbf{g}_2, \ldots\}$ is a basis if and only if $\det G \neq 0$. Using almost standard terminology, we shall call G the *Jacobian matrix* of $\{\mathbf{g}_1,\mathbf{g}_2, \ldots\}$ and $J \equiv \det G$ the *Jacobian* of $\{\mathbf{g}_1,\mathbf{g}_2, \ldots\}$.[1]

The Summation Convention

The summation convention, invented by Einstein, gives tensor analysis much of its appeal. Observe in (2.1) that the *dummy index* of summation i is *repeated*. Moreover, its range, 1 to 3, is already known from the context of the discussion. Therefore, without any loss of information, we may drop the summation symbol in (2.1) and write, simply,

$$\mathbf{v} = v^i\mathbf{g}_i. \tag{2.2}$$

"Dummy" means that the symbol i in (2.1) or (2.2) may be replaced by any other symbol without affecting the value of the sum. Thus $v^i\mathbf{g}_i = v^j\mathbf{g}_j = v^k\mathbf{g}_k$, etc. Replacing one dummy index by another is one of the first tricks a would-be index slinger must learn.

We attach, however, one useful proviso: *the summation convention applies only when one dummy index is "on the roof" and the other is "in the*

[1] The label "Jacobian" is usually applied when the components of $\mathbf{g}_1,\mathbf{g}_2, \ldots$ are partial derivatives, a situation we shall meet in Chapter III.

cellar". Thus $v^i v_i = v^1 v_1 + v^2 v_2 + v^3 v_3$, but $v^i v^i = v^1 v^1$ *or* $v^2 v^2$ *or* $v^3 v^3$. Generally, repeated roof or cellar indices occur in quantities that are *not* invariants or *not* the components of invariants. (*Cartesian tensor notation,* discussed in Exercise 2.21, is the one exception to this proviso.)

Computing the Dot Product in a General Basis

Suppose we wish to compute the dot product of a vector $\mathbf{u} = u^i \mathbf{g}_i$ with a vector $\mathbf{v} = v^i \mathbf{g}_i$. In doing so we must replace one pair of dummy indices, say the second, by another. Otherwise, blind application of the summation convention yields, incorrectly, $\mathbf{u}\cdot\mathbf{v} = u^i v^i \mathbf{g}_i \cdot \mathbf{g}_i = u^1 v^1 \mathbf{g}_1\cdot\mathbf{g}_1 + u^2 v^2 \mathbf{g}_2\cdot\mathbf{g}_2 + u^3 v^3 \mathbf{g}_3\cdot\mathbf{g}_3$. The correct expression is

$$\begin{aligned}\mathbf{u}\cdot\mathbf{v} &= u^i v^j \mathbf{g}_i \cdot \mathbf{g}_j \\ &= u^1 v^1 \mathbf{g}_1\cdot\mathbf{g}_1 + u^1 v^2 \mathbf{g}_1\cdot\mathbf{g}_2 + u^1 v^3 \mathbf{g}_1\cdot\mathbf{g}_3 \\ &\quad + u^2 v^1 \mathbf{g}_2\cdot\mathbf{g}_1 + \cdots + u^3 v^3 \mathbf{g}_3\cdot\mathbf{g}_3. \end{aligned} \tag{2.3}$$

This *extended expression* for $\mathbf{u}\cdot\mathbf{v}$ is a nine-term mess. We can clean it up by introducing a set of reciprocal base vectors.

Reciprocal Base Vectors

What these are and how they simplify things is seen most easily in 2-dimensions. Given a basis $\{\mathbf{g}_1,\mathbf{g}_2\}$, we represent, as before, the first factor in the dot product $\mathbf{u}\cdot\mathbf{v}$ as $\mathbf{u} = u^1 \mathbf{g}_1 + u^2 \mathbf{g}_2$. (In this two-dimensional example, formulas are short enough that we do not need the summation convention.) However, let us represent the second factor in the dot product in terms of some new (and as yet unknown) basis $\{\mathbf{g}^1,\mathbf{g}^2\}$, writing $\mathbf{v} = v_1 \mathbf{g}^1 + v_2 \mathbf{g}^2$. Then

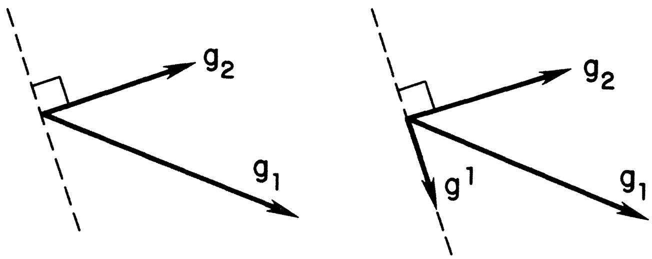
$$\begin{aligned}\mathbf{u}\cdot\mathbf{v} &= (u^1 \mathbf{g}_1 + u^2 \mathbf{g}_2)\cdot(v_1 \mathbf{g}^1 + v_2 \mathbf{g}^2) \\ &= u^1 v_1 \mathbf{g}_1\cdot\mathbf{g}^1 + u^1 v_2 \mathbf{g}_1\cdot\mathbf{g}^2 + u^2 v_1 \mathbf{g}_2\cdot\mathbf{g}^1 + u^2 v_2 \mathbf{g}_2\cdot\mathbf{g}^2.\end{aligned}$$

The idea is now to choose $\mathbf{g}^1$ and $\mathbf{g}^2$ so that the above expression reduces to

$$\mathbf{u}\cdot\mathbf{v} = u^1 v_1 + u^2 v_2.$$

Thus we require that $\mathbf{g}_1\cdot\mathbf{g}^1 = \mathbf{g}_2\cdot\mathbf{g}^2 = 1$ and $\mathbf{g}_1\cdot\mathbf{g}^2 = \mathbf{g}_2\cdot\mathbf{g}^1 = 0$. Given $\mathbf{g}_1$ and $\mathbf{g}_2$, we may construct $\mathbf{g}^1$ and $\mathbf{g}^2$ geometrically as follows:
As in Fig. 2.3a, draw a line $\perp$ to $\mathbf{g}_2$. Since $\mathbf{g}^1\cdot\mathbf{g}_2 = 0$, $\mathbf{g}^1$ must lie along this line. Now adjust the direction and length of $\mathbf{g}^1$ until $\mathbf{g}^1\cdot\mathbf{g}_1 = 1$, as indicated in Fig. 2.3b. Repeat to construct $\mathbf{g}^2$.

With the geometry well-understood (especially after the completion of Exercise 2.22), we now develop an algebraic method (that a computer can use) of constructing reciprocal base vectors in n-dimensions. This will take us on



(a) Figure 2.3 (b)

a small detour before returning to the problem of simplifying the component form of (2.3).

Let $\{\mathbf{g}_1, \mathbf{g}_2, \ldots\} \equiv \{\mathbf{g}_j\}$ be a basis. Then, det $G \neq 0$. As you recall from matrix theory, this implies that G^{-1} exists. The elements in the ith row of G^{-1} may be regarded as the Cartesian components of a vector $\mathbf{g}^i$, i.e.,

$$G^{-1} = \begin{bmatrix} \mathbf{g}^1 \\ \mathbf{g}^2 \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} = [\mathbf{g}^1, \mathbf{g}^2, \ldots]^T \equiv [\mathbf{g}^i],$$

where T denotes transpose. (Consistent with this notation we may set $G \equiv [\mathbf{g}_j]$ when we wish to regard G as a collection of column vectors). The law of matrix multiplication says that the element in the ith row and jth column of $G^{-1}G$ is the sum of the products of corresponding elements in the ith row of G^{-1} and the jth column of G. Thus, $G^{-1}G = I$ is equivalent to the statement

$$\mathbf{g}^i \cdot \mathbf{g}_j = \delta^i_j \equiv \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}. \tag{2.4}$$

The symbol δ^i_j is called the *Kronecker delta*. It is ubiquitous in tensor analysis. The set $\{\mathbf{g}^1, \mathbf{g}^2, \ldots\} \equiv \{\mathbf{g}^i\}$ is called a *reciprocal basis* and its elements, *reciprocal base vectors*. The name basis is justified because det $G^{-1} \neq 0$.

PROBLEM 2.2.

Find the reciprocal base vectors for the basis given in Problem 2.1.

SOLUTION.

We must compute G^{-1} and then read off the Cartesian components of $\mathbf{g}^1$, $\mathbf{g}^2$, and $\mathbf{g}^3$. A systematic way of computing G^{-1} is to reduce G to I by a sequence of elementary row operations (involving, possibly, row interchanges). This same sequence of operations applied to I will produce G^{-1}. We adjoin I to G and carry out these operations simultaneously, as follows.

$$\begin{bmatrix} 1 & 0 & -1 & | & 1 & 0 & 0 \\ -1 & 1 & -2 & | & 0 & 1 & 0 \\ 2 & 1 & 1 & | & 0 & 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -1 & | & 1 & 0 & 0 \\ 0 & 1 & -3 & | & 1 & 1 & 0 \\ 0 & 1 & 3 & | & -2 & 0 & 1 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 1 & 0 & -1 & | & 1 & 0 & 0 \\ 0 & 1 & -3 & | & 1 & 0 & 0 \\ 0 & 0 & 6 & | & -3 & -1 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 & | & \frac{1}{2} & -\frac{1}{6} & \frac{1}{6} \\ 0 & 1 & 0 & | & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 & | & -\frac{1}{2} & -\frac{1}{6} & \frac{1}{6} \end{bmatrix}.$$

Thus $\mathbf{g}^1 \sim (1/6)(3,-1,1)$, $\mathbf{g}^2 \sim (1/2)(-1,1,1)$, $\mathbf{g}^3 \sim (1/6)(-3,-1,1)$.

The Roof (Contravariant) and Cellar (Covariant) Components of a Vector

If $\{\mathbf{g}_i\}$ is a basis then not only may we express any vector $\mathbf{v}$ as $v^i\mathbf{g}_i$, we may also represent $\mathbf{v}$ as a linear combination of the reciprocal base vectors, thus:

$$\mathbf{v} = v_i\mathbf{g}^i. \tag{2.5}$$

Breaking with tradition, we shall call the coefficients v^i the *roof components* of $\mathbf{v}$ and the $\mathbf{g}_i$ the *cellar base vectors*. Likewise, in (2.5), the v_i shall be called the *cellar components* of $\mathbf{v}$ and the $\mathbf{g}^i$ the *roof base vectors*. The conventional names for v^i and v_i are the "contravariant" and "covariant" components of $\mathbf{v}$—names that seem to me awkward and meaningless. "Roof" and "cellar" also have mnemonic value in matrix theory where A^i_j is sometimes used to denote the element of a matrix A that sits in the ith row and jth column. The following diagram indicates how to remember which is the row and which is the column index:

$$A^{R(\text{ow})}_{C(\text{olumn})} \leftrightarrow A^R_C \leftrightarrow A^{R(\text{oof})}_{C(\text{ellar})}$$

PROBLEM 2.3.

Compute the cellar components of the vector $\mathbf{v}$ given in Problem 2.1.

SOLUTION.

Writing (2.5) in extended form and noting (2.4), we have

$$\begin{aligned} \mathbf{v}\cdot\mathbf{g}_1 &= (v_1\mathbf{g}^1 + v_2\mathbf{g}^2 + v_3\mathbf{g}^3)\cdot\mathbf{g}_1 \\ &= v_1\mathbf{g}^1\cdot\mathbf{g}_1 + v_2\mathbf{g}^2\cdot\mathbf{g}_1 + v_3\mathbf{g}^3\cdot\mathbf{g}_1 \\ &= v_1. \end{aligned}$$

But from the information given in Problem 2.1, $\mathbf{v}\cdot\mathbf{g}_1 = (3)(1) + (3)(-1) + (6)(2) = 12$. Likewise,

$$\begin{aligned} v_2 &= \mathbf{v}\cdot\mathbf{g}_2 = (3)(0) + (3)(1) + (6)(1) = 9 \\ v_3 &= \mathbf{v}\cdot\mathbf{g}_3 = (3)(-1) + (3)(-2) + (6)(1) = -3. \end{aligned}$$

The solution of Problem 2.3 illustrates a useful fact:

$$v_i = \mathbf{v}\cdot\mathbf{g}_i. \tag{2.6}$$

One would also hope that

$$v^i = \mathbf{v}\cdot\mathbf{g}^i. \tag{2.7}$$

To get some practice with the Kronecker delta, let us establish (2.7) starting from (2.2) and (2.4), but without writing out anything in extended form. (After all, tensor notation is designed to keep things compressed.) Taking the dot product of both sides of (2.2) with $\mathbf{g}^j$ we have

$$\begin{aligned} \mathbf{v}\cdot\mathbf{g}^j &= (v^i\mathbf{g}_i)\cdot\mathbf{g}^j \\ &= v^i\mathbf{g}^j\cdot\mathbf{g}_i \\ &= v^i\delta_i^j \\ &= v^j. \qquad \square \end{aligned} \tag{2.8}$$

Let's take this derivation one line at a time. The second follows from the first because $(\alpha\mathbf{u} + \beta\mathbf{v} + \ldots)\cdot\mathbf{w} = \alpha\mathbf{u}\cdot\mathbf{w} + \beta\mathbf{v}\cdot\mathbf{w} + \ldots$. The third line follows because (2.4) obviously holds if i and j on both sides are replaced by any other distinct symbols, say j and i. You've seen such renaming with functions where, for example, $f(x,y) = x/y$ implies that $f(y,x) = y/x$.

How is the last line in (2.8) obtained? Consider one of the possible values that the *free index* j may assume in the preceding line, say 2. Then, summing over the repeated index i, we have $\mathbf{v}\cdot\mathbf{g}^2 = v^1\delta_1^2 + v^2\delta_2^2 + v^3\delta_3^2$. But $\delta_1^2 = \delta_3^2 = 0$ while $\delta_2^2 = 1$. Thus the sum collapses to simply v^2. The step from the next to last to the last line of (2.8) shows that the Kronecker delta may be regarded as a *replacement operator*. That is, multiplying v^i by δ_i^j replaces the index i on v by the index j. Finally, note that $\mathbf{v}\cdot\mathbf{g}^j = v^j$ and $v^i = \mathbf{v}\cdot\mathbf{g}^i$ are equivalent statements, just as $f(x) \equiv x^2$ and $z^2 \equiv f(z)$ are.

Simplification of the Component Form of the Dot Product in a General Basis

Let us return to the problem of simplifying the extended component form of $\mathbf{u}\cdot\mathbf{v}$. Referring both vectors to the same basis led to an explosion of terms when we wrote (2.3) in extended form. Instead, let us set $\mathbf{u} = u^i\mathbf{g}_i$ but $\mathbf{v} = v_j\mathbf{g}^j$. Then

$$\mathbf{u}\cdot\mathbf{v} = u^iv_j\mathbf{g}_i\cdot\mathbf{g}^j = u^iv_j\delta_i^j = u^iv_i = u^1v_1 + u^2v_2 + u^3v_3, \tag{2.9}$$

a formula no more complicated than that involving the Cartesian components of $\mathbf{u}$ and $\mathbf{v}$. An alternate way to compute the dot product is to set $\mathbf{u} = u_i\mathbf{g}^i$ and $\mathbf{v} = v^j\mathbf{g}_j$. Then

$$\mathbf{u}\cdot\mathbf{v} = u_iv^i. \tag{2.10}$$

PROBLEM 2.4.

Given $\mathbf{u} = 2\mathbf{g}_1 - \mathbf{g}_2 + 4\mathbf{g}_3$ and $\mathbf{w} = -3\mathbf{g}^1 + 2\mathbf{g}^2 - 2\mathbf{g}^3$, compute $\mathbf{u}\cdot\mathbf{v}$, $\mathbf{w}\cdot\mathbf{v}$, and $\mathbf{u}\cdot\mathbf{w}$, where $\mathbf{v}$ and the cellar basis vectors are given in Problem 2.1 and the roof basis vectors are given in the Solution to Problem 2.2.

SOLUTION.

The roof components of $\mathbf{u}$ are given, while the cellar components of $\mathbf{v}$ are given in the solution to Problem 2.3. Thus, by (2.9),

$$\mathbf{u}\cdot\mathbf{v} = (2)(12) + (-1)(9) + (4)(-3) = 3.$$

The cellar components of $\mathbf{w}$ are given while the roof components of $\mathbf{v}$ are given in the Solution to Problem 2.1. Thus, by (2.10),

$$\mathbf{w}\cdot\mathbf{v} = (-3)(2) + (2)(3) + (-2)(-1) = 2.$$

Finally, from (2.9)

$$\mathbf{u}\cdot\mathbf{w} = (2)(-3) + (-1)(2) + (4)(-2) = -16.$$

Computing the Cross Product in a General Basis

Sometimes it is convenient to denote the roof and cellar components of a vector $\mathbf{v}$ by $(\mathbf{v})^i$ and $(\mathbf{v})_i$, respectively. With this notation,

$$\mathbf{u} \times \mathbf{v} = (\mathbf{u} \times \mathbf{v})_k \mathbf{g}^k. \tag{2.11}$$

To compute the cellar components of $\mathbf{u} \times \mathbf{v}$ we set $\mathbf{u} = u^i\mathbf{g}_i$ and $v = v^j\mathbf{g}_j$, so obtaining

$$\begin{aligned}(\mathbf{u} \times \mathbf{v})_k &= (\mathbf{u} \times \mathbf{v})\cdot\mathbf{g}_k \\ &= u^i v^j(\mathbf{g}_i \times \mathbf{g}_j)\cdot\mathbf{g}_k \\ &\equiv u^i v^j \epsilon_{ijk}.\end{aligned} \tag{2.12}$$

The $3 \times 3 \times 3 = 27$ symbols ϵ_{ijk} are called the *cellar components of the permutation tensor* (See Exercise 2.13). Let us examine their properties. From (2.12) and (1.26),

$$\epsilon_{132} = (\mathbf{g}_1 \times \mathbf{g}_2)\cdot\mathbf{g}_3 = J, \text{ the Jacobian.}$$

If we interchange two indices there is a change in sign, e.g., switching 2 and 3 above yields

$$\epsilon_{132} = (\mathbf{g}_1 \times \mathbf{g}_3)\cdot\mathbf{g}_2 = (\mathbf{g}_2 \times \mathbf{g}_1)\cdot\mathbf{g}_3 = -(\mathbf{g}_1 \times \mathbf{g}_2)\cdot\mathbf{g}_3 = -J,$$

and switching 1 and 2 in this expression yields

$$\epsilon_{231} = (\mathbf{g}_2 \times \mathbf{g}_3)\cdot\mathbf{g}_1 = (\mathbf{g}_1 \times \mathbf{g}_2)\cdot\mathbf{g}_3 = J.$$

Moreover, if two or more indices are equal, $\epsilon_{ijk} = 0$ because we can always permute the vector triple product $(\mathbf{g}_i \times \mathbf{g}_j)\cdot\mathbf{g}_k$ so that the first term is the cross product of a vector with itself.

In summary,

$$\epsilon_{ijk} = \begin{cases} +J & \text{if } (i,j,k) \text{ is an even permutation of } (1,2,3) \\ -J & \text{if } (i,j,k) \text{ is an odd permutation of } (1,2,3) \\ 0 & \text{if two or more indices are equal.} \end{cases} \tag{2.13}$$

Thus, returning to (2.11), we have

$$\mathbf{u} \times \mathbf{v} = \epsilon_{ijk} u^i v^j \mathbf{g}^k. \tag{2.14}$$

PROBLEM 2.5.

Given $\mathbf{u} = 2\mathbf{g}_1 - \mathbf{g}_2 + 4\mathbf{g}_3$, $\mathbf{v} = 2\mathbf{g}_1 + 3\mathbf{g}_2 - \mathbf{g}_3$, where the basis $\{\mathbf{g}_i\}$ is given in Problem 2.1, compute the cellar components of $\mathbf{u} \times \mathbf{v}$.

SOLUTION.

Let us write (2.12) in extended form for $k = 1$, taking note of (2.14).

$$\begin{aligned}(\mathbf{u} \times \mathbf{v})_1 = \epsilon_{ij1} u^i v^j &= \cancelto{0}{\epsilon_{111}} u^1 v^1 + \cancelto{0}{\epsilon_{121}} u^1 v^2 + \cancelto{0}{\epsilon_{131}} u^1 v^3 \\ &+ \cancelto{0}{\epsilon_{211}} u^2 v^1 + \cancelto{0}{\epsilon_{221}} u^2 v^2 + \cancelto{J}{\epsilon_{231}} u^2 v^3 \\ &+ \cancelto{0}{\epsilon_{311}} u^3 v^1 + \cancelto{-J}{\epsilon_{321}} u^3 v^2 + \cancelto{0}{\epsilon_{331}} u^3 v^3 \\ &= J(u^2 v^3 - u^3 v^2).\end{aligned}$$

The roof components of $\mathbf{u}$ and $\mathbf{v}$ are given, and from Problem 2.1,

$$J = \begin{vmatrix} 1 & 0 & -1 \\ -1 & 1 & -2 \\ 2 & 1 & 1 \end{vmatrix} = 6.$$

Hence

$$(\mathbf{u} \times \mathbf{v})_1 = 6[(-1)(-1) - (4)(3)] = -66.$$

In like fashion, we find that for $k = 2$ and $k = 3$,

$$\begin{aligned}(\mathbf{u} \times \mathbf{v})_2 &= J(u^3 v^1 - u^1 v^3) = 6[(4)(2) - (2)(-1)] = 60 \\ (\mathbf{u} \times \mathbf{v})_3 &= J(u^1 v^2 - u^2 v^1) = 6[(2)(3) - (-1)(2)] = 48.\end{aligned}$$

To compute the roof components of $\mathbf{u} \times \mathbf{v}$, we mimic the above procedure, but use roof instead of cellar base vectors. We obtain

$$(\mathbf{u} \times \mathbf{v})^k = \epsilon^{ijk} u_i v_j, \tag{2.15}$$

where u_i and v_j are the cellar components of $\mathbf{u}$ and $\mathbf{v}$ and the ϵ^{ijk} are the 27 *roof components of the permutation tensor* defined as follows:

$$\epsilon^{ijk} \equiv (\mathbf{g}^i \times \mathbf{g}^j)\cdot\mathbf{g}^k = \begin{cases} +J^{-1} & \text{if } (i,j,k) \text{ is an even permutation of } (1,2,3) \\ -J^{-1} & \text{if } (i,j,k) \text{ is an odd permutation of } (1,2,3) \\ 0 & \text{if two or more indices are equal.} \end{cases} \quad (2.16)$$

In other words, $\epsilon_{ijk} = J^2\epsilon^{ijk}$.

Here is an interesting identity connecting the components of the permutation tensor and the Kronecker delta that you are asked to establish in Exercise 2.5:

$$\epsilon^{ijk}\epsilon_{pqr} \equiv \begin{vmatrix} \delta^i_p & \delta^i_q & \delta^i_r \\ \delta^j_p & \delta^j_q & \delta^j_r \\ \delta^k_p & \delta^k_q & \delta^k_r \end{vmatrix}. \quad (2.17)$$

If we expand the determinant and set $r = k$, we obtain

$$\epsilon^{ijk}\epsilon_{pqk} \equiv \delta^i_p\delta^j_q - \delta^i_q\delta^j_p, \quad (2.18)$$

which is intimately related to the vector triple product identity, (1.22). See Exercise 2.7.

A Second Order Tensor Has Four Sets of Components in General

This comes about as follows. Given a second order tensor $\mathbf{T}$ and a general basis $\{\mathbf{g}_j\}$, the action of $\mathbf{T}$ on each of the basis vectors is known, say

$$\mathbf{T}\mathbf{g}_j = \mathbf{T}_j. \quad (2.19)$$

Now each vector $\mathbf{T}_j$ may be expressed as a linear combination of the given basis vectors or their reciprocals. Choosing the latter, we may write

$$\mathbf{T}_j = T_{ij}\mathbf{g}^i. \quad (2.20)$$

The 9 coefficients T_{ij} are called *the cellar components* of $\mathbf{T}$. Explicitly,

$$T_{ij} = \mathbf{g}_i\cdot\mathbf{T}\mathbf{g}_j. \quad (2.21)$$

To represent $\mathbf{T}$ in terms of its components, we proceed just as we did in Chapter I. Thus if $\mathbf{v}$ is an arbitrary vector,

$$\begin{aligned} \mathbf{T}\mathbf{v} &= \mathbf{T}(v^j\mathbf{g}_j), \text{ representing } \mathbf{v} \text{ in the basis } \{\mathbf{g}_j\} \\ &= v^j\mathbf{T}\mathbf{g}_j, \text{ because } \mathbf{T} \text{ is linear} \\ &= v^jT_{ij}\mathbf{g}^i, \text{ by (2.19) and (2.20)} \\ &= T_{ij}\mathbf{g}^i(\mathbf{g}^j\cdot\mathbf{v}), \text{ by (2.7)} \\ &= T_{ij}\mathbf{g}^i\mathbf{g}^j(\mathbf{v}), \text{ by (1.29).} \end{aligned} \quad (2.22)$$

As $\mathbf{v}$ is arbitrary, (2.22) implies that

$$\mathbf{T} = T_{ij}\mathbf{g}^i\mathbf{g}^j, \tag{2.23}$$

and we see that $\{\mathbf{g}^i\mathbf{g}^j\}$ is a basis for the set of all 2nd order tensors.

Repeating the above line of reasoning, but with the roles of the cellar and roof base vectors reversed, we have

$$\mathbf{T}\mathbf{g}^j = \mathbf{T}^j = T^{ij}\mathbf{g}_i, \qquad \mathbf{T} = T^{ij}\mathbf{g}_i\mathbf{g}_j. \tag{2.25}$$

The 9 coefficients T^{ij} are called the *roof components* of $\mathbf{T}$. In other words, the T^{ij} are the components of $\mathbf{T}$ in the basis $\{\mathbf{g}_i\mathbf{g}_j\}$. The analogue of (2.21) is

$$T^{ij} = \mathbf{g}^i \cdot \mathbf{T}\mathbf{g}^j. \tag{2.26}$$

The alert reader will now realize that there are two additional sets of components that can be defined, namely

$$T^i_{.j} \equiv \mathbf{g}^i \cdot \mathbf{T}\mathbf{g}_j \tag{2.27}$$

and

$$T_j^{.i} \equiv \mathbf{g}_j \cdot \mathbf{T}\mathbf{g}^i. \tag{2.28}$$

These are called the *mixed components of* $\mathbf{T}$. The dots are used as distinguishing marks because, in general, $T^i_{.j} \neq T_j^{.i}$. It is easy to show that $\mathbf{T}$ has the following representations in terms of its mixed components:

$$\mathbf{T} = T^i_{.j}\mathbf{g}_i\mathbf{g}^j = T_j^{.i}\mathbf{g}^j\mathbf{g}_i, \tag{2.29}$$

i.e., the $T^i_{.j}$ are the components of $\mathbf{T}$ in the basis $\{\mathbf{g}_i\mathbf{g}^j\}$ and the $T_j^{.i}$ are the components of $\mathbf{T}$ in the basis $\{\mathbf{g}^j\mathbf{g}_i\}$.

Note that if $\mathbf{T}$ is symmetric, then $T_{ij} = \mathbf{g}_i \cdot \mathbf{T}\mathbf{g}_j = \mathbf{g}_j \cdot \mathbf{T}^T\mathbf{g}_i = \mathbf{g}_j \cdot \mathbf{T}\mathbf{g}_i = T_{ji}$. Likewise $T^i_{.j} = T_j^{.i}$ and $T^{ij} = T^{ji}$. However, $\mathbf{T} = \mathbf{T}^T$ does not imply that the matrices $[T^i_{.j}]$ and $[T_j^{.i}]$ are symmetric. See Exercise 2.23.

PROBLEM 2.6.

Compute the cellar, roof, and mixed components of the tensor given in Problem 1.4, using the base vectors given in Problem 2.1 and the associated reciprocal base vectors given in the Solution to Problem 2.2.

SOLUTION.

We are given that

$$\mathbf{Tv} \sim (-2v_x + 3v_z, \ -v_z, \ v_x + 2v_y)$$

and

$$\mathbf{g}_1 \sim (1, -1, 2), \ \mathbf{g}_2 \sim (0, 1, 1), \ \mathbf{g}_3 \sim (-1, -2, 1).$$

Hence,

$$\mathbf{T}\mathbf{g}_1 \sim (4, -2, -1), \ \mathbf{T}\mathbf{g}_2 \sim (3, -1, 2), \ \mathbf{T}\mathbf{g}_3 \sim (5, -1, -5),$$

and so

$$[T_{ij}] = [\mathbf{g}_i \cdot \mathbf{T}\mathbf{g}_j] = \begin{bmatrix} \mathbf{g}_1\cdot\mathbf{T}\mathbf{g}_1 & \mathbf{g}_1\cdot\mathbf{T}\mathbf{g}_2 & \cdot \\ \mathbf{g}_2\cdot\mathbf{T}\mathbf{g}_1 & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} = \begin{bmatrix} 4 & 8 & -4 \\ -3 & 1 & -6 \\ -1 & 1 & -8 \end{bmatrix}.$$

From Problem 2.2,

$$\mathbf{g}^1 \sim (1/6)(3,-1,1),\ \mathbf{g}^2 \sim (1/2)(-1,1,1),\ \mathbf{g}^3 \sim (1/6)(-3,-1,1),$$

and so

$$[T^i{}_j] = [\mathbf{g}^i \cdot \mathbf{T}\mathbf{g}_j] = \begin{bmatrix} \mathbf{g}^1\cdot\mathbf{T}\mathbf{g}_1 & \mathbf{g}^1\cdot\mathbf{T}\mathbf{g}_2 & \cdot \\ \mathbf{g}^2\cdot\mathbf{T}\mathbf{g}_1 & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} = (1/6)\begin{bmatrix} 13 & 12 & 11 \\ -21 & -6 & -33 \\ -11 & -6 & -19 \end{bmatrix}.$$

Finally, we have

$$\mathbf{T}\mathbf{g}^1 \sim (1/6)(-3,-1,1),\ \mathbf{T}\mathbf{g}^2 \sim (1/2)(5,-1,1),\ \mathbf{T}\mathbf{g}^3 \sim (1/6)(9,-1,-5),$$

yielding

$$[T_j{}^i] = [\mathbf{g}_j \cdot \mathbf{T}\mathbf{g}^i] = \begin{bmatrix} 0 & 0 & 1 \\ 4 & 0 & -1 \\ 0 & -1 & -2 \end{bmatrix},$$

$$[T^{ij}] = [\mathbf{g}^i \cdot \mathbf{T}\mathbf{g}^j] = (1/36)\begin{bmatrix} -7 & 51 & 23 \\ 9 & -45 & -45 \\ 11 & -39 & -31 \end{bmatrix}.$$

Change of Basis

Within a given frame, vectors and tensors are blissfully unaware of the bases we choose to represent them. That is, they are *geometric invariants*. Under a change of basis it is their components that change, not they themselves. A major aim of tensor analysis is to provide recipes for computing new components from old ones once a change of basis has been specified.

To concoct these recipes, let us start by assuming that each element of the new basis is a known linear combination of the elements of the old, say

$$\tilde{\mathbf{g}}_1 = A_1^1\mathbf{g}_1 + A_1^2\mathbf{g}_2 + A_1^3\mathbf{g}_3,\ \tilde{\mathbf{g}}_2 = A_2^1\mathbf{g}_1 + \dots,\ \tilde{\mathbf{g}}_3 = \dots. \tag{2.30}$$

We may summarize (2.30) in either matrix or index form as

$$\tilde{G} = GA \quad \text{or} \quad \tilde{\mathbf{g}}_j = A_j^i\mathbf{g}_i. \tag{2.31}$$

We shall need the inverse of (2.31) as well as the analogous relations between the new and old reciprocal bases. As $\det\tilde{G} = (\det G)(\det A)$ and $\{\mathbf{g}_i\}$ and $\{\tilde{\mathbf{g}}_i\}$ are bases, $\det A \neq 0$. Therefore, $A^{-1} \equiv [(A^{-1})_j^i]$ exists so, from (2.31),

$$G = \tilde{G}A^{-1} \quad \text{or} \quad \mathbf{g}_j = (A^{-1})_j^i\tilde{\mathbf{g}}_i. \qquad (2.32)^2$$

Furthermore, recalling that if a matrix B can be postmultiplied by a matrix C,

[2] An advantage of index notation over matrix notation is that the order of multiplication is immaterial. Thus $(A^{-1})_j^i\tilde{\mathbf{g}}_i = \tilde{\mathbf{g}}_i(A^{-1})_j^i$, but, in general, $\tilde{G}A^{-1} \neq A^{-1}\tilde{G}$.

then each row of BC will be a linear combination of the rows of C, we have

$$\tilde{G}^{-1} = A^{-1}G^{-1} \quad \text{or} \quad \tilde{\mathbf{g}}^i = (A^{-1})^i_j\mathbf{g}^j \tag{2.33}$$

$$G^{-1} = A\tilde{G}^{-1} \quad \text{or} \quad \mathbf{g}^i = A^i_j\tilde{\mathbf{g}}^j. \tag{2.34}$$

The relations between the new and old components of any vector $\mathbf{v}$ follow immediately because

$$\tilde{v}_j = \tilde{\mathbf{g}}_j\cdot\mathbf{v} = A^i_j\mathbf{g}_i\cdot\mathbf{v},\ \tilde{v}^i = \tilde{\mathbf{g}}^i\cdot\mathbf{v} = (A^{-1})^i_j\mathbf{g}^j\cdot\mathbf{v} \tag{2.35}$$

That is,

$$\tilde{v}_j = A^i_jv_i,\ \tilde{v}^i = (A^{-1})^i_jv^j. \tag{2.36}$$

Likewise, for any 2nd order tensor $\mathbf{T}$,

$$\begin{aligned}\tilde{T}_{ij} &= \tilde{\mathbf{g}}_i\cdot\mathbf{T}\tilde{\mathbf{g}}_j = A^k_i\mathbf{g}_k\cdot\mathbf{T}A^p_j\mathbf{g}_p,\ \tilde{T}^i_{.j} = \tilde{\mathbf{g}}^i\cdot\mathbf{T}\tilde{\mathbf{g}}_j = (A^{-1})^i_k\mathbf{g}^k\cdot\mathbf{T}A^p_j\mathbf{g}_p\\ \tilde{T}_j^{.i} &= \tilde{\mathbf{g}}_j\cdot\mathbf{T}\tilde{\mathbf{g}}^i = A^k_j\mathbf{g}_k\cdot\mathbf{T}(A^{-1})^i_p\mathbf{g}^p,\ \tilde{T}^{ij} = \tilde{\mathbf{g}}^i\cdot\mathbf{T}\tilde{\mathbf{g}}^j = (A^{-1})^i_k\mathbf{g}^k\cdot\mathbf{T}(A^{-1})^j_p\mathbf{g}^p\end{aligned} \tag{2.37}$$

That is,

$$\begin{aligned}\tilde{T}_{ij} &= A^k_iA^p_jT_{kp},\ \tilde{T}^i_{.j} = (A^{-1})^i_kA^p_jT^k_{.p}\\ \tilde{T}_j^{.i} &= A^k_j(A^{-1})^i_pT_k^{.p},\ \tilde{T}^{ij} = (A^{-1})^i_k(A^{-1})^j_pT^{kp}.\end{aligned} \tag{2.38}$$

PROBLEM 2.7.

If

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \quad \text{and } (v^1,v^2,v^3) = (2,3,-1),$$

find $(\tilde{v}^1,\tilde{v}^2,\tilde{v}^3)$.

SOLUTION.

Using the row reduction algorithm, we find that

$$A^{-1} = \begin{bmatrix} -\tfrac{1}{2} & 1 & \tfrac{1}{2} \\ 1 & -1 & -1 \\ -\tfrac{1}{2} & 1 & \tfrac{3}{2} \end{bmatrix}.$$

Hence, from $(2.36)_2$,

$$\begin{aligned}\tilde{v}^1 &= (-\tfrac{1}{2})(2) + (1)(3) + (\tfrac{1}{2})(-1) = \tfrac{3}{2}\\ \tilde{v}^2 &= (1)(2) + (-1)(3) + (-1)(-1) = 0\\ \tilde{v}^3 &= (-\tfrac{1}{2})(2) + (1)(3) + (\tfrac{3}{2})(-1) = \tfrac{1}{2}.\end{aligned}$$

PROBLEM 2.8.

Using the change of basis defined by the matrix A in Problem 2.7, compute $\tilde{T}_{21}$ and $\tilde{T}_2^{.3}$, where $\mathbf{T}$ is defined in Problem 2.6.

SOLUTION.

From $(2.38)_{1,3}$ we have

$$\begin{aligned}\tilde{T}_{21} &= A_2^1(A_1^1 T_{11} + A_1^2 T_{12} + A_1^3 T_{13}) \\ &\quad + A_2^2(A_1^1 T_{21} + A_1^2 T_{22} + A_1^3 T_{23}) \\ &\quad + A_2^3(A_1^1 T_{31} + A_1^2 T_{32} + A_1^3 T_{33}) \\ \tilde{T}_2^{\cdot 3} &= A_2^1[(A^{-1})_1^3 T_1^{\cdot 1} + (A^{-1})_2^3 T_1^{\cdot 2} + (A^{-1})_3^3 T_1^{\cdot 3}] \\ &\quad + A_2^2[(A^{-1})_1^3 T_2^{\cdot 1} + (A^{-1})_2^3 T_2^{\cdot 2} + (A^{-1})_3^3 T_2^{\cdot 3}] \\ &\quad + A_2^3[(A^{-1})_1^3 T_3^{\cdot 1} + (A^{-1})_2^3 T_3^{\cdot 2} + (A^{-1})_3^3 T_3^{\cdot 3}].\end{aligned}$$

The mixed and cellar components of $\mathbf{T}$ are given in the Solution to Problem 2.6 and the elements of the matrices A and A^{-1} in the statement and Solution of Problem 2.7. Thus all we need do is plug and chug:

$$\begin{aligned}\tilde{T}_{21} &= (2)[(1)(4) + (2)(8) + (-1)(-4)] \\ &\quad + (1)[(1)(-3) + (2)(1) + (-1)(-6)] = 53. \\ \tilde{T}_2^{\cdot 3} &= (2)[(-1/2)(0) + (1)(4) + (3/2)(0)] \\ &\quad + (1)[(-1/2)(0) + (1)(0) + (3/2)(-1)] = \frac{13}{2}.\end{aligned}$$

Exercises

2.1. Which of the following sets of vectors is a basis?
(a). $\mathbf{g}_1 \sim (4,6,2)$, $\mathbf{g}_2 \sim (1,0,1)$, $\mathbf{g}_3 \sim (1,3,0)$
(b). $\mathbf{g}_1 \sim (1,1,0)$, $\mathbf{g}_2 \sim (0,2,2)$, $\mathbf{g}_3 \sim (3,0,3)$
(c). $\mathbf{g}_1 \sim (1,1,1)$, $\mathbf{g}_2 \sim (1,-1,1)$, $\mathbf{g}_3 \sim (-1,1,-1)$

2.2. Let $\mathbf{g}_1 \sim (-1,0,0)$, $\mathbf{g}_2 \sim (1,1,0)$, $\mathbf{g}_3 \sim (1,1,1)$ and $\mathbf{v} \sim (1,2,3)$. Compute the reciprocal base vectors and the roof and cellar components of $\mathbf{v}$.

2.3. First simplify and then carry out explicitly any implied summations:
(a). $\delta_j^i v_i u^j$ (b). $\delta_j^2 \delta_k^j v^k$ (c). $\delta_j^3 \delta_1^j$ (d). $\epsilon_{i3k}\delta_p^i v^k$.

2.4. The vectors $\mathbf{g}_1$, $\mathbf{g}_2$, $\mathbf{g}_3$, given in Exercise 2.2 form a basis. If the roof components of $\mathbf{u}$ and $\mathbf{v}$ are (2,2,1) and (−3,1,2) respectively, compute the three cellar components of $\mathbf{u} \times \mathbf{v}$.

2.5. Establish (2.17) by first setting $(i,j,k) = (p,q,r) = (1,2,3)$ and then arguing about the value of both sides when ijk and pqr are various permutations and combinations of (1,2,3).

2.6. (a). Establish (2.18) by expanding the determinant in (2.17) by, say, its first row and then setting $r = k$.
(b). Use (2.18) to show that $\epsilon^{ijk}\epsilon_{pjk} = 2\delta_p^i$.

2.7. Establish the vector triple product identity (1.22) by first computing the roof components of $\mathbf{u} \times \mathbf{v}$ and then the cellar components of $(\mathbf{u} \times \mathbf{v}) \times \mathbf{w}$. Finally, use the identity (2.18).

2.8. *Components of the identity tensor.* Let

$$g_{ij} \equiv \mathbf{g}_i \cdot \mathbf{g}_j, \qquad g^{ij} \equiv \mathbf{g}^i \cdot \mathbf{g}^j.$$

Using the fact that $\mathbf{1v} = \mathbf{v}$ for all $\mathbf{v}$, show that the various components of $\mathbf{1}$ are

$$1_{ij} = g_{ij},\ 1^{i}_{.j} = \delta^i_j,\ 1_j^{\cdot i} = \delta^i_j,\ 1^{ij} = g^{ij}.$$

Thus

$$\mathbf{1} = \mathbf{g}^i\mathbf{g}_i = \mathbf{g}_i\mathbf{g}^i.$$

2.9. *Raising and lowering of indices.* Show that

$$\mathbf{g}^i = g^{ik}\mathbf{g}_k,\ \mathbf{g}_i = g_{ik}\mathbf{g}^k,\ v^i = g^{ik}v_k,\ v_i = g_{ik}v^k$$
$$T^i_{.j} = g^{ik}T_{kj},\ T^{ij} = g^{ik}T_k^{\cdot j},\ T_{ij} = g_{ik}T^k_{.j},\ \text{etc.}$$

As indicated, the effect of multiplying a component of a vector or tensor by g^{ik} and summing on k is to raise or lower an index. Thus, for example, the index on v_k is raised, becoming an i, by multiplying v_k by g^{ik}.

2.10. Noting that $[g_{ij}] = G^TG$, show that
(a). $\det [g_{ij}] = J^2$
(b). $g^{ik}g_{kj} = \delta^i_j$.

2.11. Show that
(a). $\mathbf{g}_i \times \mathbf{g}_j = \epsilon_{ijk}\mathbf{g}^k$
(b). $\mathbf{g}^k = \frac{1}{2}\epsilon^{ijk}\mathbf{g}_i \times \mathbf{g}_j$.

2.12. If $\mathbf{u} = u^i\mathbf{g}_i = u_i\mathbf{g}^i$, find formulas for the four different components of the 2nd order tensor $\mathbf{u}\times$. Use the following notation and procedure.

$$(\mathbf{u}\times)_{ij} = \mathbf{g}_i\cdot(\mathbf{u}\times\mathbf{g}_j) = \mathbf{u}\cdot(\mathbf{g}_j\times\mathbf{g}_i) = -\mathbf{u}\cdot\epsilon_{ijk}\mathbf{g}^k = \ldots$$

2.13. *A 3rd order tensor* is a linear operator that sends vectors into 2nd order tensors. Any 3rd order tensor may be represented as a linear combination of *triads*. A triad is a direct product of three vectors (**u**,**v**,**w**), denoted by **uvw** and defined by its action on any vector **x** as follows:

$$\mathbf{uvw}\cdot\mathbf{x} = \mathbf{uv}(\mathbf{w}\cdot\mathbf{x}).$$

Here is where the dot product, as a contraction operation, comes into its own. The *double dot product* of a triad acting on a dyad is a linear operator that produces a vector according to the rule

$$\mathbf{uvw}\cdot\cdot\mathbf{xy} = \mathbf{u}(\mathbf{w}\cdot\mathbf{x})(\mathbf{v}\cdot\mathbf{y}).$$

Finally, and obviously, we define the triple dot product of a triad acting on a triad as a linear operation that produces a scalar, according to the rule

$$\mathbf{uvw}\cdot\cdot\cdot\mathbf{xyz} = (\mathbf{w}\cdot\mathbf{x})(\mathbf{v}\cdot\mathbf{y})(\mathbf{u}\cdot\mathbf{z}).$$

The 3rd order *permutation tensor* may be denoted and defined by

$$\mathbf{P} \equiv \epsilon_{ijk}\mathbf{g}^i\mathbf{g}^j\mathbf{g}^k. \qquad (*)$$

Show that
(a). $\mathbf{P}\cdot\cdot\cdot\mathbf{wvu} = \mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})$
(b). $\mathbf{P}\cdot\cdot\mathbf{vu} = \mathbf{u}\times\mathbf{v}$
(c). $\mathbf{P}\cdot\mathbf{u} = -\mathbf{u}\times$.

2.14. Show, using (2.17), that

$$\begin{aligned}(\mathbf{a}\times\mathbf{b})(\mathbf{c}\times\mathbf{d}) &= \mathbf{da}(\mathbf{b}\cdot\mathbf{c}) + \mathbf{cb}(\mathbf{d}\cdot\mathbf{a}) - \mathbf{db}(\mathbf{c}\cdot\mathbf{a}) - \mathbf{ca}(\mathbf{d}\cdot\mathbf{b})\\ &\quad + [(\mathbf{a}\cdot\mathbf{c})(\mathbf{b}\cdot\mathbf{d}) - (\mathbf{a}\cdot\mathbf{d})(\mathbf{b}\cdot\mathbf{c})]\,\mathbf{1}\\ &= \mathbf{T}\cdot\mathbf{S} - (\tfrac{1}{2}\mathbf{T}\cdot\cdot\mathbf{S})\mathbf{1},\end{aligned} \tag{2.39}$$

where, recalling Exercise 1.13(b),

$$\mathbf{S} = \mathbf{ba} - \mathbf{ab} = (\mathbf{a}\times\mathbf{b})\times\ , \quad \mathbf{T} = \mathbf{dc} - \mathbf{cd} = (\mathbf{c}\times\mathbf{d})\times\ .$$

This result is useful in describing finite rotations. See Exercise 2.19.

2.15. Any change of basis $\tilde{\mathbf{g}}_j = A^i_j\mathbf{g}_i$ can be effected by a nonsingular 2nd order tensor $\mathbf{B}$ such that $\tilde{\mathbf{g}}_j = \mathbf{B}\mathbf{g}_j$. Show that
(a). $B^i_{\cdot j} = A^i_j$.
(b). $\tilde{\mathbf{g}}^i = (\mathbf{B}^{-1})^T\mathbf{g}^i$.

2.16. *An orthogonal tensor* $\mathbf{Q}$ *satisfies* $\mathbf{Q}^T\mathbf{Q} = \mathbf{1}$, i.e. $\mathbf{Q}^T = \mathbf{Q}^{-1}$.
(a). Let $Q = [\mathbf{e}_i\cdot\mathbf{Q}\mathbf{e}_j]$, where $\{\mathbf{e}_i\}$ is the standard Cartesian basis. Show that $Q^TQ = I$.
(b). Show that $\det Q = \pm 1$. Hint: $\det AB = (\det A)\,(\det B)$ (If $\det Q = +1$, $\mathbf{Q}$ is called *a proper* ⊥ *tensor* or *rotator*)
(c). Show that under the ⊥ *change of coordinates* $\tilde{\mathbf{g}}_j = \mathbf{Q}\mathbf{g}_j$, the cellar and roof components of $\mathbf{1}$ (i.e., g_{ij} and g^{ij}) are unchanged.
(d). Explain why the columns or rows of Q may be regarded as the Cartesian components of mutually ⊥ unit vectors.
(e). If $\mathbf{Q} = Q^i_{\cdot j}\mathbf{g}_i\mathbf{g}^j = Q_j^{\cdot i}\mathbf{g}^j\mathbf{g}_i$,

show that

$$\mathbf{Q}^T = Q^i_{\cdot j}\mathbf{g}^j\mathbf{g}_i = Q_j^{\cdot i}\mathbf{g}_i\mathbf{g}^j,$$

and hence that

$$Q_k^{\cdot i}Q^k_{\cdot j} = \delta^i_j.$$

2.17. *A reflector* is an ⊥ tensor $\mathbf{H}$ that reflects vectors across a plane with normal $\mathbf{n}$, as indicated in Fig. 2.4.
(a). Show, geometrically, that

$$\mathbf{H} = \mathbf{1} - 2\bar{\mathbf{n}}\bar{\mathbf{n}}. \tag{2.40}$$

Hint: What vector when added to $\mathbf{v}$ gives $\mathbf{H}\mathbf{v}$?
(b). Show that $\mathbf{H} = \mathbf{H}^T$ and hence that $\mathbf{H}^2 = \mathbf{1}$.
(c). With the aid of Cartesian components, show that $\det H = -1$. (Hint: take $\mathbf{e}_x = \bar{\mathbf{n}}$).
(d). If $\mathbf{n} \sim (1,-2,3)$, and $\mathbf{v} \sim (v_x, v_y, v_z)$, fill in the blanks:

$$\mathbf{H}\mathbf{v} \sim (\_\_\_\_\_\_,\ \_\_\_\_\_\_,\ \_\_\_\_\_\_).$$

(e). If $\mathbf{u}$ and $\mathbf{v}$ are distinct unit vectors, show that

$$\mathbf{H} = \mathbf{1} + \frac{\mathbf{uv} + \mathbf{vu} - \mathbf{uu} - \mathbf{vv}}{1 - \mathbf{u}\cdot\mathbf{v}} \tag{2.41}$$

is a reflector that sends $\mathbf{u}$ into $\mathbf{v}$ and $\mathbf{v}$ into $\mathbf{u}$.

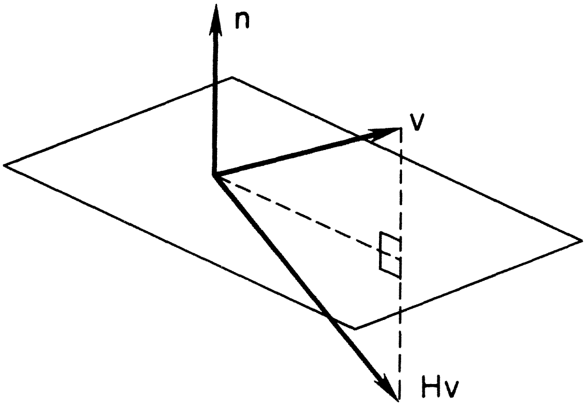


Figure 2.4

2.18. *A 3-dimensional rotator R* is characterized by an *axis of rotation,* with direction **e**, and an *angle of rotation* θ, reckoned positive by the right-hand rule. Fig. 2.5 shows a vector **u** that has been rotated about **e** through θ into the vector $\mathbf{v} = \mathsf{R}\mathbf{u}$.

(a). By resolving **v** into components along the three mutually $\perp$ vectors $\mathbf{e}$, $\mathbf{e} \times \mathbf{u}$, and $\mathbf{e} \times (\mathbf{e} \times \mathbf{u})$, show that $\mathbf{v} = \mathbf{u} + (\sin\theta)\mathbf{e} \times \mathbf{u} + (1 - \cos\theta)\mathbf{e} \times (\mathbf{e} \times \mathbf{u}) = (\cos\theta)\mathbf{u} + (1 - \cos\theta)\mathbf{e}(\mathbf{e}\cdot\mathbf{u}) + (\sin\theta)\mathbf{e} \times \mathbf{u}$, and hence that

$$\mathsf{R} = (\cos\theta)\mathsf{1} + (1 - \cos\theta)\mathbf{ee} + (\sin\theta)\mathbf{e}\times . \tag{2.42}$$

(b). Show that $\mathsf{R}\mathbf{e} = \mathbf{e}$.

(c). Using Cartesian components, show that $\det R = +1$.
Hint: take $\mathbf{e}_x = \mathbf{e}$.

(d). Show that in terms of the *finite rotation vector*

$$\mathbf{r} = 2\tan(\theta/2)\mathbf{e}, \tag{2.43}$$

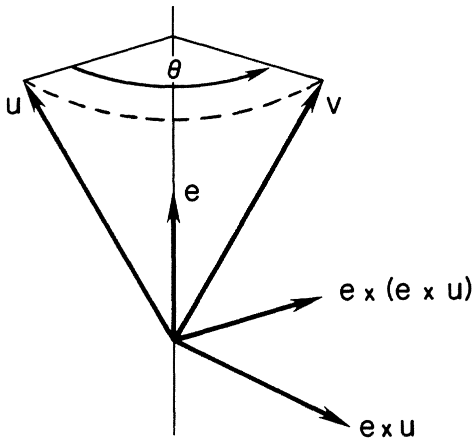


Figure 2.5

$$\mathbf{R} = (1 + \tfrac{1}{4}\mathbf{r}\cdot\mathbf{r})^{-1}[(1 - \tfrac{1}{4}\mathbf{r}\cdot\mathbf{r})\mathbf{1} + \tfrac{1}{2}\mathbf{rr} + \mathbf{r}\times\,]. \tag{2.44}$$

(The advantage of introducing $\mathbf{r}$ is that (2.44) is a rational function of the components of $\mathbf{r}$; the disadvantage hits when $\theta = \pi$.)

(e). If $\mathbf{r} \sim (1,1,1)$, compute θ and the matrix of Cartesian components of $\mathbf{R}$.

2.19. Given two non-parallel, three-dimensional unit vectors $\mathbf{u}$ and $\mathbf{v}$, show, by taking $\sin\theta\,\mathbf{e} = \mathbf{u}\times\mathbf{v}$ and using (2.39), that the rotator that sends $\mathbf{u}$ into $\mathbf{v}$ about an axis $\perp$ to $\mathbf{u}$ and $\mathbf{v}$ is given by

$$\mathbf{R} = \mathbf{1} + \mathbf{vu} - \mathbf{uv} + (1 + \mathbf{u}\cdot\mathbf{v})^{-1}[(\mathbf{uv} + \mathbf{vu})(\mathbf{u}\cdot\mathbf{v}) - (\mathbf{uu} + \mathbf{vv})]. \tag{2.45}$$

As (2.45) is free of the cross product, it is valid in any number of dimensions! (With (2.40) and (2.41), or (2.42) and (2.45), we can construct, in two different ways, $\perp$ tensors with determinant equal to -1 or $+1$, respectively.)

2.20. Find the 2×2 matrix of the Cartesian components of the rotator that sends $\mathbf{e}_x$ into the unit vector $\mathbf{e}_x \cos\theta + \mathbf{e}_y \sin\theta$.

2.21. *Cartesian tensor notation* means writing all indices as subscripts and may be used whenever $\mathbf{g}^i = \mathbf{g}_i$. Show that this happens if and only if $\mathbf{g}_i = \mathbf{Q}\mathbf{e}_i$, where $\mathbf{Q}$ is an $\perp$ tensor. Invariant forms are often most easily established by introducing the standard Cartesian basis $\{\mathbf{e}_i\}$ and using Cartesian tensor notation. Thus, with the aid of (2.18),

$$\begin{aligned}(\mathbf{u}\times\mathbf{v})\cdot(\mathbf{u}\times\mathbf{v}) &\equiv \epsilon_{ijk}u_j v_k \epsilon_{iqr} u_q v_r \equiv (\delta_{jq}\delta_{kr} - \delta_{jr}\delta_{kq})\, u_j v_k u_q v_r \\ &\equiv u_q v_r u_q v_r - u_r v_q u_q v_r \equiv (\mathbf{u}\cdot\mathbf{u})(\mathbf{v}\cdot\mathbf{v}) - (\mathbf{u}\cdot\mathbf{v})^2.\end{aligned}$$

2.22. Given a basis $\{\mathbf{g}_1,\mathbf{g}_2,\mathbf{g}_3\}$, explain the geometrical construction of the reciprocal basis $\{\mathbf{g}^1,\mathbf{g}^2,\mathbf{g}^3\}$.

2.23. Given the 2-dimensional 2nd order tensor $\mathbf{T}\mathbf{v} \sim (v_x - v_y, v_x)$, where $\mathbf{v} \sim (v_x, v_y)$,

(a). Determine its symmetric and skew parts, $\mathbf{S}$ and $\mathbf{A}$.

(b). If $\mathbf{g}_1 \sim (1,0)$ and $\mathbf{g}_2 \sim (1,1)$, compute the matrices $[T^i_{.j}]$, $[T_j^{\ .i}]$, $[S^i_{.j}]$ and $[S_j^{\ .i}]$. Are these last two matrices symmetric?

2.24. Let $\{\mathbf{g}_i\}$ be a basis, $\mathbf{T}$ any 2nd order tensor, and set $\mathbf{h}_i \equiv \mathbf{T}\mathbf{g}_i$.

(a). Show that $\mathbf{T} = \mathbf{h}_i\mathbf{g}^i$.

(b). Show that $\tfrac{1}{2}(\mathbf{h}_i\mathbf{g}^i - \mathbf{g}^i\mathbf{h}_i)$ is the skew part of $\mathbf{T}$.

(c). If $\mathbf{T}$ is 3-dimensional, show that $\tfrac{1}{2}\mathbf{g}^i \times \mathbf{h}_i$ is the axis of its skew part. (See Exercise 1.18).

(d). If $\{\mathbf{g}_i\}$ is the basis given in Problem 2.1 and $\mathbf{h}_1 \sim (1,0,-1)$, $\mathbf{h}_2 \sim (2,1,0)$, $\mathbf{h}_3 \sim (0,1,1)$, compute the Cartesian components of $\mathbf{T}$ and the axis of its skew part. Note that $\{\mathbf{g}^i\}$ is given in the solution to Problem 2.2.

CHAPTER III

Newton's Law and Tensor Calculus

Newton's Law of Motion is studied in introductory courses in calculus, physics, and dynamics. Being familiar, fundamental, and simple, Newton's Law is an ideal vehicle for introducing many of the key ideas in tensor calculus.

In its most primitive form, Newton's Law states that, in an *inertial frame,* the motion of a mass-point p obeys

$$\mathbf{f} = m\ddot{\mathbf{x}}, \tag{3.1}$$[1]

where $\mathbf{f}$ is the force acting on p, m is its mass, $\mathbf{x}$ is its position with respect to a fixed origin O in the inertial frame, and a dot—Newton's notation—denotes differentiation with respect to time.

Rigid Bodies

Rigid bodies in *classical mechanics* are composed of mass-points that remain a fixed distance apart and exert on one another mutually parallel forces only. It may be inferred from (3.1) that such bodies obey the *gross form of Newton's Law,*

$$\mathbf{F} = \dot{\mathbf{L}}, \tag{3.2}$$

where $\mathbf{F}$ is the net external force acting on the body,

$$\mathbf{L} = M\dot{\mathbf{X}} \tag{3.3}$$

is its *linear momentum, M* is its mass, and $\mathbf{X}$ the position of its center of

[1] According to Truesdell, Newton never stated his law in this form; Euler first did!

mass.[2] It may also be inferred from (3.1) that a rigid body of classical mechanics obeys the equation of *conservation of rotational momentum,*

$$\mathbf{X} \times \mathbf{F} + \mathbf{T} = \dot{\mathbf{R}}_0, \tag{3.4}$$

where $\mathbf{T}$ is the net torque on the body and $\mathbf{R}_0$ is the *rotational momentum* about O.

In *continuum mechanics,* (3.2) and (3.4) are taken as *postulates,* valid for *any* body, while (3.3) follows from the definition of center of mass; rigid bodies are defined by a special constitutive assumption, without any reference to "atomic" forces. See Exercises 4.19 and 4.22.

Combining (3.2) and (3.3), we obtain

$$\mathbf{F} = M\ddot{\mathbf{X}}, \tag{3.5}$$

This law, though a gross one, yields, by itself, useful results for the motion of bodies that are nearly rigid and either nearly spherical or else rotating at a known rate, and small in the sense that their diameters are much less than the scale of the field of interest. I have in mind raindrops falling on my head, golf balls soaring over fairways, a record throw of a shot (but not one of a javelin or discus), a satellite tumbling in space, or an asteroid orbiting the earth.

As (3.1) and (3.5) are of the same form, we shall work with (3.1), but leave the physical interpretation open.

New Conservation Laws

The simplicity of (3.1)—the coordinate-free form on Newton's Law—suggests the following manipulations.

(a) Take the dot product of both sides of (3.1) with $\dot{\mathbf{x}}$ to obtain

$$\mathbf{f}\cdot\dot{\mathbf{x}} = m\ddot{\mathbf{x}}\cdot\dot{\mathbf{x}} = (\tfrac{1}{2}m\dot{\mathbf{x}}\cdot\dot{\mathbf{x}})^{\cdot} \equiv \dot{k}. \tag{3.6}$$

In words, the *external power* ($\mathbf{f}\cdot\dot{\mathbf{x}}$) is equal to the rate of change of *kinetic energy* (k).

(b) Take the cross product of both sides of (3.1) with $\mathbf{x}$ and note that $\dot{\mathbf{x}} \times \dot{\mathbf{x}} = \mathbf{o}$ to obtain

$$\mathbf{t}_O \equiv \mathbf{x} \times \mathbf{f} = \mathbf{x} \times m\ddot{\mathbf{x}} = (\mathbf{x} \times m\dot{\mathbf{x}})^{\cdot} \equiv \dot{\mathbf{r}}_O, \tag{3.7}$$

i.e., the *torque* $\mathbf{t}_O$ about the origin is equal to the rate of change of the *rotational momentum* $\mathbf{r}_O$ about the origin.

(c) If there exists a potential $v(\mathbf{x})$ such that

$$\mathbf{f}\cdot\dot{\mathbf{x}} = -\dot{v}, \tag{3.8}$$

then $\mathbf{f}$ is said to be *conservative,* and (3.6) implies that

[2] See, for example, Goldstein's *Classical Mechanics,* 2nd Ed.

$$k + v = c, \text{ a constant,} \tag{3.9}$$

i.e., *energy* is conserved. Often we can solve simple problems with the aid of (3.9) and thus avoid having to solve a differential equation. See Exercise 3.2.

(d) See Problem 3.2.

Nomenclature

The set of all points C occupied by a mass-point p between two times, say a and b, is called the *orbit* of p. The point on C occupied by p at time t may be denoted by $P(t)$. If C is represented in the *parametric form*

$$C{:}\mathbf{x} = \hat{\mathbf{x}}(t) = \hat{x}(t)\mathbf{e}_x + \hat{y}(t)\mathbf{e}_y + \hat{z}(t)\mathbf{e}_z, \quad a \le t \le b, \qquad (3.10)^3$$

then the vector function $\hat{\mathbf{x}}$ is called the *trajectory* of p, and $\mathbf{x}$ the *radius vector* to C.

If $\mathbf{x}$ is differentiable, then

$$\dot{\mathbf{x}} \equiv \hat{\mathbf{v}}(t), \quad a < t < b, \tag{3.11}$$

is called the *velocity* of p (at t) and $|\hat{\mathbf{v}}(t)|$ its *speed*. $\hat{\mathbf{x}}$ is said to *be smooth at* t if $\dot{\mathbf{x}}$ is continuous at t and *smooth* (everywhere) if $\dot{\mathbf{x}}$ is continuous for all $t\epsilon(a,b)$.[4]

If $\hat{\mathbf{x}}$ is smooth, then $\hat{\mathbf{v}}(t)$ is tangent to C at $P(t)$. This follows immediately from the definition

$$\hat{\mathbf{v}}(t) = \lim_{\Delta t \to 0} \frac{\Delta \mathbf{x}}{\Delta t}, \quad \Delta\mathbf{x} \equiv \hat{\mathbf{x}}(t + \Delta t) - \hat{\mathbf{x}}(t)$$

and from Fig. 3.1, which shows that the vector $\Delta\mathbf{x}/\Delta t$, being parallel to the vector $\Delta\mathbf{x}$ from $P(t)$ to $P(t + \Delta t)$, must approach the direction of the tangent line at $P(t)$ as $\Delta t \to 0$.

The *length* s along a smooth trajectory, measured from a point $P(t_1)$, satisfies the differential equation

$$ds/dt = |\hat{\mathbf{v}}(t)|, \quad s(t_1) = 0, \tag{3.12}$$

i.e.,

$$s = \int_{t_1}^{t} |\hat{\mathbf{v}}(\tau)| d\tau. \tag{3.13}$$

If $\dot{\mathbf{x}}$ is differentiable, then

[3] Here and henceforth we adopt a useful notation: if $\hat{f}$ denotes a function, then its *value at t* is denoted by either $\hat{f}(t)$ or f, the latter notation to be used whenever the functional form of $\hat{f}$ is not important.

[4] Note that a smooth trajectory need not have a smooth orbit. Consider a point p on the rim of a rolling wheel. Its trajectory is smooth but its orbit, a cycloid, is not: at points where p touches the ground, the velocity is zero but the orbit has a cusp.

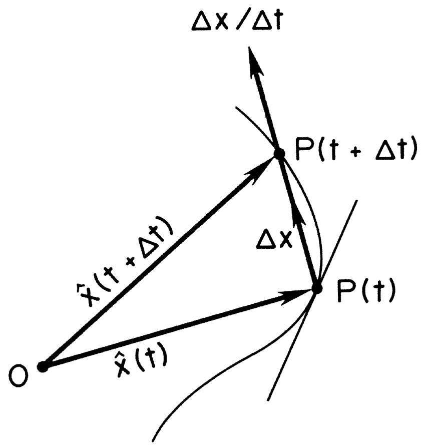


Figure 3.1

$$\ddot{\mathbf{x}} = \dot{\mathbf{v}} \equiv \hat{\mathbf{a}}(t) \tag{3.14}$$

is called the *acceleration of p* (at t).

PROBLEM 3.1.

If $\hat{\mathbf{x}}$ is smooth, show that the area swept out by $\mathbf{x}$ as t varies from a to b is given by

$$A(b,a) = \tfrac{1}{2} \int_a^b |\mathbf{x} \times \dot{\mathbf{x}}| dt. \tag{3.15}$$

SOLUTION.

From Fig. 3.2 it is seen that the area of the shaded triangle, being half the area of the parallelogram having $\mathbf{x}$ and $\Delta\mathbf{x}$ as co-terminal edges, is given by

$$\Delta A = \tfrac{1}{2} |\mathbf{x} \times \Delta\mathbf{x}|.$$

As $\alpha|\mathbf{v}| = |\alpha\mathbf{v}|$, $\Delta A/\Delta t = \tfrac{1}{2} |\mathbf{x} \times \Delta\mathbf{x}/\Delta t|$. Hence, in the limit as $\Delta t \to 0$,

$$dA/dt = \tfrac{1}{2} |\mathbf{x} \times \dot{\mathbf{x}}|, \qquad (3.16)^5$$

which implies (3.15).

PROBLEM 3.2.

Show that in a central force field ($\mathbf{f}$ parallel to $\mathbf{x}$)

(i) $\mathbf{x}$ lies in a plane.
(ii) Kepler's Law holds: $\mathbf{x}$ sweeps out equal areas in equal times.

[5] For those who like a little more precision and detail: Assume that $\dot{\mathbf{x}} = \lim \Delta\mathbf{x}/\Delta t$ exists at $P(t)$. Then the right side of (3.16) exists at $P(t)$. By the triangle inequality, $|\tfrac{1}{2}|\mathbf{x} \times \dot{\mathbf{x}}| - \tfrac{1}{2}|\mathbf{x} \times \Delta\mathbf{x}/\Delta t|| \le \tfrac{1}{2} |\mathbf{x} \times \dot{\mathbf{x}} - \mathbf{x} \times \Delta\mathbf{x}/\Delta t| = \tfrac{1}{2}|\mathbf{x} \times (\dot{\mathbf{x}} - \Delta\mathbf{x}/\Delta t)| \le \tfrac{1}{2}|\mathbf{x}||\dot{\mathbf{x}} - \Delta\mathbf{x}/\Delta t| \to 0$.

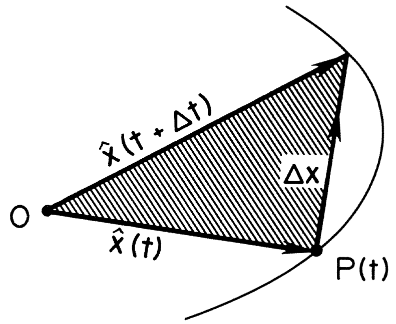


Figure 3.2

SOLUTION.

(i) If $\mathbf{f}$ is parallel to $\mathbf{x}$, then $\mathbf{x} \times \mathbf{f} = \mathbf{o}$ and (3.7) implies that

$$\mathbf{x} \times \dot{\mathbf{x}} = \mathbf{c}, \tag{3.17}$$

a constant vector. Thus

$$\mathbf{x}\cdot\mathbf{c} = \mathbf{x}\cdot(\mathbf{x} \times \dot{\mathbf{x}}) = \dot{\mathbf{x}}\cdot(\mathbf{x} \times \mathbf{x}) = 0, \tag{3.18}$$

i.e., $\mathbf{x}$ lies in a plane.

(ii) It follows from (3.15) and problem 3.1 that $dA/dt = ½|\mathbf{c}|$. i.e., $\hat{A}(t + \Delta t) - \hat{A}(t) = ½|\mathbf{c}|\Delta t$. This is Kepler's Law.

Newton's Law in Cartesian Components

The four conservation laws derived under (a)–(d) are simple and useful and epitomize the virtues of expressing the laws of mechanics in coordinate-free form. However, except in special cases, these conservation laws alone cannot provide all of the details of the motion of a mass-point.To extract full information from Newton's Law, we must express (3.1) in component form. In specific cases, we try to choose a coordinate system in which the components of $\mathbf{f}$ simplify.

In the simplest case, the Cartesian components of $\mathbf{f}$ are known functions of position and time, so we set

$$\mathbf{f} = f_x\mathbf{e}_x + f_y\mathbf{e}_y + f_z\mathbf{e}_z. \tag{3.19}$$

From (3.10)

$$\ddot{\mathbf{x}} = \ddot{x}\mathbf{e}_x + \ddot{y}\mathbf{e}_y + \ddot{z}\mathbf{e}_z, \tag{3.20}$$

whence

$$f_x = m\ddot{x}, \quad f_y = m\ddot{y}, \quad f_z = m\ddot{z}. \tag{3.21}$$

Newton's Law in Plane Polar Coordinates

To study the motion of a planet about the sun or, more generally, the motion of a point-mass under any central force (say a ball-bearing, attached to a rubber band, whirling around on a horizontal, frictionless table), we would want to introduce *plane polar coordinates* (r, θ) defined by

$$x = r \cos \theta, \quad y = r \sin \theta. \tag{3.22}$$

With the change of variables (3.22), the position of a mass-point takes the form

$$\mathbf{x} = \hat{\mathbf{x}}(r,\theta) = r \cos \theta \mathbf{e}_x + r \sin \theta \mathbf{e}_y, \quad \Big| \quad \mathbf{x} = \hat{\mathbf{x}}(u^j) = \hat{x}^i(u^j)\mathbf{e}_i \qquad (3.23)^6$$

where r and θ are unknown functions of time. By the chain rule,

$$\mathbf{v} = \dot{\mathbf{x}} = (\partial \mathbf{x}/\partial r)\dot{r} + (\partial \mathbf{x}/\partial \theta)\dot{\theta}. \quad \Big| \quad \mathbf{v} = (\partial \mathbf{x}/\partial u^i)\dot{u}^i \tag{3.24}$$

From (3.23),

$$\mathbf{g}_r \equiv \partial \mathbf{x}/\partial r = \cos \theta \mathbf{e}_x + \sin \theta \mathbf{e}_y$$

$$\mathbf{g}_\theta \equiv \partial \mathbf{x}/\partial \theta = -r \sin \theta \mathbf{e}_x + r \cos \theta \mathbf{e}_y. \quad \Big| \quad \mathbf{g}_i \equiv \partial \mathbf{x}/\partial u^i = (\partial x^k/\partial u^i)\mathbf{e}_k. \tag{3.25}$$

Since

$$J(\mathbf{g}_r, \mathbf{g}_\theta) = \begin{vmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{vmatrix} = r, \quad \Big| \quad J(\mathbf{g}_1, \ldots) = \det [x^i{}_{,j}] \tag{3.26}$$

$\{\mathbf{g}_r, \mathbf{g}_\theta\}$ is a basis at every point of the plane except the origin. (Such *exceptional points* are the bane and beauty of many useful coordinate systems. Often the field equations of physics and mechanics admit solutions that are *singular* at exceptional points. Sometimes these solutions must be discarded on physical grounds. Yet other times they represent such useful idealizations as sources, sinks, vortex lines, concentrated forces, or black holes!)

We shall call $\mathbf{g}_r$ and $\mathbf{g}_\theta$ the *cellar base vectors of the (r, θ) coordinate system,*[7] and denote their coefficients in (3.24) by

$$v^r = \dot{r}, \quad v^\theta = \dot{\theta}. \quad \Big| \quad v^i = \dot{u}^i. \tag{3.27}$$

These are the *roof components* of $\mathbf{v}$. With this notation, (3.24) reads

$$\mathbf{v} = v^r \mathbf{g}_r + v^\theta \mathbf{g}_\theta. \quad \Big| \quad \mathbf{v} = v^i \mathbf{g}_i. \tag{3.28}$$

[6] A number of equations that follow have two parts, separated by a vertical dashed line. The right half represents the general tensor form and will be discussed presently.

[7] Here and henceforth we omit the qualifier "aside from exceptional points" whenever we speak of the cellar (or roof) base vectors of a coordinate system.

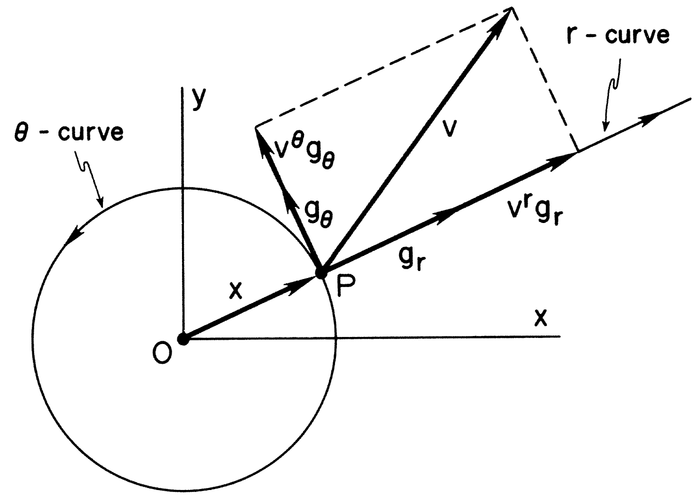


Figure 3.3

The geometric interpretation of $\mathbf{g}_r$ and $\mathbf{g}_\theta$ is simple. Through any fixed point P_* with polar coordinates (r_*,θ_*), except the origin, there passes a unique *r-coordinate curve* with the parametric representation $\hat{\mathbf{x}}(r,\theta_*)$, and a unique *θ-coordinate curve* with the parametric representation $\hat{\mathbf{x}}(r_*,\theta)$. By the same argument we used to show that $\mathbf{v} = \dot{\mathbf{x}}$ is tangent to the curve traced out by $\mathbf{x}$ as t varies, $\mathbf{g}_r = \partial\mathbf{x}/\partial r$ and $\mathbf{g}_\theta = \partial\mathbf{x}/\partial\theta$ are tangent, respectively, to the curves traced out by $\mathbf{x}$ as r or θ varies. Fig. 3.3 shows the r- and θ-coordinate curves that pass through a typical point P and the associated base vectors $\mathbf{g}_r$ and $\mathbf{g}_\theta$. Also indicated is the decomposition (3.28).

The Physical Components of a Vector

The physical component of a vector $\mathbf{w}$ in the direction of a vector $\mathbf{u}$ is defined to be $\mathbf{w}\cdot\bar{\mathbf{u}}$. Thus if $v^{(r)}$ and $v^{(\theta)}$ denote, respectively, the physical components of $\mathbf{v}$ in the directions of the roof base vectors

$$\mathbf{g}^r = \cos\theta\mathbf{e}_x + \sin\theta\mathbf{e}_y, \quad \mathbf{g}^\theta = r^{-1}(-\sin\theta\mathbf{e}_x + \cos\theta\mathbf{e}_y), \tag{3.29}$$

then from (3.27) and (3.28)

$$\left.\begin{aligned} v^{(r)} &= \mathbf{v}\cdot\bar{\mathbf{g}}^r = v^r = \dot{r} \\ v^{(\theta)} &= \mathbf{v}\cdot\bar{\mathbf{g}}^\theta = rv^\theta = r\dot{\theta}. \end{aligned}\right| \quad v^{(i)} = \mathbf{v}\cdot\bar{\mathbf{g}}^i = v^i/|\mathbf{g}^i|. \tag{3.30}$$

These expressions agree with the ordinary definitions of radial and angular velocity.

To compute the acceleration in plane polar coordinates, we differentiate (3.28) with respect to time, so obtaining

$$\mathbf{a} = \dot{\mathbf{v}} = \dot{v}^r \mathbf{g}_r + v^r \dot{\mathbf{g}}_r + \dot{v}^\theta \mathbf{g}_\theta + v^\theta \dot{\mathbf{g}}_\theta. \quad \Bigg| \quad \mathbf{a} = \dot{v}^i \mathbf{g}_i + v^i \dot{\mathbf{g}}_i. \tag{3.31}$$

The Christoffel Symbols

To compute $\dot{\mathbf{g}}_r$ and $\dot{\mathbf{g}}_\theta$ we must use the chain rule, for by (3.25), $\mathbf{g}_r$ and $\mathbf{g}_\theta$ are functions of r and θ and these, in turn, are (unknown) functions of t. Thus

$$\begin{aligned} \dot{\mathbf{g}}_r &= (\partial \mathbf{g}_r/\partial r)\dot{r} + (\partial \mathbf{g}_r/\partial \theta)\dot{\theta} \\ &= v^r \mathbf{g}_{r,r} + v^\theta \mathbf{g}_{r,\theta} \\ \dot{\mathbf{g}}_\theta &= (\partial \mathbf{g}_\theta/\partial r)\dot{r} + (\partial \mathbf{g}_\theta/\partial \theta)\dot{\theta} \\ &= v^r \mathbf{g}_{\theta,r} + v^\theta \mathbf{g}_{\theta,\theta}, \end{aligned} \quad \Bigg| \quad \begin{aligned} \dot{\mathbf{g}}_i &= (\partial \mathbf{g}_i/\partial u^j)\dot{u}^j \\ &= v^j \mathbf{g}_{i,j} \end{aligned} \tag{3.32}$$

[8]

where a comma indicates partial differentiation with respect to the variable(s) that follow it. Being derivatives of vectors, $\mathbf{g}_{r,r}$, $\mathbf{g}_{r,\theta} = \mathbf{g}_{\theta,r}$, and $\mathbf{g}_{\theta,\theta}$ are themselves vectors. Therefore each can be written as a linear combination of $\mathbf{g}_r$ and $\mathbf{g}_\theta$, say

$$\begin{aligned} \mathbf{g}_{r,r} &= \Gamma^r_{rr}\mathbf{g}_r + \Gamma^\theta_{rr}\mathbf{g}_\theta \\ \mathbf{g}_{r,\theta} = \mathbf{g}_{\theta,r} &= \Gamma^r_{r\theta}\mathbf{g}_r + \Gamma^\theta_{r\theta}\mathbf{g}_\theta \\ \mathbf{g}_{\theta,\theta} &= \Gamma^r_{\theta\theta}\mathbf{g}_r + \Gamma^\theta_{\theta\theta}\mathbf{g}_\theta. \end{aligned} \quad \Bigg| \quad \mathbf{g}_{i,j} = \mathbf{g}_{j,i} = \Gamma^k_{ij}\mathbf{g}_k \tag{3.33}$$

The coefficients of $\mathbf{g}_r$ and $\mathbf{g}_\theta$ in (3.33) are called the *Christoffel symbols* of the (r, θ) coordinate system. Substituting, (3.33) into (3.32), we have

$$\begin{aligned} \dot{\mathbf{g}}_r &= v^r(\Gamma^r_{rr}\mathbf{g}_r + \Gamma^\theta_{rr}\mathbf{g}_\theta) + v^\theta(\Gamma^r_{r\theta}\mathbf{g}_r + \Gamma^\theta_{r\theta}\mathbf{g}_\theta) \\ \dot{\mathbf{g}}_\theta &= v^r(\Gamma^r_{r\theta}\mathbf{g}_r + \Gamma^\theta_{r\theta}\mathbf{g}_\theta) + v^\theta(\Gamma^r_{\theta\theta}\mathbf{g}_r + \Gamma^\theta_{\theta\theta}\mathbf{g}_\theta) \end{aligned} \quad \Bigg| \quad \dot{\mathbf{g}}_i = v^j \Gamma^k_{ij}\mathbf{g}_k. \tag{3.34}$$

Finally, we place these relations into our expression for the acceleration, (3.31), and collect coefficients of $\mathbf{g}_r$ and $\mathbf{g}_\theta$, to obtain

[8] Note that $\mathbf{g}_{r,\theta} = \mathbf{x}_{,r\theta} = \mathbf{x}_{,\theta r} = \mathbf{g}_{\theta,r}$. | $\mathbf{g}_{i,j} = \mathbf{x}_{,ij} = \mathbf{x}_{,ji} = \mathbf{g}_{j,i}$.

$$\mathbf{a} = [\dot{v}^r + (\Gamma^r_{rr}v^r + \Gamma^r_{r\theta}v^\theta)v^r + (\Gamma^r_{r\theta}v^r + \Gamma^r_{\theta\theta}v^\theta)v^\theta]\mathbf{g}_r$$
$$+ [\dot{v}^\theta + (\Gamma^\theta_{rr}v^r + \Gamma^\theta_{r\theta}v^\theta)v^r + (\Gamma^\theta_{r\theta}v^r + \Gamma^\theta_{\theta\theta}v^\theta)v^\theta]\mathbf{g}_\theta$$
$$\equiv a^r\mathbf{g}_r + a^\theta\mathbf{g}_\theta .$$

$$\mathbf{a} = (\dot{v}^k + v^iv^j\Gamma^k_{ij})\mathbf{g}_k \equiv a^k\mathbf{g}_k . \tag{3.35}$$

We call a^r and a^θ the *roof components of* **a** in the (r, θ) coordinate system.

Things are not really as complicated as they appear because a number of the Christoffel symbols vanish. From (3.25)

$$\mathbf{g}_{r,r} = \mathbf{0}$$
$$\mathbf{g}_{r,\theta} = \mathbf{g}_{\theta,r} = -\sin\theta\mathbf{e}_x + \cos\theta\mathbf{e}_y = r^{-1}\mathbf{g}_\theta$$
$$\mathbf{g}_{\theta,\theta} = -r\cos\theta\mathbf{e}_x - r\sin\theta\mathbf{e}_y = -r\mathbf{g}_r .$$

$$\mathbf{g}_{i,j} = x^p_{,ij}\mathbf{e}_p = x^p{}_{,ij}u^k_{,p}\mathbf{g}_k . \tag{3.36}$$

Comparing (3.33) with the above, we have, by inspection,

$$\Gamma^r_{rr} = \Gamma^\theta_{rr} = \Gamma^r_{r\theta} = \Gamma^\theta_{\theta\theta} = 0$$
$$\Gamma^\theta_{r\theta} = r^{-1},\ \Gamma^r_{\theta\theta} = -r,$$

$$\Gamma^k_{ij} = x^p_{,ij}u^k_{,p}, \tag{3.37}$$

so that the formula for **a** with $v^r = \dot{r}$ and $v^\theta = \dot{\theta}$ reduces to

$$\mathbf{a} = (\ddot{r} - r\dot{\theta}^2)\mathbf{g}_r + (\ddot{\theta} + 2r^{-1}\dot{r}\dot{\theta})\mathbf{g}_\theta = a^r\mathbf{g}_r + a^\theta\mathbf{g}_\theta .$$

$$\mathbf{a} = (\ddot{u}^k + \dot{u}^i\dot{u}^jx^p_{,ij}u^k_{,p})\mathbf{g}_k = a^k\mathbf{g}_k . \tag{3.38}$$

Note that the physical components of the acceleration in the directions of $\mathbf{g}^r$ and $\mathbf{g}^\theta$, are given, respectively, by

$$a^{(r)} = \mathbf{a}\cdot\bar{\mathbf{g}}^r = a^r = \ddot{r} - r\dot{\theta}^2$$
$$a^{(\theta)} = \mathbf{a}\cdot\bar{\mathbf{g}}^\theta = ra^\theta = r\ddot{\theta} + 2\dot{r}\dot{\theta},$$

$$a^{(i)} = \mathbf{a}\cdot\bar{\mathbf{g}}^i = (\ddot{u}^i + \dot{u}^p\dot{u}^q\Gamma^i_{pq})/|\mathbf{g}^i| \tag{3.39}$$

formulas that you will recognize from calculus, physics, and dynamics.

To get the component form of $\mathbf{f} = m\mathbf{a}$, we assume that the roof components of **f** are known. Then Newton's Law may be written as

$$f^r\mathbf{g}_r + f^\theta\mathbf{g}_\theta = m(a^r\mathbf{g}_r + a^\theta\mathbf{g}_\theta).$$

$$f^i\mathbf{g}_i = ma^i\mathbf{g}_i \tag{3.40}$$

But $\{\mathbf{g}_r, \mathbf{g}_\theta\}$ is a basis, which implies that

$$f^r = ma^r = m(\ddot{r} - r\dot{\theta}^2) \qquad \Big| \qquad f^i = m(\ddot{u}^i + \dot{u}^p\dot{u}^q\Gamma^i_{pq})$$
$$f^\theta = ma^\theta = m(\ddot{\theta} + 2r^{-1}\dot{r}\dot{\theta}). \qquad \Big| \tag{3.41}$$

Now the motive for introducing polar coordinates in the first place was the expectation that, in a central force field, the component form of Newton's Law would simplify. And indeed it does. Note that in $(3.41)_2$, $\ddot{\theta} + 2r^{-1}\dot{r}\dot{\theta} = r^{-2}(r^2\dot{\theta})^{\cdot}$. Hence, if $f^\theta = 0$, $r^2\dot{\theta} = r_0$, a constant—Kepler's Law! Combined with $(3.41)_1$, this relation allows us to reduce the determination of the orbit of a mass-point to the solution of a first order differential equation, provided that f^r depends on r only. See Exercise 3.7.

General Three-Dimensional Coordinates

A convenient way to describe the position of a satellite orbiting the earth is to give its distance ρ from the center of the earth, its co-latitude ϕ (i.e., the co-latitude of a point on the earth's surface directly beneath the satellite), and its azimuth θ from a non-rotating plane Π through the polar axis. Such a system of *spherical coordinates* (ρ, ϕ, θ) is defined by

$$\begin{aligned} x &= \rho \sin\phi \cos\theta \\ y &= \rho \sin\phi \sin\theta \\ z &= \rho \cos\phi, \end{aligned} \tag{3.42}$$

where (x,y,z) is a set of Cartesian coordinates with its origin at the center of the earth and the xy- and xz-planes coincident, respectively, with the plane of the equator and the plane Π. With the restrictions $0 \leq \rho$, $0 \leq \theta \leq \pi$, $0 \leq \theta < 2\pi$, the transformation (3.42) becomes 1:1, except along the z-axis which is a locus of exceptional points.

A study of the vibrations of a triclinic crystal in which the atoms of a cell lie at the vertices of a parallelopiped provides another example where a coordinate system other than Cartesian is preferable. Here we might introduce *oblique Cartesian coordinates* (u,v,w) defined by

$$\begin{aligned} x &= A^1_1 u + A^1_2 v + A^1_3 w \\ y &= A^2_1 u + A^2_2 v + A^2_3 w \\ z &= A^3_1 u + A^3_2 v + A^3_3 w, \end{aligned} \tag{3.43}$$

where $-\infty < u,v,w < \infty$. The constants A^i_j would be chosen so that the coordinate curves are parallel to the skewed lines along which the atoms lie. As $\det [A^i_j] \neq 0$, the transformation is 1:1 and there are no exceptional points.

Still other problems (for example the determination of the trajectory of a particle rising in a hot, tall chimney) call for *circular cylindrical coordinates* (r,θ,z), defined by

$$x = r\cos\theta,\ y = r\sin\theta,\ z = z, \tag{3.44}$$

where $0 \leq r$, $0 \leq \theta < 2\pi$, $-\infty < z < \infty$. This transformation is 1:1 except along the z-axis.

Equations (3.42) to (3.44) are each examples of a coordinate transformation of the form

$$x = \hat{x}(u,v,w),\ y = \hat{y}(u,v,w),\ z = \hat{z}(u,v,w). \tag{3.45}$$

Instead of using different letters to denote different coordinates, let us set $x = x^1$, $y = x^2$, $z = x^3$ and $u = u^1$, $v = u^2$, $w = u^3$. (The superscripts represent *indices*, not powers.) Then (3.45) reads

$$\begin{aligned} x^1 &= \hat{x}^1(u^1,u^2,u^3) \\ x^2 &= \hat{x}^2(u^1,u^2,u^3) \\ x^3 &= \hat{x}^3(u^1,u^2,u^3). \end{aligned} \tag{3.46}$$

Each of these relations is of the same form. We emphasize this by rewriting (3.46) as

$$x^i = \hat{x}^i(u^1,u^2,u^3), \quad i = 1,2,3 \tag{3.47}$$

However, there is no need to continually remind ourselves that we are working in 3-dimensions. Indeed, the beauty of tensor notation is that it reveals those relations that hold in any (finite) number of dimensions. Thus we shall drop explicit reference to the range of the free index and so replace (3.47) with

$$x^i = \hat{x}^i(u^1,u^2,u^3). \tag{3.48}$$

One final compression of notation. Let us agree that in the argument of a function, a sequence of variables such as u^1,u^2,u^3 may be replaced by a single symbol, say u^j. The only provision is that the symbol for the index (j in this case) be distinct from the symbols for any other indices in the term in which $\hat{x}^i$ appears. This boils (3.48) down to its essence:

$$x^i = \hat{x}^i(u^j). \tag{3.49}$$

The index j in (3.49) may be called an *argument index*.

Newton's Law in General Coordinates

Newton's law in general coordinates is obtained by mimicking the steps that led us from the expression for the position of a mass-point in polar coordinates, (3.23), to the formula for its acceleration in terms of roof components and cellar base vectors, (3.38). Our first thought is that things are bound to be more elaborate. However, because we refuse to tie ourselves to any particular coordinate system, certain features of our equations that are common to all coordinate systems come into focus, each cluster of formulas exhibiting a simple pattern that is captured by tensor notation.

Several of the equations that follow have a counterpart in polar coordinates, indicated by an equation number in brackets. Each of these latter equations has its generalized tensor form listed on its right. Your respect for the economy of tensor notation will grow as we proceed.

We begin by relabelling the Cartesian base vectors as follows:

$$\mathbf{e}_1 = \mathbf{e}_x, \quad \mathbf{e}_2 = \mathbf{e}_y, \quad \mathbf{e}_3 = \mathbf{e}_z. \tag{3.50}$$

With the change from Cartesian to general coordinates defined by (3.49), the representation for the position of a mass-point takes the form

$$\begin{aligned}\mathbf{x} &= \hat{\mathbf{x}}(u^j) = \hat{x}^1(u^j)\mathbf{e}_1 + \hat{x}^2(u^j)\mathbf{e}_2 + \hat{x}^3(u^j)\mathbf{e}_3 \\ &= \hat{x}^k(u^j)\mathbf{e}_k. \qquad (3.51)\ [3.23]\end{aligned}$$

As $\mathbf{x}$ depends on the u^j and these, in turn, are (unknown) functions of t, we have, by the chain rule,

$$\begin{aligned}\mathbf{v} = \dot{\mathbf{x}} &= (\partial\mathbf{x}/\partial u^1)\dot{u}^1 + (\partial\mathbf{x}/\partial u^2)\dot{u}^2 + (\partial\mathbf{x}/u^3)\dot{u}^3 \\ &= (\partial\mathbf{x}/\partial u^i)\dot{u}^i. \qquad (3.52)\ [3.24]\end{aligned}$$

The three vectors

$$\mathbf{g}_i \equiv \partial\mathbf{x}/\partial u^i = \mathbf{x}_{,i} = x^k_{,i}\mathbf{e}_k \qquad (3.53)\ [3.25]$$

are called *the cellar base vectors of the u^j-coordinate system*. They form a basis at any point P where $J(\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3) \neq 0$ and are tangent to the associated coordinate curves that pass through P. We represent these base vectors by arrows with tails at P, although occasionally, as with the Cartesian base vectors $\mathbf{e}_i$, we shall place the tails at the origin. Fig. 3.4 is typical.

PROBLEM 3.3.

Compute the cellar base vectors and the Jacobian for the system of spherical coordinates (3.42).

SOLUTION.

When dealing with specific three-dimensional coordinate systems, it is frequently more tidy, typographically, to revert to using distinct letters for different coordinates. Thus, with

$$\mathbf{x} = x\mathbf{e}_x + y\mathbf{e}_y + z\mathbf{e}_z$$

we have, from (3.42) and (3.53),

$$\mathbf{g}_1 = \mathbf{g}_\rho = \mathbf{x}_{,\rho} = \sin\phi\cos\theta\,\mathbf{e}_x + \sin\phi\sin\theta\,\mathbf{e}_y + \cos\phi\,\mathbf{e}_z \tag{3.54}$$

$$\mathbf{g}_2 = \mathbf{g}_\phi = \mathbf{x}_{,\phi} = \rho\cos\phi\cos\theta\,\mathbf{e}_x + \rho\cos\phi\sin\theta\,\mathbf{e}_y - \rho\sin\phi\,\mathbf{e}_z \tag{3.55}$$

$$\mathbf{g}_3 = \mathbf{g}_\theta = \mathbf{x}_{,\theta} = -\rho\sin\phi\sin\theta\,\mathbf{e}_x + \rho\sin\phi\cos\theta\,\mathbf{e}_y \tag{3.56}$$

$$J(\mathbf{g}_\rho, \mathbf{g}_\phi, \mathbf{g}_\theta) = \begin{vmatrix} \sin\phi\cos\theta & \sin\phi\sin\theta & \cos\phi \\ \rho\cos\phi\cos\theta & \rho\cos\phi\sin\theta & -\rho\sin\phi \\ -\rho\sin\phi\sin\theta & \rho\sin\phi\cos\theta & 0 \end{vmatrix} = \rho^2\sin\phi. \tag{3.57}$$

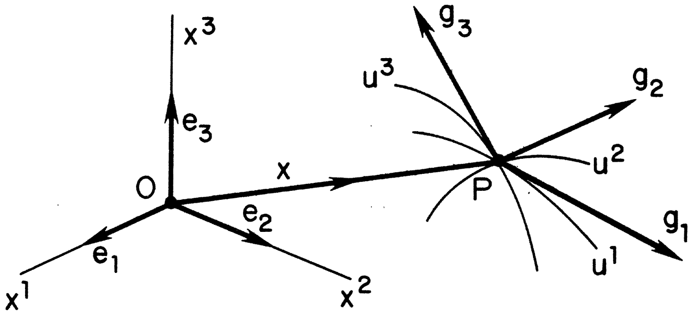


Figure 3.4

The coefficients

$$v^i = \dot{u}^i \qquad (3.58)\ [3.27]$$

of the $\mathbf{g}_i$ in (3.52) are called the *roof* (or contravariant) components of $\mathbf{v}$. In the new notation defined by (3.53) and (3.58), the component representation for the velocity reads

$$\mathbf{v} = v^i \mathbf{g}_i. \qquad (3.59)\ [3.28]$$

The physical components of $\mathbf{v}$ in the directions of the reciprocal base vectors $\mathbf{g}^i$ are, by definition,

$$\mathbf{v}^{(i)} \equiv \mathbf{v}\cdot\bar{\mathbf{g}}^i = \mathbf{v}\cdot\mathbf{g}^i/|\mathbf{g}^i| = v^i/|\mathbf{g}^i|. \qquad (3.60)^9\ [3.30]$$

To compute the acceleration, we differentiate (3.59) with respect to time, so obtaining

$$\mathbf{a} = \dot{\mathbf{v}} = \dot{v}^i\mathbf{g}_i + v^i\dot{\mathbf{g}}_i. \qquad (3.61)\ [3.31]$$

The $\mathbf{g}_i$ are functions of the general coordinates which, along the trajectory, are (unknown) functions of time. Thus, by the chain rule,

$$\begin{aligned}\dot{\mathbf{g}}_i &= (\partial\mathbf{g}_i/\partial u^1)\dot{u}^1 + (\partial\mathbf{g}_i/\partial u^2)\dot{u}^2 + (\partial\mathbf{g}_i/\partial u^3)\dot{u}^3 \\ &= (\partial\mathbf{g}_i/\partial u^j)\dot{u}^j \\ &= v^j\mathbf{g}_{i,j}, \qquad (3.62)\ [3.32]\end{aligned}$$

where we have used, successively, the summation convention, the comma convention for partial differentiation, and (3.58).

The vectors $\mathbf{g}_{i,j}$ can be expressed as linear combinations of the $\mathbf{g}_k$ in the form

$$\mathbf{g}_{i,j} = \Gamma^1_{ij}\mathbf{g}_1 + \Gamma^2_{ij}\mathbf{g}_2 + \Gamma^3_{ij}\mathbf{g}_3 = \Gamma^k_{ij}\mathbf{g}_k. \qquad (3.63)\ [3.33]$$

Note that $\mathbf{g}_{i,j} = \partial^2\mathbf{x}/\partial u^i\partial u^j = \partial^2\mathbf{x}/\partial u^j\partial u^i = \mathbf{g}_{j,i}$, i.e., $\mathbf{g}_{i,j}$ is *symmetric* in the indices i and j. This implies that $\Gamma^k_{ij} = \Gamma^k_{ji}$. The Γ^k_{ij} are called the *Christoffel symbols of the u^j-coordinate system.*

[9] Recall that two dummy indicies on the same level (on the roof in this case) are not to be summed over.

Substituting (3.63) into (3.62) and (3.62) into (3.61), we obtain for the acceleration,

$$\mathbf{a} = \dot{v}^i\mathbf{g}_i + v^i v^j \Gamma^k_{ij}\mathbf{g}_k. \tag{3.64}$$

We now throw a move that is typical in tensor analysis: we change the dummy index in the first term on the right from i to k. This allows us to write

$$\begin{aligned}\mathbf{a} &= \dot{v}^k\mathbf{g}_k + v^i v^j \Gamma^k_{ij}\mathbf{g}_k \\ &= (\dot{v}^k + v^i v^j \Gamma^k_{ij})\mathbf{g}_k \equiv a^k\mathbf{g}_k.\end{aligned} \qquad (3.65)\ [3.35]$$

The a^k are the roof components of $\mathbf{a}$ (with respect to the basis $\{\mathbf{g}_k\}$).

The physical components of $\mathbf{a}$ in the directions of the reciprocal base vectors $\mathbf{g}^i$ are, by definition, and from (3.58) and (3.65),

$$a^{(i)} \equiv \mathbf{a}\cdot\overline{\mathbf{g}}^i = a^i/|\mathbf{g}^i| = (\ddot{u}^i + \dot{u}^p\dot{u}^q\Gamma^i_{pq})/|\mathbf{g}^i|. \qquad (3.66)\ [3.39]$$

When $\mathbf{f}$ is represented in terms of the basis $\{\mathbf{g}_k\}$, Newton's Law, $\mathbf{f} = m\mathbf{a}$, takes the form

$$f^k\mathbf{g}_k = ma^k\mathbf{g}_k \qquad (3.67)\ [3.40]$$

which implies that

$$f^k = ma^k = m(\ddot{u}^k + \dot{u}^i\dot{u}^j\Gamma^k_{ij}). \qquad (3.68)\ [3.41]$$

Computation of the Christoffel Symbols

Computation of the Christoffel symbols in any specified coordinate system is straightforward. Indeed, there exist computer programs that do this by manipulating symbols—not numbers—just as we are about to do.

(a) Start from (3.51) and compute the Cartesian components of $\mathbf{g}_{i,j}$.
(b) Compute the roof base vectors, i.e. find the inverse of the Jacobian matrix.
(c) Take the dot product of both sides of (3.63) with $\mathbf{g}^p$, p being a free index. This yields $\mathbf{g}^p\cdot\mathbf{g}_{i,j} = \Gamma^k_{ij}\mathbf{g}^p\cdot\mathbf{g}_k = \Gamma^k_{ij}\delta^p_k = \Gamma^p_{ij}$. Or, renaming the free index,

$$\Gamma^k_{ij} = \mathbf{g}^k\cdot\mathbf{g}_{i,j}. \tag{3.69}$$

PROBLEM 3.4.

Compute the Christoffel symbols for the u^j-coordinate system defined by

$$\begin{aligned}x^1 &= u^1u^2 \\ x^2 &= (u^3)^2 \\ x^3 &= (u^1)^2 - (u^2)^2.\end{aligned}$$

SOLUTION.

As mentioned earlier, when dealing with specific coordinate systems, it often simplifies the typography to set $x = x^1$, $y = x^2$, . . . , $x^1_{,1} = x_{,u}$, $\mathbf{g}_1 = \mathbf{g}_u$, etc. Thus, recalling our abc's, we have

(a)
$$\mathbf{x} = \hat{x}^i(u^j)\mathbf{e}_i \sim (uv, w^2, u^2 - v^2).$$
$$\mathbf{g}_u \sim (v,0,2u),\ \mathbf{g}_v \sim (u,0,-2v),\ \mathbf{g}_w \sim (0,2w,0)$$
$$\mathbf{g}_{u,u} \sim (0,0,2),\ \mathbf{g}_{u,v} \sim (1,0,0),\ \mathbf{g}_{u,w} \sim (0,0,0)$$
$$\mathbf{g}_{v,v} \sim (0,0,-2),\ \mathbf{g}_{v,w} \sim (0,0,0)$$
$$\mathbf{g}_{w,w} \sim (0,2,0).$$

(b)
$$G = \begin{bmatrix} v & u & 0 \\ 0 & 0 & 2w \\ 2u & -2v & 0 \end{bmatrix} \text{ and hence } J = 4w(u^2 + v^2).$$

Rather than carry out the simultaneous row reduction of $[G|I]$ to obtain $[I|G^{-1}]$, it is simpler, since we are dealing with a 3×3 matrix, to use the formula (no sum on i or j)

$$G^{-1} = J^{-1}[(-1)^{i+j}M_{ji}], \tag{3.70}$$

where M_{ij} is the determinant obtained from G by deleting its ith row and jth column. (See Hildebrand, *Methods of Applied Mathematics*, 2nd Ed., p. 16–17.) Thus

$$G^{-1} = J^{-1}\begin{bmatrix} 4vw & 0 & 2uw \\ 4uw & 0 & -2vw \\ 0 & 2u^2 + 2v^2 & 0 \end{bmatrix} \sim \begin{bmatrix} \mathbf{g}^u \\ \mathbf{g}^v \\ \mathbf{g}^w \end{bmatrix}.$$

(c) From (a) and (b)

$$\Gamma^u_{uu} = \mathbf{g}^u \cdot \mathbf{g}_{u,u} = J^{-1}[(4vw)(0) + (0)(0) + (2uw)(2)]$$
$$= u/(u^2 + v^2)$$
$$\Gamma^u_{uv} = \mathbf{g}^u \cdot \mathbf{g}_{uv} = J^{-1}[(4vw)(1) + (0)(0) + (2uw)(0)]$$
$$= v/(u^2 + v^2).$$

Continuing in this fashion, we find that

$$-\Gamma^u_{vv} = \Gamma^v_{uv} = u/(u^2 + v^2),\ -\Gamma^v_{uu} = \Gamma^v_{vv} = v/(u^2 + v^2)$$
$$\Gamma^w_{ww} = 1/w,$$

while the remaining 11 Christoffel symbols are all zero.

An Alternate Formula for Computing the Christoffel Symbols

An alternate formula for computing the Christoffel symbols is

$$\Gamma^k_{ij} = \tfrac{1}{2} g^{kp}(g_{ip,j} + g_{jp,i} - g_{ij,p}), \tag{3.71}$$

where, as explained in Exercise 2.8, $g^{ij} \equiv \mathbf{g}^i \cdot \mathbf{g}^j$ and $g_{ij} \equiv \mathbf{g}_i \cdot \mathbf{g}_j$ are the roof and cellar components of the identity tensor $\mathbf{I}$. This formula has two advantages over (3.69).

First, it is applicable as it stands to general relativity where the indices range from 1 to 4 and where the g_{ij}—there called the covariant components of the metric tensor—reflect the distribution of matter in the universe.

Second, (3.71) is efficient in *orthogonal* coordinates, i.e., in coordinate systems where the $\mathbf{g}_i$ are mutually $\perp$ (though not necessarily of unit length). In an orthogonal coordinate system, the matrix $[g_{ij}]$ is *diagonal,* as is $[g^{ij}]$. In this event, (3.71) reduces to

$$\Gamma^k_{ij} = \tfrac{1}{2} g^{kk}(g_{ik,j} + g_{jk,i} - g_{ij,k}) \tag{3.72}$$

This formula may now be broken into 4 mutually exclusive cases.

(a) *i, j, and k are distinct.* In this case,

$$\Gamma^k_{ij} = 0,\ i \neq j \neq k \neq i, \tag{3.73}$$

because $[g_{ij}]$ is diagonal. In n-dimensions there are $n(n-1)/2$ Christoffel symbols of this type.

(b) *i and j are distinct but k is not.* In this case, because $\Gamma^k_{ij} = \Gamma^k_{ji}$ and $g_{ij} = 0$,

$$\Gamma^i_{ij} = \tfrac{1}{2} g^{ii} g_{ii,j},\ i \neq j \tag{3.74}$$

In n-dimensions there are $n(n-1)$ Christoffel symbols of this type.

(c) *i and j are equal but distinct from k.* In this case $g_{ik} = g_{jk} = 0$ and (3.71) reduces to

$$\Gamma^k_{ii} = -\tfrac{1}{2} g^{kk} g_{ii,k},\ k \neq i \tag{3.75}$$

In n-dimensions there are $n(n-1)$ Christoffel symbols of this type.

(d) *i, j, and k are all equal.* Then

$$\Gamma^i_{ii} = \tfrac{1}{2} g^{ii} g_{ii,i} \tag{3.76}$$

In n-dimensions there are n Christoffel symbols of this type.

Before deriving (3.71), let us illustrate the above results.

PROBLEM 3.5.

Compute the Christoffel symbols in spherical coordinates.

SOLUTION.

From (3.54) to (3.56),

$$[g_{ij}] = \begin{bmatrix} \mathbf{g}_\rho\cdot\mathbf{g}_\rho & \mathbf{g}_\rho\cdot\mathbf{g}_\phi & \mathbf{g}_\rho\cdot\mathbf{g}_\theta \\ \cdot & \mathbf{g}_\phi\cdot\mathbf{g}_\phi & \mathbf{g}_\phi\cdot\mathbf{g}_\theta \\ \cdot & \cdot & \mathbf{g}_\theta\cdot\mathbf{g}_\theta \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \rho^2 & 0 \\ 0 & 0 & \rho^2 \sin^2\phi \end{bmatrix}.$$

Hence, by Exercise 2.10(b),

$$[g^{ij}] = [g_{ij}]^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \rho^{-2} & 0 \\ 0 & 0 & \rho^{-2}\csc^2\phi \end{bmatrix}.$$

Taking cases (a)–(d) in order, we have

(a) $\Gamma^1_{23} = \Gamma^\rho_{\phi\theta} = 0,\ \Gamma^2_{13} = \Gamma^\phi_{\rho\theta} = 0,\ \Gamma^3_{12} = \Gamma^\theta_{\rho\phi} = 0$

(b) $\Gamma^1_{12} = \Gamma^\rho_{\rho\phi} = \tfrac{1}{2} g^{11} g_{11,2} = \tfrac{1}{2}(1)(0) = 0$

$\Gamma^1_{13} = \Gamma^\rho_{\rho\theta} = \tfrac{1}{2} g^{11} g_{11,3} = \tfrac{1}{2}(1)(0) = 0$

$$\Gamma^2_{21} = \Gamma^\phi_{\phi\rho} = \tfrac{1}{2} g^{22} g_{22,1} = \tfrac{1}{2}(\rho^{-2})(2\rho) = \rho^{-1}$$
$$\Gamma^2_{23} = \Gamma^\phi_{\phi\theta} = \tfrac{1}{2} g^{22} g_{22,3} = \tfrac{1}{2}(\rho^{-2})(0) = 0$$
$$\Gamma^3_{31} = \Gamma^\theta_{\theta\rho} = \tfrac{1}{2} g^{33} g_{33,1} = \tfrac{1}{2}(\rho^{-2} \csc^2 \phi)(2\rho \sin^2 \phi) = \rho^{-1}$$
$$\Gamma^3_{32} = \Gamma^\theta_{\theta\phi} = \tfrac{1}{2} g^{33} g_{33,2} = \tfrac{1}{2}(\rho^{-2} \csc^2 \phi)(2\rho^2 \sin \phi \cos \phi) = \cot \phi$$

(c)
$$\Gamma^1_{22} = \Gamma^\rho_{\phi\phi} = -\tfrac{1}{2} g^{11} g_{22,1} = -\tfrac{1}{2}(1)(2\rho) = -\rho$$
$$\Gamma^1_{33} = \Gamma^\rho_{\theta\theta} = -\tfrac{1}{2} g^{11} g_{33,1} = -\tfrac{1}{2}(1)(2\rho \sin^2 \phi) = -\rho \sin^2 \phi$$
$$\Gamma^2_{11} = \Gamma^\phi_{\rho\rho} = -\tfrac{1}{2} g^{22} g_{11,2} = -\tfrac{1}{2}(\rho^{-2})(0) = 0$$
$$\Gamma^2_{33} = \Gamma^\phi_{\theta\theta} = -\tfrac{1}{2} g^{22} g_{33,2} = -\tfrac{1}{2}(\rho^{-2})(2\rho^2 \sin \phi \cos \phi) = -\sin \phi \cos \phi$$
$$\Gamma^3_{11} = \Gamma^\theta_{\rho\rho} = -\tfrac{1}{2} g^{33} g_{11,3} = -\tfrac{1}{2}(\rho^{-2} \csc^2 \phi)(0) = 0$$
$$\Gamma^3_{22} = \Gamma^\theta_{\phi\phi} = -\tfrac{1}{2} g^{33} g_{22,3} = -\tfrac{1}{2}(\rho^{-2} \csc^2 \phi)(0) = 0$$

(d)
$$\Gamma^1_{11} = \Gamma^\rho_{\rho\rho} = \tfrac{1}{2} g^{11} g_{11,1} = \tfrac{1}{2}(1)(0) = 0$$
$$\Gamma^2_{22} = \Gamma^\phi_{\phi\phi} = \tfrac{1}{2} g^{22} g_{22,2} = \tfrac{1}{2}(\rho^{-2})(0) = 0$$
$$\Gamma^3_{33} = \Gamma^\theta_{\theta\theta} = \tfrac{1}{2} g^{33} g_{33,3} = \tfrac{1}{2}(\rho^{-2} \csc^2 \phi)(0) = 0.$$

While these results are fresh in our minds let us work another problem.

PROBLEM 3.6.

Compute the roof and physical components a^2 and $a^{(2)}$ of the acceleration vector in spherical coordinates.

SOLUTION.

From (3.65),

$$a^2 = \dot{v}^2 + (v^1)^2\Gamma^2_{11} + 2v^1v^2\Gamma^2_{12} + 2v^1v^3\Gamma^2_{13} + (v^2)^2\Gamma^2_{22} + 2v^2v^3\Gamma^2_{23} + (v^3)^2\Gamma^2_{33}.$$

From Problem 3.5 we see that the only non-zero Christoffel symbols in the above expression are Γ^2_{12} $(= \Gamma^2_{21})$ and Γ^2_{33}. Thus, with $a^2 = a^\phi$, etc. we have

$$\begin{aligned} a^\phi &= \dot{v}^\phi + 2v^\rho v^\phi \Gamma^\phi_{\rho\phi} + (v^\theta)^2 \Gamma^\phi_{\theta\theta} \\ &= \ddot{\phi} + 2\rho^{-1}\dot{\rho}\dot{\phi} - \dot{\theta}^2 \sin \phi \cos \phi. \end{aligned}$$

Again, from Problem 3.5, $\mathbf{g}^2 \cdot \mathbf{g}^2 = g^{22} = \rho^{-2}$. Hence $|\mathbf{g}^2| = \rho^{-1}$ and we obtain

$$a^{(\phi)} = a^{(2)} = a^2/|\mathbf{g}^2| = \rho\ddot{\phi} + 2\dot{\rho}\dot{\phi} - \rho\dot{\theta}^2 \sin \phi \cos \phi.$$

Note that the physical component $a^{(\phi)}$, unlike the roof component a^ϕ, has the dimensions of acceleration, [length/(time)^2].

To derive (3.71) note first that

$$g_{ij,k} = (\mathbf{g}_i \cdot \mathbf{g}_j)_{,k} = \mathbf{g}_{i,k} \cdot \mathbf{g}_j + \mathbf{g}_i \cdot \mathbf{g}_{j,k}, \tag{3.77}$$

which, with the definition (3.63) of the Christoffel symbols, takes the form

$$g_{ij,k} = \Gamma^p_{ik}\mathbf{g}_p \cdot \mathbf{g}_j + \Gamma^p_{jk}\mathbf{g}_p \cdot \mathbf{g}_i = \Gamma^p_{ik} g_{pj} + \Gamma^p_{jk} g_{pi}. \tag{3.78}$$

Next interchange i and k and then j and k in (3.78) to get

$$\begin{aligned} g_{kj,i} &= \Gamma^p_{ki} g_{pj} + \Gamma^p_{ji} g_{pk}. \\ g_{ik,j} &= \Gamma^p_{ij} g_{pk} + \Gamma^p_{kj} g_{pi}. \end{aligned} \tag{3.79}$$

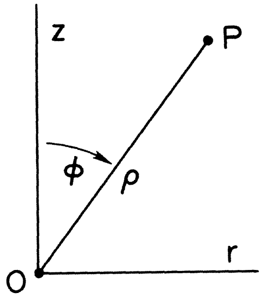


Figure 3.5

Now add (3.78) to $(3.79)_1$ and subtract $(3.79)_2$ from the resulting sum. As $\Gamma^p_{jk} = \Gamma^p_{kj}$, etc., there follows

$$g_{ij,k} + g_{kj,i} - g_{ik,j} = 2\Gamma^p_{ik} g_{pj}. \tag{3.80}$$

Finally, multiply both sides by $\frac{1}{2} g^{qj}$ and sum on the repeated index j. As $g^{qj} g_{pj} = \delta^q_p$, the resulting expression reduces to

$$\Gamma^q_{ik} = \tfrac{1}{2} g^{qj}(g_{ij,k} + g_{kj,i} - g_{ik,j}), \tag{3.81}$$

which is just (3.71) with the indices relabeled.

A Change of Coordinates

A change of coordinates from a u^j-system to, say, a $\tilde{u}^k$-system is defined by a transformation of the form

$$u^j = \hat{u}^j(\tilde{u}^k). \tag{3.82}$$

In what follows we shall assume that (3.82) is 1 : 1 and continuously differentiable except, possibly, at certain exceptional points. For example, a change from circular cylindrical coordinates (r, θ, z) to spherical coordinates (ρ, ϕ, θ), as indicated in Fig. 3.5, takes the form

$$r = \rho \sin \phi \tag{3.83}$$

$$\theta = \theta \tag{3.84}$$

$$z = \rho \cos \phi, \tag{3.85}$$

where $0 \leq \rho$, $0 \leq \phi \leq \pi$, $0 \leq \theta < 2\pi$, $\rho = 0$ being the only exceptional point.

Suppose that the various components of a vector **v** or a second order tensor **T** are known in the u^j-system. How do we compute the corresponding components in the $\tilde{u}^k$-system? As we saw in (2.36) and (2.38), this can be done once we have expressed the base vectors $\tilde{\mathbf{g}}_j = \partial \mathbf{x}/\partial \tilde{u}^j$ as a linear combination of the base vectors $\mathbf{g}_i = \partial \mathbf{x}/\partial u^i$ or *vice versa*. But if, initially, $\mathbf{x} = \hat{\mathbf{x}}(u^j)$, then by the chain rule

$$\tilde{\mathbf{g}}_j = (\partial \mathbf{x}/\partial u^1)\partial u^1/\partial \tilde{u}^j + (\partial \mathbf{x}/\partial u^2)\partial u^2/\partial \tilde{u}^j + \ldots$$
$$= \partial u^i/\partial \tilde{u}^j \mathbf{g}_i. \qquad (3.86)\ [3.36][10]$$

Thus (3.86) is of the same form as (2.31) with

$$A^i_j = \partial u^i/\partial \tilde{u}^j. \qquad (3.87)$$

Of course, now, in general, *the A^i_j vary from point to point*. It is also apparent from symmetry—just reverse the roles of u^j and $\tilde{u}^k$—that

$$(A^{-1})^i_j = \partial \tilde{u}^i/\partial u^j. \qquad (3.88)$$

It therefore follows immediately from (2.36) and (2.38) that

$$\tilde{v}_j = (\partial u^i/\partial \tilde{u}^j)v_i, \quad \tilde{v}^i = (\partial \tilde{u}^i/\partial u^j)v^j \qquad (3.89)$$

$$\begin{aligned} \tilde{T}_{ij} &= (\partial u^k/\partial \tilde{u}^i)(\partial u^p/\partial \tilde{u}^j)T_{kp} \\ \tilde{T}^i_{\cdot j} &= (\partial \tilde{u}^i/\partial u^k)(\partial u^p/\partial \tilde{u}^j)T^k_{\cdot p} \\ \tilde{T}_j^{\cdot i} &= (\partial u^k/\partial \tilde{u}^j)(\partial \tilde{u}^i/\partial u^p)T_k^{\cdot p} \\ \tilde{T}^{ij} &= (\partial \tilde{u}^i/\partial u^k)(\partial \tilde{u}^j/\partial u^p)T^{kp}. \end{aligned} \qquad (3.90)$$

These formulas are important enough to be boxed. Many texts on tensor analysis virtually start by taking (3.89) and (3.90) as the *definitions* of co- and contravariant vectors and co-, contravariant and mixed tensors of 2nd order. That is, any collection of objects $(v_1, v_2, \ldots)$, or v_i for short, is said to represent a *covariant vector* providing its components transform according to $(3.89)_1$ with similar definitions for contravariant vectors v^i, covariant tensors T_{ij}, etc. This viewpoint is taken sometimes in general relativity and in the theory of shells and membranes because in both cases one is dealing with *curved* continua.

PROBLEM 3.7.

Suppose that in circular cylindrical coordinates, a 2nd order tensor **T** has the mixed components

$$[T_j^{\cdot i}] = \begin{bmatrix} 2 & -1 & 1 \\ 0 & 1 & 2 \\ 3 & 0 & -2 \end{bmatrix}$$

at the point $r = 1$, $\theta = \pi/4$, $z = -\sqrt{3}$. Find the component $\tilde{T}_2^{\cdot 1}$ of **T** in spherical coordinates.

SOLUTION.

With $(u^1, u^2, u^3) = (r, \theta, z)$ and $(\tilde{u}^1, \tilde{u}^2, \tilde{u}^3) = (\rho, \phi, \theta)$, we have, from (3.83) to (3.85),

[10] To obtain the 2nd line of (3.36) from (3.86), take $i = k$, $j = p$ and $\tilde{u}^p = x^p$.

$$[\partial u^k/\partial \bar{u}^j] = \begin{bmatrix} \partial r/\partial \rho & \partial r/\partial \phi & \partial r/\partial \theta \\ \partial \theta/\partial \rho & \partial \theta/\partial \phi & \partial \theta/\partial \theta \\ \partial z/\partial \rho & \partial z/\partial \phi & \partial z/\partial \theta \end{bmatrix} = \begin{bmatrix} \sin\phi & \rho\cos\phi & 0 \\ 0 & 0 & 1 \\ \cos\phi & -\rho\sin\phi & 0 \end{bmatrix}.$$

Either by inverting this matrix or else by inverting (3.83) and (3.85) and then computing $\partial\rho/\partial r$, etc. we obtain:

$$[\partial \bar{u}^i/\partial u^p] = \begin{bmatrix} \partial \rho/\partial r & \partial \rho/\partial \theta & \partial \rho/\partial z \\ \partial \phi/\partial r & \partial \phi/\partial \theta & \partial \phi/\partial z \\ \partial \theta/\partial r & \partial \theta/\partial \theta & \partial \theta/\partial z \end{bmatrix} = \begin{bmatrix} \sin\phi & 0 & \cos\phi \\ \rho^{-1}\cos\phi & 0 & -\rho^{-1}\sin\phi \\ 0 & 1 & 0 \end{bmatrix}.$$

From $(3.90)_3$,

$$\begin{aligned} \bar{T}_{2}^{\cdot 1} &= (\partial \bar{u}^1/\partial u^p)[T_k^{\cdot p}(\partial u^k/\partial \bar{u}^2)] \\ &= (\partial \bar{u}^1/\partial u^1)[T_1^{\cdot 1}(\partial u^1/\partial \bar{u}^2) + T_2^{\cdot 1}(\partial u^2/\partial \bar{u}^2) + \dots] \\ &\quad + (\partial \bar{u}^1/\partial u^2)[T_1^{\cdot 2}(\partial u^1/\partial \bar{u}^2) + \dots \quad] \\ &\quad + (\partial \bar{u}^1/\partial u^3)[\dots \quad + T_3^{\cdot 3}(\partial u^3/\partial \bar{u}^2)] \\ &= (\sin\phi)[T_1^{\cdot 1}(\rho\cos\phi) + T_3^{\cdot 1}(-\rho\sin\phi)] \\ &\quad + (\cos\phi)[T_1^{\cdot 3}(\rho\cos\phi) + T_3^{\cdot 3}(-\rho\sin\phi)]. \end{aligned}$$

At the given point, $\rho = 2$, $\sin\phi = \frac{1}{2}$, and $\cos\phi = -\sqrt{3}/2$. Hence,

$$\bar{T}_{2}^{\cdot 1} = \left(\frac{1}{2}\right)[(2)(-\sqrt{3}) + (1)(-1)] + (-\sqrt{3}/2)[(3)(-\sqrt{3}) + (-2)(-1)] = 4 - 2\sqrt{3}.$$

Transformation of the Christoffel Symbols

The transformation of the Christoffel symbols is a bit elaborate but straightforward. By definition

$$\bar{\Gamma}^i_{jk} = \bar{\mathbf{g}}^i \cdot (\partial \bar{\mathbf{g}}_j/\partial \bar{u}^k). \tag{3.91}$$

From (2.33) and (3.88),

$$\bar{\mathbf{g}}^i = (\partial \bar{u}^i/\partial u^p)\mathbf{g}^p. \tag{3.92}$$

Furthermore, from (3.86),

$$\begin{aligned} \frac{\partial \bar{\mathbf{g}}_j}{\partial \bar{u}^k} &= \frac{\partial}{\partial \bar{u}^k}\left(\frac{\partial u^q}{\partial \bar{u}^j}\mathbf{g}_q\right) \\ &= \frac{\partial^2 u^q}{\partial \bar{u}^j \partial \bar{u}^k}\mathbf{g}_q + \frac{\partial u^q}{\partial \bar{u}^j}\frac{\partial \mathbf{g}_q}{\partial \bar{u}^k}. \end{aligned} \tag{3.93}$$

But the base vectors $\mathbf{g}_q$ are functions of the u^r which, in turn, are functions of the $\bar{u}^k$ according to the given change of coordinates (3.82). Hence, by the chain rule and the definition of the Christoffel symbols,

$$\frac{\partial \mathbf{g}_q}{\partial \bar{u}^k} = \frac{\partial \mathbf{g}_q}{\partial u^r}\frac{\partial u^r}{\partial \bar{u}^k}$$

$$= \frac{\partial u^r}{\partial \tilde{u}^k} \Gamma^s_{qr} \mathbf{g}_s. \qquad (3.94)$$

Substituting (3.94) into (3.93), and (3.93) and (3.92) into (3.91) and noting that $\mathbf{g}^p \cdot \mathbf{g}_q = \delta^p_q$ and $\mathbf{g}^p \cdot \mathbf{g}_s = \delta^p_s$, we obtain

$$\tilde{\Gamma}^i_{jk} = \frac{\partial \tilde{u}^i}{\partial u^p} \left(\underline{\frac{\partial^2 u^p}{\partial \tilde{u}^j \tilde{u}^k}} + \frac{\partial u^q}{\partial \tilde{u}^j} \frac{\partial u^r}{\partial \tilde{u}^k} \Gamma^p_{qr} \right). \qquad (3.95)$$

Because of the underlined term, the new Christoffel symbols are *more* than just a linear combination of the old. This is why the Christoffel *cannot* be regarded as the components of a 3rd order tensor (and therefore why some authors write $\{^{\ i}_{jk}\}$ instead of Γ^i_{jk}). Whenever Cartesian coordinates can be introduced[11], (3.95) provides a useful, alternate way of computing the Christoffel symbols—take the old coordinates u^i as a set of Cartesian coordinates x^i and the new coordinates $\tilde{u}^i$ as a set of general coordinates u^i. Since the Christoffel symbols vanish in any set of Cartesian coordinates, (3.95) reduces to

$$\Gamma^i_{jk} = \frac{\partial u^i}{\partial x^p} \frac{\partial^2 x^p}{\partial u^j \partial u^k}. \qquad (3.96)\ [3.37]$$

Exercises

3.1. Cite three different sources that define *inertial frame*. Then give a definition in your own words.

3.2. A ball-bearing of mass m is shot into the air vertically from a spring-loaded cannon whose muzzle is flush with the ground. The spring is linear with spring-constant k and is retracted a distance D from the muzzle. If the spring imparts all of its stored energy to the ball-bearing, how high does the ball-bearing fly? Neglect air drag and take the force of gravity to be constant.

3.3. If $\mathbf{x} = t\mathbf{e}_x + t^2\mathbf{e}_y + t^3\mathbf{e}_z$, $-\infty < t < \infty$, find: a) $\dot{\mathbf{x}}$; b) $\ddot{\mathbf{x}}$.

3.4. Let

$$\mathbf{x} = \hat{\mathbf{x}}(t),\ \alpha \le t \le \beta$$

be the parametric equation of a curve C. If $\mathbf{x}$ is differentiable, the velocity $\mathbf{v} = \dot{\mathbf{x}}$ may, in view of (3.12), be represented in the form

$$\mathbf{v} = |\mathbf{v}|\mathbf{t} = \dot{s}\mathbf{t}. \qquad (3.97)$$

$\mathbf{t}$ is called *the unit tangent to C*. (At points along C where $\mathbf{v} = \mathbf{0}$, $\mathbf{t}$ is undefined, but the representation (3.97) remains valid since, at such points, $\mathbf{v} = 0\mathbf{t} = \mathbf{0}$.) If $\dot{s}$ and $\mathbf{t}$ are differentiable functions of t, then

$$\ddot{\mathbf{x}} = \dot{\mathbf{v}} = \mathbf{a} = \ddot{s}\mathbf{t} + \dot{s}\dot{\mathbf{t}} \equiv \mathbf{a}_p + \mathbf{a}_c. \qquad (3.98)$$

[11]Not possible on intrinsically curved continua such as balloons.

$\mathbf{a}_p$ is called the *path acceleration* and $\mathbf{a}_c$ the *centripetal acceleration*. Note that if we differentiate both sides of the identity $\mathbf{t}\cdot\mathbf{t} = 1$, we obtain $\dot{\mathbf{t}}\cdot\mathbf{t} + \mathbf{t}\cdot\dot{\mathbf{t}} = 0$. Thus $\mathbf{t}\cdot\dot{\mathbf{t}} = 0$, which implies that $\mathbf{a}_p$ and $\mathbf{a}_c$ are orthogonal.

(a). Show that

$$\mathbf{a}_p = (\ddot{\mathbf{x}}\cdot\dot{\mathbf{x}}/\dot{\mathbf{x}}\cdot\dot{\mathbf{x}})\dot{\mathbf{x}}, \quad \mathbf{a}_c = \dot{\mathbf{x}} \times (\ddot{\mathbf{x}} \times \dot{\mathbf{x}})/\dot{\mathbf{x}}\cdot\dot{\mathbf{x}}. \tag{3.99}$$

(b). Compute $\mathbf{a}_p$ and $\mathbf{a}_c$ for the curve in Exercise 3.3.

If we regard the unit tangent $\mathbf{t}$ as a function of arc length s, then we may always represent $\mathbf{t}' = d\mathbf{t}/ds$ in the form

$$\mathbf{t}' = |\mathbf{t}'|\bar{\mathbf{t}}' \equiv \kappa\mathbf{n}. \tag{3.100}$$

κ is called *the curvature of C* and $\mathbf{n}$, which is the direction of $\mathbf{t}'$, the *unit normal to C*.

(c). Using the chain rule, show that

$$\mathbf{a}_c = \kappa\dot{s}^2\mathbf{n}$$

and, from (a), that

$$\kappa = |\dot{\mathbf{x}} \times \ddot{\mathbf{x}}|/|\dot{\mathbf{x}}|^3. \tag{3.101}$$

(d). Compute κ for the curve in Exercise 3.3.

(e). The unit vector $\mathbf{b} = \mathbf{t} \times \mathbf{n}$ is called the *binormal to C*.

Show that $\mathbf{b}'(= d\mathbf{b}/ds)$ and $\mathbf{n}'$ may be expressed in the form

$$\mathbf{b}' = -\tau\mathbf{n}. \tag{3.102}$$

$$\mathbf{n}' = -\kappa\mathbf{t} + \tau\mathbf{b}. \tag{3.103}$$

The scalar τ is called the *torsion of C*.

(f). Show that

$$\tau = (\dot{\mathbf{x}} \times \ddot{\mathbf{x}})\cdot\dddot{\mathbf{x}}/|\dot{\mathbf{x}} \times \ddot{\mathbf{x}}|^2.$$

(g). Compute τ for the curve in Exercise 3.3

The three linear, first order, vector-valued differential equations (3.100), (3.102), and (3.103) are called the *Serret–Frenet equations* and are fundamental in the study of space curves. Given κ and τ as functions of arc length, the solution of the Serret-Frenet equations determine a curve uniquely to within a rigid body movement. See Struik, *Differential Geometry*, 2nd Ed., Sect. 1-8.

3.5. The parametric equation

$$H: \mathbf{x} = a(\cos\theta\mathbf{e}_x + \sin\theta\mathbf{e}_y) + b\theta\mathbf{e}_z, \quad 0 \le \theta \le 2\pi,$$

represents a segment of a *right-handed circular helix* that lies on and wraps once around a right circular cylinder of radius a.

(a). Compute the arc length s along H as a function of θ, starting at $\theta = 0$.

(b). Compute the curvature and torsion of H.

3.6. A rigid whirling arm with a tilting seat at its end may be used to reproduce accelerations that a person might feel on a ride in an amusement park or in an air or spacecraft. With reference to Fig. 3.6, the location of the center of the head of the person seated in the tilting chair is

$$\mathbf{x} = (a + h\sin\phi)\mathbf{e}_r + h\cos\phi\mathbf{e}_z,$$

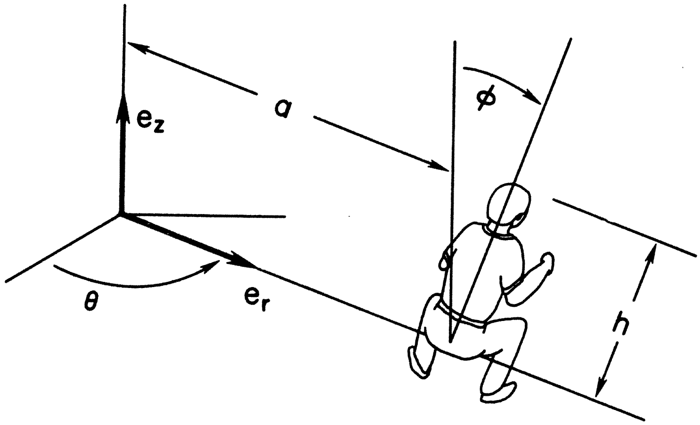


Figure 3.6

where a, h, and $\mathbf{e}_z$ are constant but ϕ and $\mathbf{e}_r$ may vary with time. Let $\{\mathbf{e}_u, \mathbf{e}_v, \mathbf{e}_\theta\}$ be a rigid frame attached to the chair, where

$$\mathbf{e}_u = \mathbf{e}_r \cos\phi - \mathbf{e}_z \sin\phi, \quad \mathbf{e}_v = \mathbf{e}_r \sin\phi + \mathbf{e}_z \cos\phi, \ \mathbf{e}_\theta = \mathbf{e}_z \times \mathbf{e}_r.$$

The acceleration of the head may be expressed as

$$\mathbf{a} = a_u \mathbf{e}_u + a_v \mathbf{e}_v + a_\theta \mathbf{e}_\theta.$$

Express a_u, a_v, and a_θ in terms of θ, ϕ and their time derivatives.

3.7. Carry out the following steps to reduce the determination of the orbit of a mass-point in a central force field to two successive quadratures:
(a). Use $r^2\dot{\theta} = r_0$ to reduce $(3.41)_1$ to a 2nd order differential equation in r.
(b). Multiply the equation so obtained by $\dot{r}$, assume that $f^r = v'(r)$, and show that the resulting equation can be cast into the form $(\quad)^{\cdot} = 0$.
(c). Integrate, solve for $\dot{r}$, and write the resulting equation in separable form, thereby reducing the solution for r to a quadrature.
(d). With $r = \hat{r}(t)$ in hand, how would you find $\theta = \hat{\theta}(t)$?

3.8. Compute the roof base vectors $\{\mathbf{g}^1, \mathbf{g}^2, \mathbf{g}^3\} = \{\mathbf{g}^\rho, \mathbf{g}^\phi, \mathbf{g}^\theta\}$ in spherical coordinates (ρ, ϕ, θ). (Hint: Since the cellar base vectors are mutually $\perp$, $\mathbf{g}^\rho = \lambda \mathbf{g}_\rho$, etc.) Check your answers by computing the matrix $[g^{ij}] = [\mathbf{g}^i \cdot \mathbf{g}^j]$ and comparing your result with the same matrix computed in the Solution to Problem 3.5.

3.9. For the (u,v,w) coordinate system defined by

$$\begin{aligned} x &= u + w \\ y &= v^2 - w \\ z &= u^2 + v, \end{aligned}$$

compute:
(a). The cellar base vectors

$$\mathbf{g}_u, \mathbf{g}_v, \mathbf{g}_w.$$

(b). The reciprocal base vectors at

$$u = -1, v = 1, w = -1.$$

(c). The Christoffel symbols at any point.
(d). The roof components of the acceleration vector.
(e). The physical components $a^{(u)}$, $a^{(v)}$ at $u = -1$, $v = 1$, $w = -1$.

3.10. Assume that we are in two dimensions. Given the transformation

$$\begin{aligned} x &= u - v^2, \\ y &= u + v, \end{aligned} \quad , \quad -\infty < u < \infty, -\frac{1}{2} < v,$$

compute

(a). The base vectors $\mathbf{g}_1 = \mathbf{g}_u$, $\mathbf{g}_2 = \mathbf{g}_v$.
(b). The reciprocal base vectors.
(c). The 6 Christoffel symbols $\Gamma^1_{11} = \Gamma^u_{uu}$, $\Gamma^1_{12} = \Gamma^u_{uv}$, etc.
(d). The roof components $a^u = a^1$ and $a^v = a^2$ of the acceleration vector.
(e). The physical components $a^{(u)}$ and $a^{(v)}$.
(f). If $\mathbf{f} = uv\mathbf{g}_v$, write down the *component form* of Newton's Law in the uv-coordinate system.

3.11. By differentiating both sides of $\mathbf{g}^i \cdot \mathbf{g}_j = \delta^i_j$, show that

$$\mathbf{g}^i_{,j} = -\Gamma^i_{jk}\mathbf{g}^k. \tag{3.104}$$

3.12. Let C be a smooth curve having, in general coordinates, the parametric equation C: $u^i = \hat{u}^i(t)$, $a \leq t \leq b$. By definition, the square of the differential element of arc length is $ds^2 = d\mathbf{x} \cdot d\mathbf{x}$, where $\mathbf{x} = \hat{\mathbf{x}}(\hat{u}^i(t))$ is the position of a point on C. By the chain rule,

$$d\mathbf{x} = (\partial \mathbf{x}/\partial u^i)(du^i/dt)dt = \mathbf{g}_i \dot{u}^i dt,$$

so

$$\begin{aligned} d\mathbf{x} \cdot d\mathbf{x} &= (\mathbf{g}_i \dot{u}^i dt) \cdot (\mathbf{g}_j \dot{u}^j dt) \\ &= g_{ij} \dot{u}^i \dot{u}^j dt^2. \end{aligned} \tag{3.105}$$

The distance along C between two points $P(t_1)$ and $P(t_2)$ is therefore equal to $\int_{t_2}^{t_1} \sqrt{g_{ij}\dot{u}^i\dot{u}^j}\, dt$. Because of this relation, the g_{ij}, especially in general relativity, are sometimes referred to, collectively, as the *metric tensor*.

Let $\theta = \phi = t$, $0 \leq t \leq \pi$ denote the parametric equations of a curve lying on a sphere of radius R. Express the length of C as a definite integral.

3.13. Compute the roof components of the vector $\mathbf{v} \sim (x^2 + y^2, z, -2)$ in spherical coordinates, (3.42).

3.14. Compute the 6 Christoffel symbols of the two-dimensional uv-coordinate system defined by

$$x = 2e^{u-v}, y = -e^{3u+2v}.$$

3.15. Generalize Exercise 3.14: Determine formulas (using matrix notation if you prefer) for the Christoffel symbols of the u^j-coordinate defined by

$$x^i = c^i \exp (A^i_j u^j),$$

where $[A^i_j]$ is a constant matrix. (Recall: no sum on i on the right because i appears twice on the roof.)

3.16. (a). Determine the transformation $x^i = \hat{x}^i(u^j)$ such that $\Gamma^i_{jk} = \alpha_k \delta^i_j$, where the α_k are constants.

(b). Does there always exist a transformation of the form $x^i = \hat{x}^i(u^j)$ such that the Γ^i_{jk} take on arbitrarily prescribed values? (Hint: First try special two-dimensional examples.)

3.17. *A class of 3-dimensional cylindrical coordinates* is defined by a transformation of the form[12]

$$x + iy = f(u + iv)$$
$$z = w,$$

where f is an analytic function. Recalling that

$$f'(u + iv) = x_{,u} + iy_{,u} = y_{,v} - ix_{,v},$$

show that

(a). $\mathbf{g}_u - i\mathbf{g}_v = (\mathbf{e}_x - i\mathbf{e}_y)f'$.

(b). the coordinate system is $\perp$.

(c). $J = |f'|^2$.

(d). $\mathbf{g}_{u,u} - i\mathbf{g}_{v,u} = (\mathbf{g}_u - i\mathbf{g}_v)f''/f'$, $\mathbf{g}_{u,v} - i\mathbf{g}_{v,v} = (\mathbf{g}_v + i\mathbf{g}_u)f''/f$.

(e). $\Gamma^u_{uu} = \Gamma^v_{uv} = -\Gamma^u_{vv} = \mathcal{R}(f''/f')$, $\Gamma^v_{uu} = -\Gamma^u_{uv} = -\Gamma^v_{vv} = \mathcal{I}(f''/f')$,

where $\mathcal{R}$ denotes "The real part of" and $\mathcal{I}$ denotes "The imaginary part of". Compute the Christoffel symbols and sketch a few of the coordinate lines in the plane $z = 0$ for

(f). *parabolic cylindrical coordinates,* defined by

$$f = \tfrac{1}{2}(u + iv), \ -\infty < u < \infty, \ v \geq 0.$$

(g). *elliptic cylindrical coordinates,* defined by

$$f = \cosh(u + iv), \ u \geq 0, \ 0 \leq v < 2\pi.$$

(h). *bipolar cylindrical coordinates,* defined by

$$f = \coth(u + iv), \ -\infty < u < \infty, \ 0 < v < \pi.$$

3.18. Express Newton's Law in component form in elliptic cylindrical coordinates.

[12] z is the 3rd Cartesian coordinate and *not* the complex variable $x + iy$.

CHAPTER IV

The Gradient Operator, Covariant Differentiation, and the Divergence Theorem

Suppose that you had a topographical map of a piece of land and wanted to indicate at a spot P on the map the slope m of the land in a direction $\mathbf{t}$. This could be done by drawing a vector $m\mathbf{t}$ from P, as indicated in Fig. 4.1. Obviously, if the terrain is smooth but not level, there is one direction from P in which the slope is a maximum. This is called *the direction of steepest ascent.*[1] The associated vector is called the *gradient* of the elevation at P. If you draw a contour line through P, you will realize that the gradient at P must be $\perp$ to this contour.

The Gradient

This 2-dimensional example suggests the following definition in n-dimensions. Let $f(\mathbf{x})$ be given in some region of E_n.[2] *The gradient of f at* $\mathbf{x}$ *is the vector having the direction and magnitude of the maximum increase of f with respect to distance away from* $\mathbf{x}$. We shall denote the gradient of f by ∇f and represent it by an arrow with its tail at P. When we wish to emphasize that ∇f is a function of position, we shall write $\nabla f(\mathbf{x})$.

The gradient arises in many physical contexts. In the theory of heat transfer, Fourier's law of heat conduction for a materially isotropic body states that the heat flux at a point is proportional to the negative of the temperature gradient at that point: $\mathbf{q} = -k\nabla T,\ k > 0$. In ideal fluid flow (no vorticity), the

[1] At a dome-like peak, all directions are directions of (zero) steepest ascent. Cone-like peaks are excluded by our assumption of smoothness.

[2] Such a function is sometimes called a *scalar field*. Likewise, vector and tensor functions may be referred to as *vector and tensor fields*, respectively.

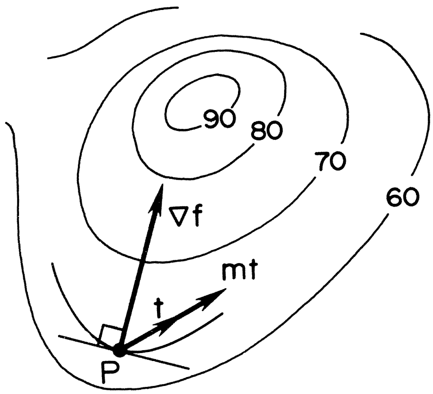


Figure 4.1

velocity at any point in the fluid is equal to the gradient of a potential: $\mathbf{v} = \nabla\phi$. And in the torsion of prismatic bars, the shear stress vector acting at any point on a cross section is equal to the cross product of a unit vector $\mathbf{k}$ along the axis of the bar with the gradient of a stress function: $\boldsymbol{\tau} = \mathbf{k} \times \nabla\psi$.

The first aim of this chapter is to express ∇f in a concise form, valid in *any* coordinate system. To this end, it is convenient to introduce the gradient indirectly, as follows.

Let a set of general coordinates be defined by a transformation of the form $x^i = \hat{x}^i(u^j)$. The position $\mathbf{x}$ of a point P is then a function of the u^j's and, therefore, so is f. Further, let C be a smooth curve having the parametric representation

$$C\colon u^j = \hat{u}^j(s),\ a \le s \le b, \tag{4.1}$$

where s is arc length. For example, in circular cylindrical coordinates, the helix H of Exercise 3.5 is described by

$$H\colon r = a,\ \theta = s/c,\ z = bs/c,\ 0 \le s \le 2\pi c,$$

where a and b are given constants and $c = \sqrt{a^2 + b^2}$.

Along C, f is a function of s via the u^j's. Assuming that the partial derivatives of f with respect to the u^j's exist, we may compute the rate of change of f with respect to distance along C by the chain rule:

$$\begin{aligned} \frac{df}{ds} &= \frac{\partial f}{\partial u^1}\frac{du^1}{ds} + \frac{\partial f}{\partial u^2}\frac{du^2}{ds} + \ldots \\ &= \frac{\partial f}{\partial u^i}\frac{du^i}{ds}\,. \end{aligned} \tag{4.2}$$

The left side of (4.2) is a *scalar invariant* called the *directional* or *path derivative* of f. It depends on $f(\mathbf{x})$ and C, but not on the coordinate system used to specify $\mathbf{x}$ and C—if f is temperature and C represents a mountain trail, I need only a thermometer and a pedometer to measure df/ds. On the other hand, the terms on the right side of (4.2) are evidently coordinate-bound. To interpret these terms geometrically, let $\hat{\mathbf{x}}(s)$ denote the position of points on C and recall from the preceding chapter that $\mathbf{t} = \hat{\mathbf{x}}'(s)$ is a unit tangent to C. As $x^i = \hat{x}^i(u^j)$ and, on C, $u^j = \hat{u}^j(s)$ we have, by the chain rule,

$$\mathbf{t} = (\partial \mathbf{x}/\partial u^i)(du^i/ds) \equiv \mathbf{g}_i t^i. \tag{4.3}$$

Thus, with our short-hand notation for partial derivatives, (4.2) takes the form

$$df/ds = f_{,i} t^i. \tag{4.4}$$

A glance at (2.10) shows that the right side of (4.4) is simply the dot product of $\mathbf{t}$ with the vector

$$\nabla \mathbf{f} \equiv f_{,i} \mathbf{g}^i, \tag{4.5}$$

i.e.,

$$df/ds = \nabla \mathbf{f} \cdot \mathbf{t}. \tag{4.6}$$

To justify our presumptuousness in denoting $f_{,i}\mathbf{g}^i$ by ∇f, we note that if the right side of (4.4) is evaluated at a fixed point P_*, then only the t^i's change when one curve through P_* is replaced by another. $f_{,i}$ at P_* is merely the partial derivative of f with respect to u^i evaluated at u^i_*, the ith coordinate of P_*. That is, df/ds at P_* is a function of $\mathbf{t}$ but not ∇f. Because $df/ds = \nabla f \cdot \mathbf{t}$, df/ds is a maximum when $\mathbf{t}$ is aligned with ∇f and, because $|\mathbf{t}| = 1$, the value of this maximum is just $|\nabla f|$. Thus ∇f fits the geometric definition of the gradient given at the beginning of the chapter.

PROBLEM 4.1.

If

$$f = xy + yz + zx,$$

compute ∇f and $|\nabla f|$ at (12, 5, −9). Compute the corresponding cellar components of ∇f in circular cylindrical coordinates.

SOLUTION.

Noting that

$$f_{,x} = y + z, f_{,y} = x + z, f_{,z} = y + x,$$

and evaluating these partial derivatives at (12, 5, −9), we find that

$$\nabla f \sim (-4, 3, 17), |\nabla f| = \sqrt{(-4)^2 + (3)^2 + (17)^2} = \sqrt{314}.$$

In circular cylindrical coordinates.

$$f = r^2 \sin\theta \cos\theta + rz \sin\theta + rz \cos\theta.$$

Hence, by (4.5), the cellar components of ∇f are

$$f_{,r} = 2r \sin\theta \cos\theta + z \sin\theta + z \cos\theta$$
$$f_{,\theta} = r^2 \cos^2\theta - r^2 \sin^2\theta + rz \cos\theta - rz \sin\theta$$

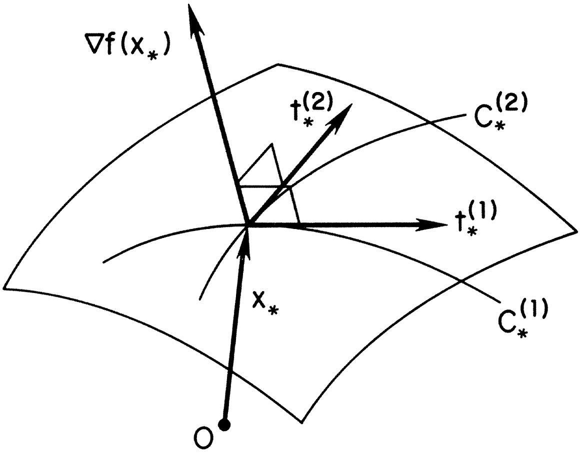


Figure 4.2

$$f_{,z} = r \sin\theta + r\cos\theta.$$

The point P_* with Cartesian coordinates (12, 5, −9) has circular cylindrical coordinates (13, $\tan^{-1} 5/12$, −9), so that at P_*,

$$f_{,r} = (2)(13)(5/13)(12/13) + (-9)(5/13) + (-9)(12/13) = -33/13$$
$$f_{,\theta} = 56, f_{,z} = 17.$$

Alternatively, we can apply $(3.89)_1$ with $v_i = f_{,i}$. In this case $(3.89)_1$ reduces to the chain rule:

$$\frac{\partial f}{\partial \bar{u}^j} = \frac{\partial u^i}{\partial \bar{u}^j}\frac{\partial f}{\partial u^i}\,.$$

With $(\bar{u}^1, \bar{u}^2, \bar{u}^3) = (r, \theta, z)$ and $(u^1, u^2, u^3) = (x, y, z)$, it follows that at P_*

$$\begin{aligned}\bar{v}^1 = f_{,r} &= f_{,x}(\partial x/\partial r) + f_{,y}(\partial y/\partial r) + f_{,z}(\partial z/\partial r)\\ &= f_{,x}\cos\theta + f_{,y}\sin\theta\\ &= (-4)(12/13) + (3)(5/13) = -33/13,\end{aligned}$$

etc.

Equation (4.6) reveals another important property of the gradient: If f is smooth in a neighborhood of a fixed point $P_*(\mathbf{x}_*)$, then $\nabla f(\mathbf{x}_*)$ is ⊥ to the surface $f(\mathbf{x}) = f(\mathbf{x}_*)$ at P_*. Why? Because f is a constant on all curves lying in the surface. Hence at P_*, on any smooth curve C_* passing through P_*,

$$(df/ds)_* = \nabla f(\mathbf{x}_*)\cdot\mathbf{t}_* = 0, \tag{4.7}$$

where $\mathbf{t}_*$ is the unit tangent to C_* at P_*. Fig. 4.2 illustrates (4.7) for curves $C_*^{(1)}$ and $C_*^{(2)}$ having distinct unit tangents $\mathbf{t}_*^{(1)}$ and $\mathbf{t}_*^{(2)}$.

Problem 4.2.

Find the Cartesian coordinates of that point $P_*(\mathbf{x}_*)$ on the ellipsoid.

$$f \equiv \frac{x^2}{1^2} + \frac{y^2}{2^2} + \frac{z^2}{3^2} = 1$$

where the outward normal vector has Cartesian components (1,1,1).

SOLUTION.

From (4.5)

$$\nabla f \sim (2x,\ y/2,\ 2z/9).$$

As ∇f points to the outside of the ellipsoid (why?), we seek a positive constant k and a vector $\mathbf{x}_* \sim (x_*, y_*, z_*)$ such that $\nabla f(\mathbf{x}_*) \sim (2x_*,\ y_*/2,\ 2z_*/9) = k(1,1,1)$. Thus $x_* = k/2$, $y_* = 2k$, $z_* = 9k/2$. Since $\mathbf{x}_*$ must lie on the ellipsoid, $(k^2/4) + k^2 + (9k^2/4) = 1$. This equation has the positive solution $k = 2/\sqrt{14}$. Hence $\mathbf{x}_* \sim (1,\ 4,\ 9)/\sqrt{14}$.

Linear and Nonlinear Eigenvalue Problems

Linear and nonlinear eigenvalue problems arise in nearly every branch of mechanics and physics. A simple but typical problem—simple to state, that is—is to determine the possible shapes of an idealized, hanging chain when its upper end is spun at a constant rate about a vertical axis. Approximate solutions to such problems may be sought by attacking related finite-dimensional problems, obtained from the original problem by applying some discretation procedure such as the finite element method. Often, the finite-dimensional problems can be phrased as follows: Find those points on the n-dimensional surface $f(\mathbf{x}) = 0$ whose distance from the origin is stationary. If f is smooth, this form of the problem may be recast as an *eigenvalue problem* involving ∇f and given a simple geometric interpretation. The idea may be illustrated in two-dimensions.

A typical smooth curve $f(\mathbf{x}) = 0$ is shown in Fig. 4.3. At any point $P(\mathbf{x})$ where a circle centered at the origin is tangent to $f(\mathbf{x}) = 0$, ∇f must be parallel to $\mathbf{x}$. That is, there is some scalar λ such that

$$\nabla f(\mathbf{x}) = \lambda \mathbf{x}. \tag{4.8}$$

Values of λ satisfying (4.8) are called *eigenvalues* and the associated solutions for $\mathbf{x}$, *eigenvectors*. If f is a *quadratic function* of the Cartesian components of $\mathbf{x}$, then ∇f becomes a linear operator or tensor, and (4.8) reduces to the familiar eigenvalue problem studied in linear algebra.

PROBLEM 4.3.

Find the point(s) on the graph $y = 1/x^4$ closest to the origin.

SOLUTION.

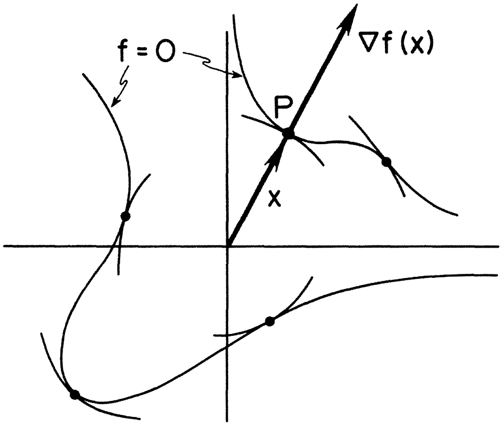


Figure 4.3

Let $f = x^4y - 1$. Then $\nabla f \sim (4x^3y,\ x^4)$ and (4.8) reduces to the two scalar equations

$$4x^3y = \lambda x \text{ and } x^4 = \lambda y.$$

(Note that these equations constitute a *nonlinear* eigenvalue problem.) As $x \neq 0$ (why?), the first equation yields $\lambda = 4x^2y$. Thence, from the second, $x^2 = 4y^2$ or $x = \pm 2y$. The points we seek must lie on $f(\mathbf{x}) = 0$. Hence $16y^5 - 1 = 0$ or $y = 16^{-1/5}$. Thus there is only one eigenvalue, $\lambda = 16^{2/5}$, but two associated eigenvectors, $\mathbf{x} \sim (\pm 16^{1/20}, 16^{-1/5})$, each lying the same distance, $16^{-1/5}\sqrt{5}$, from the origin.

The Del or Gradient Operator

The Leibniz notation *df/ds* has two advantages. It allows us to manipulate the path derivative of *f* as the quotient of two differentials and it allows us to think of the path derivative of *f* as the result of applying the *linear differential operator d/ds* to the function *f*. This second interpretation (but not the first!) carries over to ∇f: application of the *del operator* ∇ to *f* produces ∇f.

The component form of the del operator,

$$\nabla = \mathbf{g}^i \partial/\partial u^i, \tag{4.9}$$

follows immediately from (4.5).

The Divergence, Curl, and Gradient of a Vector Field

The divergence, curl, and gradient of a vector field $\hat{\mathbf{v}}(u^j)$ arise by taking the three possible vector products of ∇ with $\mathbf{v}$. Thus we have

(a). The dot product

$$\nabla\cdot\mathbf{v} = \mathbf{g}^i\cdot\mathbf{v}_{,i}. \tag{4.10}$$

This is a scalar field called the *divergence of* **v** (sometimes denoted by div**v**).

(b). The cross product

$$\nabla \times \mathbf{v} = \mathbf{g}^i \times \mathbf{v}_{,i}. \tag{4.11}$$

This is a vector field called the *curl of* **v** (sometimes denoted by curl**v**).

(c). The direct product

$$\nabla\mathbf{v} = \mathbf{g}^i\mathbf{v}_{,i}. \tag{4.12}$$

This is a 2nd order tensor field called the *gradient of* **v** (sometimes denoted by grad**v**).

By taking ∇ as the second factor in the above vector products, we arrive at the scalar, vector, and tensor operators $\mathbf{v}\cdot\nabla$, $\mathbf{v}\times\nabla$, and $\mathbf{v}\nabla$. All of these are used in continuum mechanics. We shall show presently how $\mathbf{v}\cdot\nabla$ arises.

PROBLEM 4.4.

If $\mathbf{v} = z\mathbf{e}_x + xy\mathbf{e}_y + xyz\mathbf{e}_z$, compute $\nabla\cdot\mathbf{v}$, $\nabla\times\mathbf{v}$, and $\nabla\mathbf{v}$.

SOLUTION.

$\mathbf{v}_{,x} = y\mathbf{e}_y + yz\mathbf{e}_z$, $\mathbf{v}_{,y} = x\mathbf{e}_y + xz\mathbf{e}_z$, $\mathbf{v}_{,z} = \mathbf{e}_x + xy\mathbf{e}_z$, and, in Cartesian coordinates, $\mathbf{g}^1 = \mathbf{e}_x$, $\mathbf{g}^2 = \mathbf{e}_y$, $\mathbf{g}^3 = \mathbf{e}_z$. Hence

$$\begin{aligned}
\nabla\cdot\mathbf{v} &= \mathbf{e}_x\cdot\mathbf{v}_{,x} + \mathbf{e}_y\cdot\mathbf{v}_y + \mathbf{e}_z\cdot\mathbf{v}_{,z}\\
&= 0 + x + xy = x(1+y).\\
\nabla\times\mathbf{v} &= \mathbf{e}_x\times\mathbf{v}_{,x} + \mathbf{e}_y\times\mathbf{v}_{,y} + \mathbf{e}_z\times\mathbf{v}_{,z}\\
&= (y\mathbf{e}_z - yz\mathbf{e}_y) + (xz\mathbf{e}_x) + (\mathbf{e}_y)\\
&= xz\mathbf{e}_x + (1-yz)\mathbf{e}_y + y\mathbf{e}_z.\\
\nabla\mathbf{v} &= \mathbf{e}_x\mathbf{v}_{,x} + \mathbf{e}_y\mathbf{v}_{,y} + \mathbf{e}_z\mathbf{v}_{,z}\\
&= \qquad y\mathbf{e}_x\mathbf{e}_y + yz\mathbf{e}_x\mathbf{e}_z\\
&\qquad + x\mathbf{e}_y\mathbf{e}_y + xz\mathbf{e}_y\mathbf{e}_z\\
&\quad + \mathbf{e}_z\mathbf{e}_x \qquad + xy\mathbf{e}_z\mathbf{e}_z.
\end{aligned}$$

In heat transfer $\nabla\cdot\mathbf{q}$, evaluated at a point P, measures the rate of heat outflow from a neighborhood of P, where $\mathbf{q}$ is the heat flux vector. In fluid dynamics, $\nabla\cdot\rho\mathbf{v}$, evaluated at a point P, measures the rate of decrease of mass in a neighborhood of P, where ρ is the mass density and $\mathbf{v}$ is the fluid velocity. In solid mechanics, $\nabla\cdot\mathbf{u}$, evaluated at a point P, measures the change in volume of the particles in a neighborhood of P when P undergoes a displacement $\mathbf{u}$.

The curl and gradient of a vector field have other important physical interpretations. In fluid dynamics, the vector $\boldsymbol{\omega} \equiv \nabla\times\mathbf{v}$ is called the *vorticity*.

The change from point to point of $\boldsymbol{\omega}$ is a measure of the amount of friction or *viscosity* in the flow, while the symmetric tensor $\mathbf{D} = \frac{1}{2}[\nabla\mathbf{v} + (\nabla\mathbf{v})^T]$, called the *strain-rate,* determines completely the change in shape of an arbitrarily small neighborhood of particles centered, instantaneously, at a point P. In classical fluid dynamics, a knowledge of $\mathbf{D}$ at P allows the stresses in the fluid in a neighorhood of P to be calculated.

The Invariance of $\nabla\cdot\mathbf{v}$, $\nabla \times \mathbf{v}$, and $\nabla\mathbf{v}$

The invariance of $\nabla\cdot\mathbf{v}$, $\nabla \times \mathbf{v}$, and $\nabla\mathbf{v}$ is implied by our coordinate-free notation, but, of course, must be proved. This can be done either by introducing the component form of these expressions and then showing that their values are the same in every coordinate system or else by giving definitions for these expressions that are coordinate-free. For brevity we illustrate the two approaches with $\nabla\cdot\mathbf{v}$.

Let $u^j = \hat{u}^j(\tilde{u}^k)$ define a second set of coordinates. Computing $\nabla\cdot\mathbf{v}$ in the $\tilde{u}^i$-coordinate system via (4.10), we have,

$$\nabla\cdot\mathbf{v} = \tilde{\mathbf{g}}^i\cdot(\partial\mathbf{v}/\partial\tilde{u}^i). \tag{4.13}$$

By the chain rule

$$\begin{aligned}\nabla\cdot\mathbf{v} &= \tilde{\mathbf{g}}^i(\partial\mathbf{v}/\partial u^j)(\partial u^j/\partial\tilde{u}^i)\\ &= \mathbf{g}^j\cdot(\partial\mathbf{v}/\partial u^j), \text{ by (2.34) and (3.87).}\end{aligned} \tag{4.14}$$

This is just (4.10) with j instead of i. Thus, whether (4.10) or (4.14) is used to compute $\nabla\cdot\mathbf{v}$, the values obtained will be the same. That is, $\nabla\cdot\mathbf{v}$ *is an invariant.*

The second way of establishing this invariance is to show that at any point P

$$\nabla\cdot\mathbf{v} = \lim \int_{\partial R} \mathbf{v}\cdot\mathbf{n}dA \text{ as } ||R|| \to 0. \tag{4.15}$$

Here R is a region of diameter $||R||$ enclosing P and having a piecewise smooth boundary ∂R with an outward unit normal $\mathbf{n}$ and differential element of surface area dA. As the integral in (4.15) can be defined without reference to a coordinate system, $\nabla\cdot\mathbf{v}$ must be invariant. Equation (4.15) is a consequence of the divergence theorem that we shall consider later.

The Covariant Derivative

The covariant derivative appears automatically when we express $\mathbf{v}_{,i}$ in terms of the roof components of $\mathbf{v}$. Thus

$$\begin{aligned}\mathbf{v}_{,i} &= (v^j\mathbf{g}_j)_{,i}\\ &= v^j_{,i}\mathbf{g}_j + v^j\mathbf{g}_{j,i}\\ &= v^j_{,i}\mathbf{g}_j + v^j\Gamma^k_{ij}\mathbf{g}_k\\ &= (v^k_{,i} + \Gamma^k_{ij}v^j)\mathbf{g}_k.\end{aligned} \tag{4.16}$$

With the abbreviation

$$\nabla_i v^k \equiv v^k_{,i} + \Gamma^k_{ij}v^j, \tag{4.17}$$

(4.16) reads

$$\mathbf{v}_{,i} = \nabla_i v^k \mathbf{g}_k. \tag{4.18}$$

The symbol $\nabla_i v^k$ is called *the covariant derivative of* v^k (with respect to u^i). It has been defined so that *the component form of* $\mathbf{v}_{,i}$ *in general coordinates looks the same as it does in Cartesian coordinates, save that the partial derivative operator* $\partial/\partial u^i$ *is replaced by the covariant derivative operator* ∇_i.

To compute the covariant derivative of other objects, we insist on the following two properties:

(a) *The covariant derivative of an invariant,* e.g. a scalar, a vector, or a 2nd order tensor, *coincides with its partial derivative.*
(b) *The covariant derivative of a product follows the same rule as does the ordinary derivative.*

Thus, consider the dot product $\phi \equiv \mathbf{u}\cdot\mathbf{v} = u^j v_j$ of two differentiable vectors. It follows, with the aid of (4.17), that

$$\begin{aligned}\phi_{,i} &= (u^j v_j)_{,i}\\ &= u^j_{,i}v_j + u^j v_{j,i}\\ &= (\nabla_i u^j - \Gamma^j_{ik}u^k)v_j + u^j v_{j,i}\\ &= v_j\nabla_i u^j + u^k(v_{k,i} - \Gamma^j_{ik}v_j).\end{aligned} \tag{4.19}$$

But if conditions (a) and (b) are to be fulfilled, we must have, with a renaming of dummy indices, $\phi_{,i} = v_j\nabla_i u^j + u^j\nabla_i v_j$, whence

$$\nabla_i v_j \equiv v_{j,i} - \Gamma^k_{ij}v_k. \tag{4.20}$$

In Exercise 4.5, you are led to the definition of the covariant derivative of the components of a tensor and in Exercise 4.16, you are asked to derive (4.20) in a different way.

Again, by (a), $\mathbf{v}_{,i} = \nabla_i v^j\mathbf{g}_j = \nabla_i v_j\mathbf{g}^j$, which, together with (b), (4.17), and (4.20), implies that

$$\nabla_i\mathbf{g}_j \equiv \mathbf{g}_{j,i} - \Gamma^k_{ij}\mathbf{g}_k = \mathbf{0},\ \nabla_i\mathbf{g}^j \equiv \mathbf{g}^j_{,i} + \Gamma^j_{ik}\mathbf{g}^k = \mathbf{0}, \tag{4.21}$$

i.e., *the base vectors* $\mathbf{g}_j$ *and* $\mathbf{g}^j$ *are covariantly constant.*

From (b) and (4.21),

$$\nabla_i(g_{jk}) = \nabla_i(\mathbf{g}_j\cdot\mathbf{g}_k) = \mathbf{g}_k\cdot\nabla_i\mathbf{g}_j + \mathbf{g}_j\cdot\nabla_i\mathbf{g}_k = 0,\ \text{etc}, \tag{4.22}$$

i.e., the components of the identity tensor $\mathbf{I}$ *are covariantly constant.*

Finally, we note that *the covariant derivative is defined on Riemannian manifolds* (where the g_{ij} determine everything). Why? Because $\nabla_i = \partial/\partial u^i +$ {terms involving the Christoffel symbols} and these latter terms, by (3.70), depend on the g_{ij} only. However, though $\partial^2/\partial u^i \partial u^j = \partial^2/\partial u^j \partial u^i$ (when applied to a sufficiently smooth function), $\nabla_i \nabla_j \neq \nabla_j \nabla_i$, unless the Riemannian manifold is flat. For a proof you should consult a more extended text such as McConnell's *Tensor Analysis*.

The Component Forms of $\nabla \cdot \mathbf{v}$, $\nabla \times \mathbf{v}$, and $\nabla\mathbf{v}$

The component forms of $\nabla \cdot \mathbf{v}$, $\nabla \times \mathbf{v}$, and $\nabla\mathbf{v}$ follow readily with the aid of the above results. In particular, whenever ∇ is applied to an invariant, we may, by (a), take $\nabla = \mathbf{g}^i \nabla_i$. Thus we have for the:

(a) *Dot product*

$$\begin{aligned}\nabla \cdot \mathbf{v} &= \mathbf{g}^i \cdot \nabla_i (v^j \mathbf{g}_j) \\ &= \mathbf{g}^i \cdot \mathbf{g}_j \nabla_i v^j \\ &= \delta^i_j \nabla_i v^j \\ &= \nabla_i v^i = v^i_{,i} + \Gamma^i_{ij} v^j. \end{aligned} \tag{4.23}$$

Now it is a remarkable fact (see Exercise 4.17) that

$$\Gamma^j_{ji} = J^{-1} J_{,i}, \tag{4.24}$$

where J is the Jacobian. We therefore obtain the extremely useful formula

$$\nabla \cdot \mathbf{v} = J^{-1}(\mathrm{J} v^i)_{,i}. \tag{4.25}$$

PROBLEM 4.5.

Compute $\nabla \cdot \mathbf{v}$ in spherical coordinates in terms of the roof and physical components of $\mathbf{v}$.

SOLUTION.

From (3.57), $J = \rho^2 \sin\phi$. Hence, with $\mathbf{v} = v^\rho \mathbf{g}_\rho + v^\phi \mathbf{g}_\phi + v^\theta \mathbf{g}_\theta$, we have, from (4.25),

$$\nabla \cdot \mathbf{v} = \frac{1}{\rho^2 \sin\phi}\left[\frac{\partial}{\partial\rho}(\rho^2 \sin\phi v^\rho) + \frac{\partial}{\partial\phi}(\rho^2 \sin\phi v^\phi) + \frac{\partial}{\partial\theta}(\rho^2 \sin\phi v^\theta)\right].$$

We recall that $v^{(i)} = \mathbf{v} \cdot \bar{\mathbf{g}}^i = v^i/|\mathbf{g}^i| = v^i/\sqrt{g^{ii}}$, where the repeated index is *not* to be summed. Thus, with the aid of the solution to Problem 3.5,

$$v^{(\rho)} = v^\rho,\ v^{(\phi)} = \rho v^\phi,\ v^{(\theta)} = \rho \sin\phi v^\theta,$$

so that, with all products differentiated explicitly,

$$\nabla \cdot \mathbf{v} = \frac{\partial v^{(\rho)}}{\partial \rho} + \frac{2}{\rho} v^{(\rho)} + \frac{1}{\rho}\frac{\partial v^{(\phi)}}{\partial \phi} + \frac{\cot\phi}{\rho} v^{(\phi)} + \frac{1}{\rho \sin\phi}\frac{\partial v^{(\theta)}}{\partial\theta}.$$

(b) *Cross product*

$$\begin{aligned}\nabla \times \mathbf{v} &= \mathbf{g}^i \times \nabla_i(v_j\mathbf{g}^j)\\ &= \mathbf{g}^i \times \mathbf{g}^j\nabla_i v_j\\ &= \epsilon^{ijk}\nabla_i v_j \mathbf{g}_k,\end{aligned} \tag{4.26}$$

i.e., the roof components of the cross product are given by

$$(\nabla \times \mathbf{v})^k = \epsilon^{ijk}\nabla_i v_j. \tag{4.27}$$

(c) *Direct product*

$$\nabla\mathbf{v} = \mathbf{g}^i\nabla_i(v_j\mathbf{g}^j) = \nabla_i v_j \mathbf{g}^i\mathbf{g}^j, \tag{4.28}$$

i.e., *the covariant derivatives $\nabla_i v_j$ are the cellar components of the tensor* $\nabla\mathbf{v}$. Likewise the $\nabla_i v^j$ are one of the two mixed sets of components of $\nabla\mathbf{v}$ because we may also write

$$\nabla\mathbf{v} = \mathbf{g}^i\nabla_i(v^j\mathbf{g}_j) = \nabla_i v^j \mathbf{g}^i\mathbf{g}_j. \tag{4.29}$$

The Kinematics of Continuum Mechanics

The kinematics of continuum mechanics describes the motion of *continuous* distributions of matter, called *bodies*. Each particle X in a body is identified with a point P that moves through E_3. At time t, the set of all such points determines a region $S(t)$ called the *shape* of the body at t. Suppose that S is known at some particular time, say $t = 0$. Let $\mathbf{y}$ be the position of P at time t and $\mathbf{x}$ its position at $t = 0$. Then a *motion of the body* is a transformation of the form

$$\mathbf{y} = \hat{\mathbf{y}}(\mathbf{x},t),\ \mathbf{x} = \hat{\mathbf{y}}(\mathbf{x},0),\ \mathbf{x}\in S(0),\ -\infty < t < \infty. \tag{4.30}$$

If distinct particles are not to coalesce in the course of their motion, then we must demand that the transformation be 1:1. This means that for each $\mathbf{y}$ in the range of the motion, there exists a unique $\mathbf{x}$ given by an inverse transformation of the form

$$\mathbf{x} = \hat{\mathbf{x}}(\mathbf{y},t),\ \mathbf{y}\in S(t),\ -\infty < t < \infty. \tag{4.31}$$

Except at wavefronts (where fields such as temperature, velocity, or stress may jump), we assume that the functions $\hat{\mathbf{y}}$ and $\hat{\mathbf{x}}$ have a sufficient number of derivatives for the various field equations of continuum mechanics to make sense. (Some of these field equations appear in Exercises 4.6, 4.8, and 4.13.)

The *velocity* of a particle with position $\mathbf{x}$ at $t = 0$ is denoted and defined by

$$\mathbf{v} \equiv \dot{\mathbf{y}} \equiv \hat{\mathbf{y}}_{,t}(\mathbf{x},t) \equiv \hat{\mathbf{v}}(\mathbf{x},t) \tag{4.32}$$

and its *acceleration* by

$$\mathbf{a} \equiv \dot{\mathbf{v}} \equiv \hat{\mathbf{v}}_{,t}(\mathbf{x},t) \equiv \hat{\mathbf{a}}(\mathbf{x},t). \tag{4.33}$$

In classical elasticity, the behavior of a body depends on the deviation of its present shape from its initial shape, so all field variables are takes as functions of t and $\mathbf{x}$, which are called *referential or Lagrangian coordinates*. Classical fluid mechanics, in contrast, assumes that the behavior depends only on the rate of change of the present shape. The initial shape being irrelevant, one usually regards all field variables as functions of t and $\mathbf{y}$, which are called *spatial or Eulerian coordinates*. In particular, from (4.31) and (4.32), the velocity field takes the functional form

$$\mathbf{v} \equiv \hat{\mathbf{v}}(\hat{\mathbf{x}}(\mathbf{y},t),t) \equiv \tilde{\mathbf{v}}(\mathbf{y},t). \tag{4.34}$$

As a consequence of this relation, the formula for computing the acceleration must be modified, because $\mathbf{v}$ now depends on t both explicitly through its second argument and implicitly through its first argument $\mathbf{y}$.

As motion occurs in E_3, we may always introduce a Cartesian basis $\{\mathbf{e}_i\}$. Then, with $\mathbf{x} = x^i\mathbf{e}_i$, $\mathbf{y} = y^i\mathbf{e}_i$ and $\mathbf{v} = v^i\mathbf{e}_i$, we have, by the chain rule,

$$\begin{aligned}\mathbf{a} &= \frac{\partial \mathbf{v}}{\partial y^i}\frac{\partial y^i}{\partial t} + \frac{\partial \mathbf{v}}{\partial t} \\ &= v^i\mathbf{v}_{,i} + \mathbf{v}_{,t}\end{aligned} \tag{4.35}$$

Recall the invariant operator $\boldsymbol{\nabla} = \mathbf{g}^i\nabla_i$. As mentioned earlier, we can, given a vector $\mathbf{v}$, form a new invariant operator $\mathbf{v}\cdot\boldsymbol{\nabla} = v^i\nabla_i$. Applied to any field F (scalar, vector, or tensor), this new operator must produce the same result regardless of the coordinate system in which we express it. In the Cartesian coordinate system (y^1,y^2,y^3), $v^i\nabla_i\mathbf{v} = v^i\mathbf{v}_{,i}$. Hence (4.35) can be cast into the invariant form

$$\mathbf{a} = (\mathbf{v}\cdot\boldsymbol{\nabla})\mathbf{v} + \mathbf{v}_{,t}. \tag{4.36}$$

The first term on the right is called the *convective acceleration* and the second term the *local acceleration*. The convective term accounts for the fact that even in steady flow ($\mathbf{v}_{,t} = 0$), there can be acceleration as, for example, when an incompressible fluid flows through a converging nozzle. The differential operator

$$\dot{(\ \ )} \equiv (\mathbf{v}\cdot\boldsymbol{\nabla} + \partial/\partial t) \tag{4.37}$$

that appears in (4.36) is sometimes called the *material derivative*; $\dot{F}$ gives the rate of change of F as we ride through the flow on a particle.

PROBLEM 4.6.

Compute the acceleration of a particle using cylindrical Eulerian coordinates. Express the answer in terms of the physical components of the velocity.

SOLUTION.

With $\mathbf{v} = v^k\mathbf{g}_k$ and $\mathbf{a} = a^k\mathbf{g}_k$, the component form of (4.36) reads

$$a^k = v^i\nabla_i v^k + v^k_{,t},$$

where $\nabla_i v^k$ is given by (4.17). One can verify easily that the only non-zero Christoffel symbols in circular cylindrical coordinates (r, θ, z) are $\Gamma^1_{22} = -r$ and $\Gamma^2_{12} = r^{-1}$. Thus

$$\begin{aligned}
&\nabla_1 v^1 = v^1_{,1} = v^r_{,r},\ \nabla_1 v^2 = v^2_{,1} + \Gamma^2_{12} v^2 = v^\theta_{,r} + r^{-1} v^\theta \\
&\nabla_2 v^1 = v^1_{,2} + \Gamma^1_{22} v^2 = v^r_{,\theta} - r v^\theta,\ \nabla_2 v^2 = v^2_{,2} + \Gamma^2_{21} v^1 = v^\theta_{,\theta} + r^{-1} v^r \\
&\nabla_1 v^3 = v^3_{,1} = v^z_{,r}, \qquad \nabla_2 v^3 = v^3_{,2} = v^z_{,\theta} \\
&\nabla_3 v^1 = v^r_{,z},\ \nabla_3 v^2 = v^\theta_{,z},\ \nabla_3 v^3 = v^z_{,z}.
\end{aligned}$$

In extended form

$$a^1 = v^1 \nabla_1 v^1 + v^2 \nabla_2 v^1 + v^3 \nabla_3 v^1 + v^1_{,t}, \text{ etc.}$$

That is,

$$\begin{aligned}
a^r &= v^r v^r_{,r} + v^\theta (v^r_{,\theta} - r v^\theta) + v^z v^r_{,z} + v^r_{,t} \\
a^\theta &= v^r (v^\theta_{,r} + r^{-1} v^\theta) + v^\theta (v^\theta_{,\theta} + r^{-1} v^r) + v^z v^\theta_{,z} + v^\theta_{,t} \\
a^z &= v^r v^z_{,r} + v^\theta v^z_{,\theta} + v^z v^z_{,z} + v^z_{,t}.
\end{aligned}$$

The relation between the roof and physical components of a vector $\mathbf{u}$ in circular cylindrical coordinates, namely $u^{(r)} = u^r$, $u^{(\theta)} = r u^{(\theta)}$, $u^{(z)} = u^z$, is found in the same way as we found (3.30) in two-dimensional polar coordinates. Hence,

$$a^{(r)} = v^{(r)} v^{(r)}_{,r} + r^{-1} v^{(\theta)} (v^{(r)}_{,\theta} - v^{(\theta)}) + v^{(z)} v^{(z)}_{,z} + v^{(r)}_{,t}, \text{ etc.}$$

The Divergence Theorem

The divergence theorem is the main tool for deriving the local (differential) equations of continuum mechanics from global (integral) statements of the fundamental laws. For example, if ρ and $\mathbf{v}$ denote, respectively, the density and velocity fields of a fluid streaming through a fixed, closed region R of space, then conservation of mass requires that

$$-\left(\int_R \rho dV\right)_{,t} = \int_{\partial R} \rho \mathbf{v} \cdot \mathbf{n} dA. \tag{4.38}$$

Here t is time and ∂R is the boundary of R, assumed to be piecewise smooth with outward unit normal $\mathbf{n}$. Because (4.38) must hold for all regions R, it may be shown that the Divergence Theorem implies

$$\rho_{,t} + \nabla \cdot (\rho \mathbf{v}) = 0 \tag{4.39}$$

in the neighborhood of any point P where the density and velocity fields are sufficiently smooth.

Most large calculus texts establish the Divergence Theorem in Cartesian coordinates.[3] In 2-dimensions it states that:

[3] E.g. Thomas and Finney, *Calculus and Analytic Geometry*, 5th Ed.

$$\int_R (P_{,x} + Q_{,y})dxdy = \int_{\partial R} Pdy - Qdx, \tag{4.40}$$

where P and Q are differentiable functions of x and y and R is some connected region in the xy-plane with a simple, piecewise smooth boundary

$$\partial R : x = \hat{x}(t),\ y = \hat{y}(t),\ \alpha \leq t \leq \beta. \tag{4.41}$$

The right side of (4.40) is short-hand for

$$\int_\alpha^\beta [P(\hat{x}(t),\hat{y}(t))\hat{y}'(t) - Q(\hat{x}(t),\hat{y}(t))\hat{x}'(t)]dt. \tag{4.42}$$

Our aim is to express the Divergence Theorem first in invariant form and then in general coordinate form. To this end, let us assume that ∂R can be parameterized by its arc length s (measured from some point on ∂R in some direction along ∂R). Being piecewise smooth, ∂R has a unit tangent $\mathbf{t}$ everywhere, except at a finite number of points. Furthermore, being simple and closed, ∂R has an inside and an outside. Let s increase in such a way that $\mathbf{n} \equiv \mathbf{t} \times \mathbf{e}_z$ points toward the outside of ∂R. With $\mathbf{v} = P\mathbf{e}_x + Q\mathbf{e}_y$, the integrand in (4.42) reduces to $\mathbf{v}\cdot\mathbf{n}$. We also recognize that $P_{,x} + Q_{,y} = \nabla\cdot\mathbf{v}$. Finally, to disguise the Cartesian coordinates completely, we denote the differential element of area $dxdy$ by dA, thereby reducing (4.40) to the invariant form

$$\int_R \nabla\cdot\mathbf{v}dA = \int_{\partial R} \mathbf{v}\cdot\mathbf{n}ds. \tag{4.43}$$

Our treatment of dA has been, admittedly, cavalier. Whatever its precise definition,[4] dA must be an invariant for the simple reason that everything else in (4.43) is!

To express (4.43) in coordinate form, we set $\mathbf{v} = v^i\mathbf{g}_i$ and $\mathbf{n} = n_i\mathbf{g}^i$. Then $\nabla\cdot\mathbf{v} = \nabla_i v^i$ and $\mathbf{v}\cdot\mathbf{n} = v^i n_i$. What about the elusive dA? I think that the easiest way to derive a coordinate-bound expression for dA is to return to basics.

Recall that the left side of (4.43) is merely a symbol for the limit of a sum of the form

$$S = \sum \phi(\overline{x}_i,\overline{y}_j)\Delta x\Delta y,\ (\overline{x}_i,\overline{y}_j)\epsilon R_{ij} \subset R. \tag{4.44}$$

Here $\phi = \nabla\cdot\mathbf{v}$ and

$$R_{ij} = \{(x,y)|x_i \leq x \leq x_i + \Delta x,\ y_j \leq y \leq y_j + \Delta y\} \tag{4.45}$$

is a Δx by Δy rectangle whose Southwest corner is the point (x_i,y_j), where $x_i = i\Delta x$ and $y_j = j\Delta y$, $i,j = 0,\pm1,\pm2,\ldots$.

In forming the sum S, a rectangular grid is chosen for convenience. Any other will do provided only that the diameter of its largest cell can be made

[4] See Buck, *Advanced Calculus,* 3rd Ed., or Spivak, *Calculus on Manifolds.*

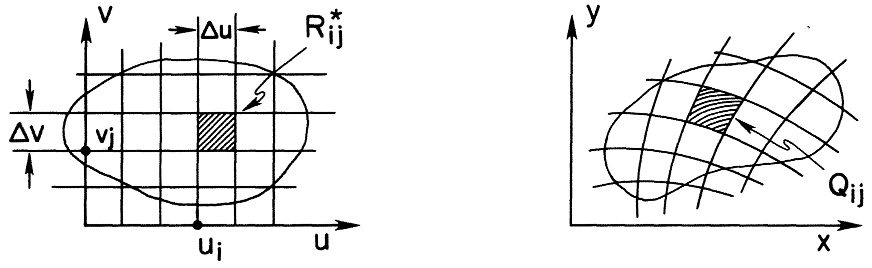


Figure 4.4

arbitrarily small. Under the change of coordinates

$$x = \hat{x}(u,v),\ y = \hat{y}(u,v), \tag{4.46}$$

R becomes the image of some region R^* in the uv-plane, and a rectangular grid in the uv-plane will map into some curvilinear grid in the xy-plane (Fig. 4.4). An element R^*_{ij} of the rectangular grid in the uv-plane is mapped into some curvilinear quadrilateral Q_{ij} in the xy-plane. As in Fig. 4.5, the area of Q_{ij} is approximated by the area $|\mathbf{g}_u \times \mathbf{g}_v|\Delta u \Delta v$ of the parallelogram having two co-terminal edges at the point $(\hat{x}(u_i),\hat{y}(v_j))$. But $|\mathbf{g}_u \times \mathbf{g}_v| = |x,_u y,_v - y,_u x,_v| = |J(u_i,v_j)|$, the absolute value of the Jacobian of the transformation (4.46) at (u_i,v_j). Thus the sum S may be approximated by a new sum of the form

$$S^* = \sum \hat{\phi}(\bar{u}_i,\bar{v}_j)|J(u_i,v_j)|\Delta u \Delta v,\ (\bar{u}_i,\bar{v}_j)\epsilon R^*_{ij} \subset R^*, \tag{4.47}$$

whose limit, as Δu, $\Delta v \to 0$, we denote by $\int_{R^*} \phi |J| du dv$. Hence

$$dA = |J(u,v)| du dv. \tag{4.48}$$

The coordinate form of (4.43) is therefore

$$\int_{R^*} \nabla_i v^i |J| du^1 du^2 = \int_{\partial R} v^i n_i ds, \qquad i = 1,2. \tag{4.49}$$

Analogous considerations in 3-dimensions show that the invariant and coordinate forms of the divergence theorem are, respectively,

$$\int_R \boldsymbol{\nabla}\cdot\mathbf{v} dV = \int_{\partial R} \mathbf{v}\cdot\mathbf{n} dA \tag{4.50}$$

and

$$\int_{R^*} \nabla_i v^i |J| du^1 du^2 du^3 = \int_{\partial R} v^i n_i |K| dv^1 dv^2. \tag{4.51}$$

On the right side of (4.51), it is assumed that the surface ∂R is represented in the parametric form

$$\partial R : \mathbf{x} = \hat{\mathbf{x}}(v^\alpha),\ a^\alpha \le v^\alpha \le b^\alpha, \qquad \alpha = 1,2. \tag{4.52}$$

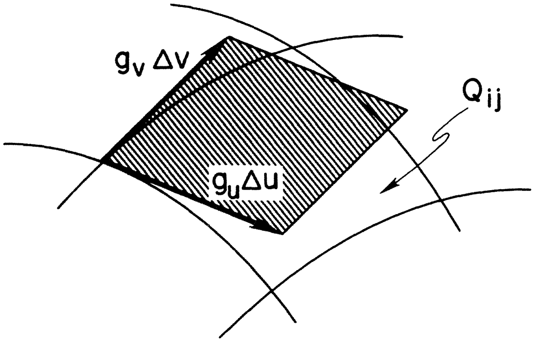


Figure 4.5

with $|K| = |\partial\mathbf{x}/\partial v^1 \times \partial\mathbf{x}/\partial v^2|$.

The arguments that lead from the global form of conservation of mass, (4.38), to its local form, (4.39), may now be spelled out. If $\rho\mathbf{v}$ is differentiable throughout R then, by the divergence theorem, the right side of (4.38) may be replaced by the integral over R of $\nabla\cdot(\rho\mathbf{v})$, giving

$$\int_R [\rho_{,t} + \nabla\cdot(\rho\mathbf{v})]dV = 0. \tag{4.53}$$

Now mass is conserved in all regions R. Hence, if we assume further that the integrand in (4.53) is continuous, it follows, because R is arbitrary, that the integrand must vanish, i.e., (4.39) must obtain. Why? Because if there were a point P_* in the flow where (4.39) failed to hold, then, because $\rho_{,t} + \nabla\cdot(\rho\mathbf{v})$ is assumed to be continuous, it would be of one sign in some small ball centered at P_*. Taking R to coincide with this ball, we would get a contradiction.

Exercises

4.1. If $f = xyz\sqrt{x^2 + y^2}$, compute df/ds at $(1,-2,3)$ in the direction of $\mathbf{v} \sim (2,1,2)$.

4.2. *The mean value theorem*: A region is said to be *convex* if a straight line joining any two points in the region lies wholly within the region. (Compare Figs. 4.6a and 4.6b)

(a). If $P(\mathbf{a})$ and $Q(\mathbf{b})$ are points in a convex region R, show that

$$S:\mathbf{x} = (1-t)\mathbf{a} + t\mathbf{b},\ 0 \le t \le 1$$

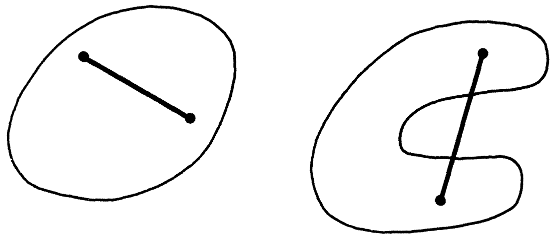
is a parametric equation of the line segment S joining P and Q.

(b). If f is a differentiable function of position in R, show that

$$f(\mathbf{b}) - f(\mathbf{a}) = \nabla f(\mathbf{x}_*)\cdot(\mathbf{b} - \mathbf{a}),$$

where $\mathbf{x}_*$ is the position of some point on S.

Hint: On $S, f(\hat{x}(t)) \equiv \tilde{f}(t)$. Apply the ordinary mean value theorem of calculus to $\tilde{f}$.



(a) convex (b) non-convex

Figure 4.6

4.3. If $f = xy + yz + zx$, compute the cellar components of ∇f in
(a). spherical coordinates (3.42)
(b). the oblique Cartesian coordinates

$$x = u + w, \; y = v - w, \; z = u + v + w.$$

4.4. Just as the action of a dyad **uv** on a vector **w** may be denoted and defined by $\mathbf{uv}\cdot\mathbf{w} \equiv \mathbf{u}(\mathbf{v}\cdot\mathbf{w})$, so the action of a vector **w** on the dyad **uv** may be denoted and defined by $\mathbf{w}\cdot\mathbf{uv} \equiv (\mathbf{w}\cdot\mathbf{u})\mathbf{v}$. By setting $\mathbf{uv} = u^j v^k \mathbf{g}_j\mathbf{g}_k$, establish the identity

$$\nabla\cdot(\mathbf{uv}) = (\boldsymbol{\nabla}\cdot\mathbf{u})\mathbf{v} + (\mathbf{u}\cdot\boldsymbol{\nabla})\mathbf{v}.$$

4.5. We were led to the definition of the covariant derivative of the roof components of a vector field $\hat{\mathbf{v}}(u^j)$ by setting $\mathbf{v},_i = (v^j\mathbf{g}_j),_i \equiv \nabla_i v^j \mathbf{g}_j$ and carrying out the indicated partial differentiation. Likewise, we may define *the covariant derivatives of the roof components of a tensor field* $\hat{\mathbf{T}}(u^j)$ by setting $\mathbf{T},_i = (T^{jk}\mathbf{g}_j\mathbf{g}_k),_i \equiv \nabla_i T^{jk}\mathbf{g}_j\mathbf{g}_k$. Show that

$$\nabla_i T^{jk} = T^{jk}_{,i} + \Gamma^j_{pi}T^{pk} + \Gamma^k_{pi}T^{jp}$$

and hence that the *divergence of a tensor field* takes the component form

$$\boldsymbol{\nabla}\cdot\mathbf{T} = \mathbf{g}^i\cdot\mathbf{T},_i = \nabla_j T^{jk}\mathbf{g}_k.$$

4.6. In continuum mechanics, the linearized equations of motion of a body may be written

$$\boldsymbol{\nabla}\cdot\mathbf{T} + \mathbf{f} = \rho\ddot{\mathbf{u}}, \; \mathbf{T} = \mathbf{T}^T, \tag{4.54}$$

where **T** is the stress tensor (see Exercise 1.20), **f** is the body force vector per unit volume and **u** is the particle displacement.
(a). Write out these equations in component form in Cartesian coordinates, where

$$\mathbf{T} = T_{xx}\mathbf{e}_x\mathbf{e}_x + T_{xy}\mathbf{e}_x\mathbf{e}_y + \ldots, \; \mathbf{f} = f_x\mathbf{e}_x + \ldots, \text{ and } \mathbf{u} = u_x\mathbf{e}_x + \ldots.$$

(b). Do the same in circular cylindrical coordinates, where

$$\mathbf{T} = T^{rr}\mathbf{g}_r\mathbf{g}_r + T^{r\theta}\mathbf{g}_r\mathbf{g}_\theta + \ldots, \; \mathbf{f} = f^r\mathbf{g}_r + \ldots, \text{ and } \mathbf{u} = u^r\mathbf{g}_r + \ldots.$$

(c). Express the roof components of **T** in (b) in terms of its physical components, $T^{(rr)}$, $T^{(r\theta)}$, etc.

4.7. Show that $\nabla\cdot(p\mathbf{I}) = \nabla p$.

4.8. The equations of motion of an ideal gas may be expressed in the coordinate-free form

$$\rho\dot{\mathbf{v}} \equiv \rho\mathbf{v}_{,t} + \rho(\mathbf{v}\cdot\nabla)\mathbf{v} = -\nabla p,$$

where t is time, ρ is the density, $\mathbf{v}$ is the velocity and p is the pressure. With the aid of (4.39) and Exercises 4.4 and 4.7, show that we may cast this equation into the *conservation form*

$$(\rho\mathbf{v})_{,t} + \nabla\cdot(\rho\mathbf{v}\mathbf{v} + p\mathbf{I}) = \mathbf{0}.$$

(This form is useful in numerical work. The term $\rho\mathbf{v}\mathbf{v}$ will be recognized from Exercise 1.19 as the momentum flux tensor.)

4.9. Establish the identities

(a). $$\nabla(\mathbf{u}\cdot\mathbf{v}) \equiv (\nabla\mathbf{u})\cdot\mathbf{v} + (\nabla\mathbf{v})\cdot\mathbf{u}$$

(b). $$\nabla\cdot(f\mathbf{v}) \equiv (\nabla\mathbf{f})\cdot\mathbf{v} + f\nabla\cdot\mathbf{v}.$$

4.10. Use 4.9(b) and (4.37) to show that the equation of conservation of mass can be rewritten in the form $\dot{\rho} + \rho\nabla\cdot\mathbf{v} = 0$. (This equation gives us two equivalent ways to define an *incompressible fluid, viz.* $\dot{\rho} = 0$ or $\nabla\cdot\mathbf{v} = 0$. The second may be preferable in that it is strictly a kinematic condition, i.e. it does not involve a material property (ρ) of the fluid.)

4.11. *The Laplacian,* denoted and defined by $\nabla^2 \equiv \nabla\cdot\nabla$, is one of the most pervasive operators in physics and continuum mechanics.
(a). Show that $\nabla^2 f = g^{ij}\nabla_i f_{,j}$.
(b). Compute $\nabla^2 f$ in Cartesian, oblique Cartesian, circular cylindrical, and spherical coordinates.

4.12. Establish the identities

(a). $\nabla\cdot(\nabla\mathbf{u}) = \nabla^2\mathbf{u}$, (b). $\nabla\cdot(\nabla\mathbf{u})^T = \nabla(\nabla\cdot\mathbf{u})$, (c). $\mathrm{tr}\,(\nabla\mathbf{u}) = \nabla\cdot\mathbf{u}$.

4.13. *The linear field equations for elastically homogeneous and isotropic bodies* consist of the equations of motion (4.54), the *stress-strain relations*

$$\mathbf{T} = \lambda\,\mathrm{tr}\,(\mathbf{E})\mathbf{I} + 2\mu\mathbf{E}, \tag{4.55}$$

and the *strain-displacement equations*

$$\mathbf{E} = \tfrac{1}{2}[\nabla\mathbf{u} + (\nabla\mathbf{u})^T]. \tag{4.56}$$

In (4.55), λ and μ are constants called the Lamé coefficients. By substituting (4.56) into (4.55) and the resulting expression into (4.54), and noting Exercises 4.4 and 4.9, show that

$$(\lambda + \mu)\nabla(\nabla\cdot\mathbf{u}) + \mu\nabla^2\mathbf{u} + \mathbf{f} = \rho\ddot{\mathbf{u}}.$$

These are called the *Navier equations*.

4.14. Assuming that ϕ and $\mathbf{v}$ are differentiable, show that

(a). $$\nabla\cdot(\phi\mathbf{v}) \equiv \nabla\phi\cdot\mathbf{v} + \phi\nabla\cdot\mathbf{v}.$$

(b). $$\nabla \times (\phi \mathbf{v}) \equiv \nabla\phi \times \mathbf{v} + \phi\nabla \times \mathbf{v}.$$

(c). $$\nabla(\phi\mathbf{v}) \equiv \nabla\phi\mathbf{v} + \phi\nabla\mathbf{v}.$$

4.15. (See Exercise 1.23). The *trace* of a 2nd order tensor $\mathbf{T}$ is denoted and defined by

$$\operatorname{tr} \mathbf{T} = T_i^{\cdot j}\mathbf{g}^i\cdot\mathbf{g}_j = T_i^{\cdot i}.$$

(a). Show that

$$\operatorname{tr} \mathbf{T} = T^i_{\cdot j} = g^{ij}T_{ij} = g_{ij}T^{ij}.$$

(b). Show that tr $\mathbf{T}$ is an *invariant*, i.e., that $T_i^{\cdot i} = \tilde{T}_i^{\cdot i}$ under the change of coordinates $u^j = \hat{u}^j(\tilde{u}^k)$.
(Equating a roof and cellar index of the mixed components of a tensor produces the components of a new object called a *contraction* of the tensor. Thus, in (b), the scalar tr $\mathbf{T}$ is the contraction of $\mathbf{T}$. If $\mathbf{U} = U^{ij}_{\cdot\cdot k}\mathbf{g}_i\mathbf{g}_j\mathbf{g}^k$, then the vectors $U^{ij}_{\cdot\cdot i}\mathbf{g}_j$ and $U^{ij}_{\cdot\cdot j}\mathbf{g}_i$ are both contractions of $\mathbf{U}$.)

4.16. Derive (4.20) by setting $\mathbf{v}_{,i} = (v_j\mathbf{g}^j)_{,i}$ and using (3.104).

4.17. Use (3.71), Exercise 2.10, and the following observations to establish (4.24):

(a). If det $[g_{ij}]$ is expanded about the row (or column) containing the element g_{ij}, the coefficient of g_{ij} is J^2g^{ij}. This follows by applying (3.70) to the matrix G^2.

(b). If a determinant is regarded as a function of its elements, then, by the chain rule,

$$\frac{\partial \det [g_{ij}]}{\partial u^k} = J^2 g^{ij}\frac{\partial g_{ij}}{\partial u^k}.$$

4.18. Compute $\nabla^2\psi$ in parabolic-, elliptic-, and bipolar cylindrical coordinates, as defined in parts *(f)*, *(g)*, *(h)* of Exercise 3.17.

4.19. *A Rigid Body Motion* $\mathbf{y} = \mathbf{r}(\mathbf{x},t)$ preserves distance between particles, i.e.

$$|\mathbf{r}(\mathbf{x},t) - \mathbf{r}(\boldsymbol{\xi},t)| = |\mathbf{x} - \boldsymbol{\xi}|, \; -\infty < t < \infty, \; \mathbf{x},\boldsymbol{\xi}\epsilon S(0).$$

Show that any such motion must be of the form

$$\mathbf{r}(\mathbf{x},t) = \mathbf{c}(t) + \mathbf{Q}(t)\mathbf{x}, \; \mathbf{c}(0) = \mathbf{0}, \; \mathbf{Q}(0) = \mathbf{I}, \tag{4.57}$$

where $\mathbf{c}$ is an arbitrary time-dependent vector and $\mathbf{Q}$ is an arbitrary time-dependent rotator (see Exercise 2.16).
Hints:

(a). Set $\mathbf{r}(\mathbf{x},t) = \mathbf{r}(\mathbf{0},t) + \mathbf{g}(\mathbf{x},t)$ and note that $\mathbf{r}(\mathbf{0},0) = \mathbf{g}(\mathbf{0},t) = \mathbf{0}$ and $\mathbf{g}(\mathbf{x},0) = \mathbf{x}$.

(b). Show that $|\mathbf{g}(\alpha\mathbf{x},t)| = |\alpha||\mathbf{x}|$.

(c). Use (1.11) to show that $\mathbf{g}$ preserves dot products, i.e.,

$$\mathbf{g}(\mathbf{x},t)\cdot\mathbf{g}(\boldsymbol{\xi},t) = \mathbf{x}\cdot\boldsymbol{\xi}, \; \mathbf{x},\boldsymbol{\xi}\epsilon S(0).$$

(d). Show that $\mathbf{g}$ is linear by showing that $|\mathbf{g}(\mathbf{x} + \boldsymbol{\xi},t) - \mathbf{g}(\mathbf{x},t) - \mathbf{g}(\boldsymbol{\xi},t)|^2 = 0$.

(e). Rename $\mathbf{r}(\mathbf{0},t)$ and $\mathbf{g}(\mathbf{x},t)$ to arrive at (4.57).

4.20. *The velocity in rigid body motion*: Starting from (4.57), show that

$$\dot{\mathbf{y}} = \dot{\mathbf{c}}(t) + \boldsymbol{\omega}(t) \times (\mathbf{y} - \mathbf{c}(t)). \tag{4.58}$$

Hint: Differentiate (4.57) with respect to time, express $\mathbf{x}$ in terms of $\mathbf{y}$ and $\mathbf{c}$, and differentiate both sides of $\mathbf{Q}\cdot\mathbf{Q}^T = \mathbf{I}$ to show that $\dot{\mathbf{Q}}\cdot\mathbf{Q}^T$ is skew and that $\boldsymbol{\omega}$ is its axis. (See Exercise 1.18a.)

4.21. *The gross equations of motion of a body,* mentioned in the introduction to Chapter III, are obtained as follows.

Let $\rho(\mathbf{x})$ denote the mass density of a body B in its initial shape $S(0) \equiv S_O$. By definition, the mass of B is

$$M = \int_{S_O} \rho dV. \tag{4.59}$$

Let (4.30) describe the motion of B in an inertial frame with origin O. The *linear momentum of B* is then defined by

$$\mathbf{L} = \int_{S_O} \rho\dot{\mathbf{y}}dV = \left(\int_{S_O} \rho\mathbf{y}dV\right)^{\cdot}, \tag{4.60}$$

and its *rotational momentum with respect to O* by

$$\mathbf{R}_O = \int_{S_O} \rho\mathbf{y} \times \dot{\mathbf{y}}dV. \tag{4.61}$$

The *center of mass* of B is the vector

$$\mathbf{X} = M^{-1} \int_{S_O} \rho\mathbf{y}dV. \tag{4.62}$$

(a). Show that if we set

$$\mathbf{y} = \mathbf{X} + \mathbf{z}, \tag{4.63}$$

then

$$\int_{S_O} \rho\mathbf{z}dV = \mathbf{0} \tag{4.64}$$

and

$$\mathbf{R}_O = M\mathbf{X} \times \dot{\mathbf{X}} + \int_{S_O} \rho\mathbf{z} \times \dot{\mathbf{z}}dV. \tag{4.65}$$

(b). Show that (3.4) reduces to

$$\mathbf{T} = \left(\int_{S_O} \rho\mathbf{z} \times \dot{\mathbf{z}}dV\right)^{\cdot} \equiv \dot{\mathbf{R}}_C. \tag{4.66}$$

$\mathbf{R}_C$ is called *the rotational momentum about the center of mass of B.*

4.22. *The moment of inertia tensor* appears when we compute the rotational momentum of a *rigid body,* i.e., a body capable of undergoing rigid motions only.

(a). Let the origin of the inertial frame coincide with the center of mass of the body at $t = 0$. With reference to Exercise 4.19 and (4.62), show that

$$\mathbf{X} = \mathbf{c},\ \mathbf{z} = \mathbf{Q}\mathbf{x},\ \dot{\mathbf{z}} = \boldsymbol{\omega} \times \mathbf{z}.$$

(b). Show that $\mathbf{z} \times \dot{\mathbf{z}} = \mathbf{Q}\cdot[(\mathbf{x}\cdot\mathbf{x})\mathbf{I} - \mathbf{x}\mathbf{x}]\cdot\mathbf{Q}^T\boldsymbol{\omega}$ and hence, since $\mathbf{Q}$ is a function of t only, that

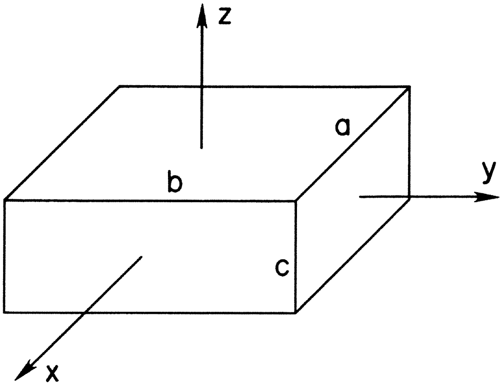


Figure 4.7

$$\mathbf{R}_C = \mathbf{Q}\cdot\mathbf{I}\cdot\mathbf{Q}^T\boldsymbol{\omega}, \qquad (4.67)$$

where

$$\mathbf{I} = \int_{S_O} \rho(\mathbf{x})[(\mathbf{x}\cdot\mathbf{x})\mathbf{I} - \mathbf{x}\mathbf{x}]dV \qquad (4.68)$$

is the *moment of inertia tensor*.

If ρ is a constant, compute the Cartesian components of $\mathbf{I}$ for

(c). A sphere of radius α.

(d). A rectangle parallipiped of dimension $a \times b \times c$ with the alignment of axes shown in Fig. 4.7.

4.23. Recalling that in the preceding exercise

$$\dot{\mathbf{Q}}\cdot\mathbf{Q}^T = \boldsymbol{\omega} \times , \qquad (4.69)$$

show that

$$\mathbf{Q}^T\dot{\mathbf{Q}} = (\mathbf{Q}^T\boldsymbol{\omega}) \times . \qquad (4.70)$$

Hint: In order, show that $\mathbf{Q}^T\boldsymbol{\omega}$ is an eigenvector of $\mathbf{Q}^T\dot{\mathbf{Q}}$; that the representation (4.70) is unique, and that, from Exercise 1.26,

$$|\mathbf{Q}^T\boldsymbol{\omega}|^2 = |\boldsymbol{\omega}|^2 = \text{tr } \dot{\mathbf{Q}}^T\dot{\mathbf{Q}} = \text{tr } \dot{\mathbf{Q}}\cdot\dot{\mathbf{Q}}^T.$$

4.24. *Euler's equations of motion for a rigid body* are

$$\mathbf{T}^* = \boldsymbol{\omega}^* \times \mathbf{I}\cdot\boldsymbol{\omega}^* + \mathbf{I}\cdot\dot{\boldsymbol{\omega}}^*. \qquad (4.71)$$

Show that (4.71) follows from (4.66), (4.67) and (4.70) upon setting

$$\boldsymbol{\omega} = \mathbf{Q}\boldsymbol{\omega}^*, \ \mathbf{R}_C = \mathbf{Q}\cdot\mathbf{R}_C^*, \ \mathbf{T} = \mathbf{Q}\cdot\mathbf{T}^*. \qquad (4.72)$$

(Physically, these equations may be regarded as a change from an inertial frame to a non-inertial, body-fixed frame. See the comments at the beginning of Chapter I and Goldstein, *op. cit.*)

Index

Undergraduate Texts in Mathematics

Apostol: Introduction to Analytic Number Theory.
1976. xii, 338 pages. 24 illus.

Bak/Newman: Complex Analysis.
1982. x, 244 pages. 69 illus.

Childs: A Concrete Introduction to Higher Algebra.
1979. xiv, 338 pages. 8 illus.

Chung: Elementary Probability Theory with Stochastic Processes.
1975. xvi, 325 pages. 36 illus.

Croom: Basic Concepts of Algebraic Topology.
1978. x, 177 pages. 46 illus.

Fleming: Functions of Several Variables. Second edition.
1977. xi, 411 pages. 96 illus.

Foulds: Optimization Techniques: An Introduction.
1981. xii, 502 pages. 72 illus.

Franklin: Methods of Mathematical Economics. Linear and Nonlinear Programming, Fixed-Point Theorems.
1980. x, 297 pages. 38 illus.

Halmos: Finite-Dimensional Vector Spaces. Second edition.
1974. viii, 200 pages.

Halmos: Naive Set Theory.
1974. vii, 104 pages.

Iooss/Joseph: Elementary Stability and Bifurcation Theory.
1980. xv, 286 pages. 47 illus.

Kemeny/Snell: Finite Markov Chains.
1976. ix, 224 pages. 11 illus.

Lax/Burstein/Lax: Calculus with Applications and Computing, Volume 1.
1976. xi, 513 pages. 170 illus.

LeCuyer: College Mathematics with A Programming Language.
1978. xii, 420 pages. 144 illus.

Macki/Strauss: Introduction to Optimal Control Theory.
1981. xiii, 168 pages. 68 illus.

Malitz: Introduction to Mathematical Logic: Set Theory - Computable Functions - Model Theory.
1979. xii, 198 pages. 2 illus.

Martin: Transformation Geometry: An Introduction to Symmetry.
1982. xii, 237 pages. 209 illus.

Millman/Parker: Geometry: A Metric Approach with Models.
1981. viii, 355 pages. 259 illus.

Prenowitz/Jantosciak: Join Geometrics: A Theory of Convex Set and Linear Geometry.
1979. xxii, 534 pages. 404 illus.

Priestley: Calculus: An Historical Approach.
1979. xvii, 448 pages. 335 illus.

Protter/Morrey: A First Course in Real Analysis.
1977. xii, 507 pages. 135 illus.

Ross: Elementary Analysis: The Theory of Calculus.
1980. viii, 264 pages. 34 illus.

Sigler: Algebra.
1976. xii, 419 pages. 27 illus.

Simmonds: A Brief on Tensor Analysis.
1982. xi, 92 pages. 28 illus.

Singer/Thorpe: Lecture Notes on Elementary Topology and Geometry.
1976. viii, 232 pages. 109 illus.

Smith: Linear Algebra.
1978. vii, 280 pages. 21 illus.

Thorpe: Elementary Topics in Differential Geometry.
1979. xvii. 253 pages. 126 illus.

Whyburn/Duda: Dynamic Topology.
1979. xiv, 338 pages. 20 illus.

Wilson: Much Ado About Calculus: A Modern Treatment with Applications Prepared for Use with the Computer.
1979. xvii, 788 pages. 145 illus.