CONTENTS

T2 Track Two; see page xvii

N Nonrelativistic (Newtonian) kinetic theory; see page 96

R Relativistic theory; see page 96

CHAPTER ONE

1

Newtonian Physics: Geometric Viewpoint

Geometry postulates the solution of these problems from mechanics and teaches the use of the problems thus solved. And geometry can boast that with so few principles obtained from other fields, it can do so much.

ISAAC NEWTON, 1687

1.1 Introduction 1.1

1.1.1 The Geometric Viewpoint on the Laws of Physics 1.1.1

In this book, we adopt a different viewpoint on the laws of physics than that in many elementary and intermediate texts. In most textbooks, physical laws are expressed in terms of quantities (locations in space, momenta of particles, etc.) that are measured in some coordinate system. For example, Newtonian vectorial quantities are expressed as triplets of numbers [e.g., $\mathbf{p} = (p_x, p_y, p_z) = (1, 9, -4)$], representing the components of a particle's momentum on the axes of a Cartesian coordinate system; and tensors are expressed as arrays of numbers (e.g.,

$$\mathbf{I} = \begin{bmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{bmatrix} \tag{1.1}$$

for the moment of inertia tensor).

By contrast, in this book we express all physical quantities and laws in *geometric forms,* i.e., in forms that are *independent of any coordinate system or basis vectors.* For example, a particle's velocity $\mathbf{v}$ and the electric and magnetic fields $\mathbf{E}$ and $\mathbf{B}$ that it encounters will be vectors described as arrows that live in the 3-dimensional, flat Euclidean space of everyday experience.[1] They require no coordinate system or basis vectors for their existence or description—though often coordinates will be useful. In other words, $\mathbf{v}$ represents the vector itself and is not just shorthand for an ordered list of numbers.

1. This interpretation of a vector is close to the ideas of Newton and Faraday. Lagrange, Hamilton, Maxwell, and many others saw vectors in terms of Cartesian components. The vector notation was streamlined by Gibbs, Heaviside, and others, but the underlying coordinate system was still implicit, and $\mathbf{v}$ was usually regarded as shorthand for (v_x, v_y, v_z).

BOX 1.1. READERS' GUIDE

- This chapter is a foundation for almost all of this book.
- Many readers already know the material in this chapter, but from a viewpoint different from our *geometric* one. Such readers will be able to understand almost all of Parts II–VI of this book without learning our viewpoint. Nevertheless, that geometric viewpoint has such power that we encourage them to learn it by browsing this chapter and focusing especially on Secs. 1.1.1, 1.2, 1.3, 1.5, 1.7, and 1.8.
- The stress tensor, introduced and discussed in Sec. 1.9, plays an important role in kinetic theory (Chap. 3) and a crucial role in elasticity (Part IV), fluid dynamics (Part V), and plasma physics (Part VI).
- The integral and differential conservation laws derived and discussed in Secs. 1.8 and 1.9 play major roles throughout this book.
- The Box labeled T2 is advanced material (Track Two) that can be skipped in a time-limited course or on a first reading of this book.

We insist that the Newtonian laws of physics all obey a *Geometric Principle:* they are all geometric relationships among geometric objects (primarily scalars, vectors, and tensors), expressible without the aid of any coordinates or bases. An example is the Lorentz force law $m d\mathbf{v}/dt = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$—a (coordinate-free) relationship between the geometric (coordinate-independent) vectors $\mathbf{v}$, $\mathbf{E}$, and $\mathbf{B}$ and the particle's scalar mass m and charge q. As another example, a body's moment of inertia tensor $\mathbf{I}$ can be viewed as a vector-valued linear function of vectors (a coordinate-independent, basis-independent geometric object). Insert into the tensor $\mathbf{I}$ the body's angular velocity vector $\mathbf{\Omega}$, and you get out the body's angular momentum vector: $\mathbf{J} = \mathbf{I}(\mathbf{\Omega})$. No coordinates or basis vectors are needed for this law of physics, nor is any description of $\mathbf{I}$ as a matrix-like entity with components I_{ij} required. Components are secondary; they only exist after one has chosen a set of basis vectors. Components (we claim) are an impediment to a clear and deep understanding of the laws of classical physics. The coordinate-free, component-free description is deeper, and—once one becomes accustomed to it—much more clear and understandable.[2]

2. This philosophy is also appropriate for quantum mechanics (see Box 1.2) and, especially, quantum field theory, where it is the invariance of the description under gauge and other symmetry operations that is the powerful principle. However, its implementation there is less direct, simply because the spaces in which these symmetries lie are more abstract and harder to conceptualize.

By adopting this geometric viewpoint, we gain great conceptual power and often also computational power. For example, when we ignore experiment and simply ask what forms the laws of physics can possibly take (what forms are allowed by the requirement that the laws be geometric), we shall find that there is remarkably little freedom. Coordinate independence and basis independence strongly constrain the laws of physics.[3]

This power, together with the elegance of the geometric formulation, suggests that in some deep sense, Nature's physical laws are geometric and have nothing whatsoever to do with coordinates or components or vector bases.

1.1.2 Purposes of This Chapter 1.1.2

The principal purpose of this foundational chapter is to teach the reader this geometric viewpoint.

The mathematical foundation for our geometric viewpoint is *differential geometry* (also called "tensor analysis" by physicists). Differential geometry can be thought of as an extension of the vector analysis with which all readers should be familiar. *A second purpose of this chapter is to develop key parts of differential geometry in a simple form well adapted to Newtonian physics.*

1.1.3 Overview of This Chapter 1.1.3

In this chapter, we lay the geometric foundations for the Newtonian laws of physics in flat Euclidean space. We begin in Sec. 1.2 by introducing some foundational geometric concepts: points, scalars, vectors, inner products of vectors, and the distance between points. Then in Sec. 1.3, we introduce the concept of a tensor as a linear function of vectors, and we develop a number of geometric tools: the tools of coordinate-free tensor algebra. In Sec. 1.4, we illustrate our tensor-algebra tools by using them to describe—without any coordinate system—the kinematics of a charged point particle that moves through Euclidean space, driven by electric and magnetic forces.

In Sec. 1.5, we introduce, for the first time, Cartesian coordinate systems and their basis vectors, and also the components of vectors and tensors on those basis vectors; and we explore how to express geometric relationships in the language of components. In Sec. 1.6, we deduce how the components of vectors and tensors transform when one rotates the chosen Cartesian coordinate axes. (These are the transformation laws that most physics textbooks use to define vectors and tensors.)

In Sec. 1.7, we introduce directional derivatives and gradients of vectors and tensors, thereby moving from tensor algebra to true differential geometry (in Euclidean space). We also introduce the Levi-Civita tensor and use it to define curls and cross

3. Examples are the equation of elastodynamics (12.4b) and the Navier-Stokes equation of fluid mechanics (13.69), which are both dictated by momentum conservation plus the form of the stress tensor [Eqs. (11.18), (13.43), and (13.68)]—forms that are dictated by the irreducible tensorial parts (Box 11.2) of the strain and rate of strain.

products, and we learn how to use *index gymnastics* to derive, quickly, formulas for multiple cross products. In Sec. 1.8, we use the Levi-Civita tensor to define vectorial areas, scalar volumes, and integration over surfaces. These concepts then enable us to formulate, in geometric, coordinate-free ways, integral and differential conservation laws. In Sec. 1.9, we discuss, in particular, the law of momentum conservation, formulating it in a geometric way with the aid of a geometric object called the *stress tensor.* As important examples, we use this geometric conservation law to derive and discuss the equations of Newtonian fluid dynamics, and the interaction between a charged medium and an electromagnetic field. We conclude in Sec. 1.10 with some concepts from special relativity that we shall need in our discussions of Newtonian physics.

1.2 Foundational Concepts

In this section, we sketch the foundational concepts of Newtonian physics without using any coordinate system or basis vectors. This is the geometric viewpoint that we advocate.

space and time

The arena for the Newtonian laws of physics is a spacetime composed of the familiar 3-dimensional Euclidean space of everyday experience (which we call *3-space*) and a universal time t. We denote points (locations) in 3-space by capital script letters, such as $\mathcal{P}$ and $\mathcal{Q}$. These points and the 3-space in which they live require no coordinates for their definition.

scalar

A *scalar* is a single number. We are most interested in scalars that directly represent physical quantities (e.g., temperature T). As such, they are real numbers, and when they are functions of location $\mathcal{P}$ in space [e.g., $T(\mathcal{P})$], we call them *scalar fields.* However, sometimes we will work with complex numbers—most importantly in quantum mechanics, but also in various Fourier representations of classical physics.

vector

A *vector* in Euclidean 3-space can be thought of as a straight arrow (or more formally a directed line segment) that reaches from one point, $\mathcal{P}$, to another, $\mathcal{Q}$ (e.g., the arrow $\Delta\mathbf{x}$ in Fig. 1.1a). Equivalently, $\Delta\mathbf{x}$ can be thought of as a direction at $\mathcal{P}$ and a number, the vector's length. Sometimes we shall select one point $\mathcal{O}$ in 3-space as an "origin" and identify all other points, say, $\mathcal{Q}$ and $\mathcal{P}$, by their vectorial separations $\mathbf{x}_\mathcal{Q}$ and $\mathbf{x}_\mathcal{P}$ from that origin.

distance and length

The Euclidean distance $\Delta\sigma$ between two points $\mathcal{P}$ and $\mathcal{Q}$ in 3-space can be measured with a ruler and so, of course, requires no coordinate system for its definition. (If one does have a Cartesian coordinate system, then $\Delta\sigma$ can be computed by the Pythagorean formula, a precursor to the invariant interval of flat spacetime; Sec. 2.2.3.) This distance $\Delta\sigma$ is also the *length* $|\Delta\mathbf{x}|$ of the vector $\Delta\mathbf{x}$ that reaches from $\mathcal{P}$ to $\mathcal{Q}$, and the square of that length is denoted

$$|\Delta\mathbf{x}|^2 \equiv (\Delta\mathbf{x})^2 \equiv (\Delta\sigma)^2. \tag{1.2}$$

Of particular importance is the case when $\mathcal{P}$ and $\mathcal{Q}$ are neighboring points and $\Delta\mathbf{x}$ is a differential (infinitesimal) quantity $d\mathbf{x}$. This *infinitesimal displacement* is a more fundamental physical quantity than the finite $\Delta\mathbf{x}$. To create a finite vector out

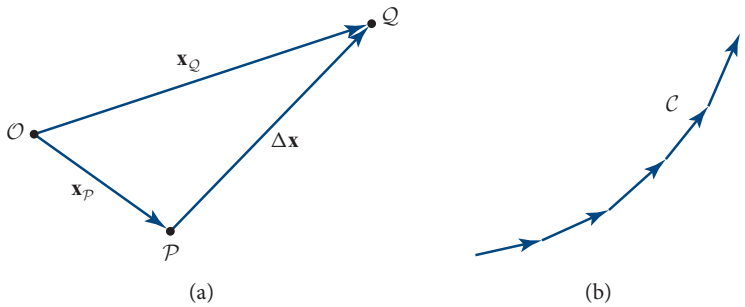

FIGURE 1.1 (a) A Euclidean 3-space diagram depicting two points $\mathcal{P}$ and $\mathcal{Q}$, their respective vectorial separations $\mathbf{x}_\mathcal{P}$ and $\mathbf{x}_\mathcal{Q}$ from the (arbitrarily chosen) origin $\mathcal{O}$, and the vector $\Delta\mathbf{x} = \mathbf{x}_\mathcal{Q} - \mathbf{x}_\mathcal{P}$ connecting them. (b) A curve $\mathcal{P}(\lambda)$ generated by laying out a sequence of infinitesimal vectors, tail-to-tip.

of infinitesimal vectors, one has to add several infinitesimal vectors head to tail, head to tail, and so on, and then take a limit. This involves *translating* a vector from one point to the next. There is no ambiguity about doing this in flat Euclidean space using the geometric notion of parallelism.[4] This simple property of Euclidean space enables us to add (and subtract) vectors at a point. We attach the tail of a second vector to the head of the first vector and then construct the sum as the vector from the tail of the first to the head of the second, or vice versa, as should be quite familiar. The point is that we do not need to add the Cartesian components to sum vectors.

We can also rotate vectors about their tails by pointing them along a different direction in space. Such a rotation can be specified by two angles. The space that is defined by all possible changes of length and direction at a point is called that point's *tangent space*. Again, we generally view the rotation as being that of a physical vector in space, and not, as it is often useful to imagine, the rotation of some coordinate system's basis vectors, with the chosen vector itself kept fixed.

tangent space

We can also construct a path through space by laying down a sequence of infinitesimal $d\mathbf{x}$s, tail to head, one after another. The resulting path is a *curve* to which these $d\mathbf{x}$s are tangent (Fig. 1.1b). The curve can be denoted $\mathcal{P}(\lambda)$, with λ a parameter along the curve and $\mathcal{P}(\lambda)$ the point on the curve whose parameter value is λ, or $\mathbf{x}(\lambda)$ where $\mathbf{x}$ is the vector separation of $\mathcal{P}$ from the arbitrary origin $\mathcal{O}$. The infinitesimal vectors that map the curve out are $d\mathbf{x} = (d\mathcal{P}/d\lambda)\, d\lambda = (d\mathbf{x}/d\lambda)\, d\lambda$, and $d\mathcal{P}/d\lambda = d\mathbf{x}/d\lambda$ is the tangent vector to the curve.

curve

tangent vector

If the curve followed is that of a particle, and the parameter λ is time t, then we have defined the *velocity* $\mathbf{v} \equiv d\mathbf{x}/dt$. In effect we are multiplying the vector $d\mathbf{x}$ by the scalar $1/dt$ and taking the limit. Performing this operation at every point $\mathcal{P}$ in the space occupied by a fluid defines the fluid's *velocity field* $\mathbf{v}(\mathbf{x})$. Multiplying a particle's velocity $\mathbf{v}$ by its scalar mass gives its *momentum* $\mathbf{p} = m\mathbf{v}$. Similarly, the difference $d\mathbf{v}$

4. The statement that there is just one choice of line parallel to a given line, through a point not lying on the line, is the famous fifth axiom of Euclid.

of two velocity measurements during a time interval dt, multiplied by $1/dt$, generates the particle's *acceleration* $\mathbf{a} = d\mathbf{v}/dt$. Multiplying by the particle's mass gives the force $\mathbf{F} = m\mathbf{a}$ that produced the acceleration; dividing an electrically produced force by the particle's charge q gives the electric field $\mathbf{E} = \mathbf{F}/q$. And so on.

We can define inner products [see Eq. (1.4a) below] and cross products [Eq. (1.22a)] of pairs of vectors at the same point geometrically; then using those vectors we can define, for example, the rate that work is done by a force and a particle's angular momentum about a point.

These two products can be expressed geometrically as follows. If we allow the two vectors to define a parallelogram, then their cross product is the vector orthogonal to the parallelogram with length equal to the parallelogram's area. If we first rotate one vector through a right angle in a plane containing the other, and then define the parallelogram, its area is the vectors' inner product.

derivatives of scalars and vectors

We can also define spatial derivatives. We associate the difference of a scalar between two points separated by $d\mathbf{x}$ at the same time with a *gradient* and, likewise, go on to define the scalar *divergence* and the vector *curl*. The freedom to translate vectors from one point to the next also underlies the association of a single vector (e.g., momentum) with a group of particles or an extended body. One simply adds all the individual momenta, taking a limit when necessary.

In this fashion (which should be familiar to the reader and will be elucidated, formalized, and generalized below), we can construct all the standard scalars and vectors of Newtonian physics. What is important is that *these physical quantities require no coordinate system for their definition.* They are geometric (coordinate-independent) objects residing in Euclidean 3-space at a particular time.

Geometric Principle

It is a fundamental (though often ignored) principle of physics that *the Newtonian physical laws are all expressible as geometric relationships among these types of geometric objects, and these relationships do not depend on any coordinate system or orientation of axes, nor on any reference frame* (i.e., on any purported velocity of the Euclidean space in which the measurements are made).[5] We call this the *Geometric Principle* for the laws of physics, and we use it throughout this book. It is the Newtonian analog of Einstein's Principle of Relativity (Sec. 2.2.2).

1.3 Tensor Algebra without a Coordinate System

In preparation for developing our geometric view of physical laws, we now introduce, in a coordinate-free way, some fundamental concepts of differential geometry: tensors, the inner product, the metric tensor, the tensor product, and contraction of tensors.

We have already defined a vector $\mathbf{A}$ as a straight arrow from one point, say $\mathcal{P}$, in our space to another, say $\mathcal{Q}$. Because our space is flat, there is a unique and obvious way to

5. By changing the velocity of Euclidean space, one adds a constant velocity to all particles, but this leaves the laws (e.g., Newton's $\mathbf{F} = m\mathbf{a}$) unchanged.

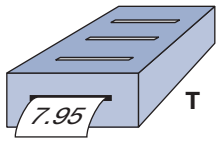

FIGURE 1.2 A rank-3 tensor **T**.

transport such an arrow from one location to another, keeping its length and direction unchanged.[6] Accordingly, we shall regard vectors as unchanged by such transport. This enables us to ignore the issue of where in space a vector actually resides; it is completely determined by its direction and its length.

tensor

A *rank-n tensor* **T** is, by definition, a real-valued linear function of n vectors.[7] Pictorially we regard **T** as a box (Fig. 1.2) with n slots in its top, into which are inserted n vectors, and one slot in its end, which prints out a single real number: the value that the tensor **T** has when evaluated as a function of the n inserted vectors. Notationally we denote the tensor by a boldfaced sans-serif character **T**:

$$\mathsf{T}(\underbrace{\_, \_, \_, \_}) \tag{1.3a}$$

$\nwarrow$ n slots in which to put the vectors.

This definition of a tensor is very different (and far simpler) than the one found in most standard physics textbooks (e.g., Marion and Thornton, 1995; Jackson, 1999; Griffiths, 1999). There, a tensor is an array of numbers that transform in a particular way under rotations. We shall learn the connection between these definitions in Sec. 1.6 below.

To illustrate this approach, if **T** is a rank-3 tensor (has 3 slots) as in Fig. 1.2, then its value on the vectors **A**, **B**, **C** is denoted $\mathsf{T}(\mathbf{A}, \mathbf{B}, \mathbf{C})$. Linearity of this function can be expressed as

$$\mathsf{T}(e\mathbf{E} + f\mathbf{F}, \mathbf{B}, \mathbf{C}) = e\mathsf{T}(\mathbf{E}, \mathbf{B}, \mathbf{C}) + f\mathsf{T}(\mathbf{F}, \mathbf{B}, \mathbf{C}), \tag{1.3b}$$

where e and f are real numbers, and similarly for the second and third slots.

inner product

We have already defined the squared length $(\mathbf{A})^2 \equiv \mathbf{A}^2$ of a vector **A** as the squared distance between the points at its tail and its tip. The *inner product* (also called the dot product) $\mathbf{A} \cdot \mathbf{B}$ of two vectors is defined in terms of this squared length by

$$\boxed{\mathbf{A} \cdot \mathbf{B} \equiv \frac{1}{4}\Big[(\mathbf{A} + \mathbf{B})^2 - (\mathbf{A} - \mathbf{B})^2\Big].} \tag{1.4a}$$

In Euclidean space, this is the standard inner product, familiar from elementary geometry and discussed above in terms of the area of a parallelogram.

6. This is not so in curved spaces, as we shall see in Sec. 24.3.4.
7. This is a different use of the word *rank* than for a matrix, whose rank is its number of linearly independent rows or columns.

One can show that the inner product (1.4a) is a real-valued linear function of each of its vectors. Therefore, we can regard it as a tensor of rank 2. When so regarded, the inner product is denoted $\mathbf{g}(\_\_, \_\_)$ and is called the *metric tensor*. In other words, the metric tensor $\mathbf{g}$ is that linear function of two vectors whose value is given by

metric tensor

$$\boxed{\mathbf{g}(\mathbf{A}, \mathbf{B}) \equiv \mathbf{A} \cdot \mathbf{B}.} \tag{1.4b}$$

Notice that, because $\mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{A}$, the metric tensor is *symmetric* in its two slots—one gets the same real number independently of the order in which one inserts the two vectors into the slots:

$$\mathbf{g}(\mathbf{A}, \mathbf{B}) = \mathbf{g}(\mathbf{B}, \mathbf{A}). \tag{1.4c}$$

With the aid of the inner product, we can regard any vector $\mathbf{A}$ as a tensor of rank one: the real number that is produced when an arbitrary vector $\mathbf{C}$ is inserted into $\mathbf{A}$'s single slot is

$$\boxed{\mathbf{A}(\mathbf{C}) \equiv \mathbf{A} \cdot \mathbf{C}.} \tag{1.4d}$$

In Newtonian physics, we rarely meet tensors of rank higher than two. However, second-rank tensors appear frequently—often in roles where one sticks a single vector into the second slot and leaves the first slot empty, thereby producing a single-slotted entity, a vector. An example that we met in Sec. 1.1.1 is a rigid body's moment-of-inertia tensor $\mathbf{I}(\_\_, \_\_)$, which gives us the body's angular momentum $\mathbf{J}(\_\_) = \mathbf{I}(\_\_, \mathbf{\Omega})$ when its angular velocity $\mathbf{\Omega}$ is inserted into its second slot.[8] Another example is the stress tensor of a solid, a fluid, a plasma, or a field (Sec. 1.9 below).

tensor product

From three vectors $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, we can construct a tensor, their *tensor product* (also called *outer product* in contradistinction to the inner product $\mathbf{A} \cdot \mathbf{B}$), defined as follows:

$$\boxed{\mathbf{A} \otimes \mathbf{B} \otimes \mathbf{C}(\mathbf{E}, \mathbf{F}, \mathbf{G}) \equiv \mathbf{A}(\mathbf{E})\mathbf{B}(\mathbf{F})\mathbf{C}(\mathbf{G}) = (\mathbf{A} \cdot \mathbf{E})(\mathbf{B} \cdot \mathbf{F})(\mathbf{C} \cdot \mathbf{G}).} \tag{1.5a}$$

Here the first expression is the notation for the value of the new tensor, $\mathbf{A} \otimes \mathbf{B} \otimes \mathbf{C}$ evaluated on the three vectors $\mathbf{E}$, $\mathbf{F}$, $\mathbf{G}$; the middle expression is the ordinary product of three real numbers, the value of $\mathbf{A}$ on $\mathbf{E}$, the value of $\mathbf{B}$ on $\mathbf{F}$, and the value of $\mathbf{C}$ on $\mathbf{G}$; and the third expression is that same product with the three numbers rewritten as scalar products. Similar definitions can be given (and should be obvious) for the tensor product of any number of vectors, and of any two or more tensors of any rank; for example, if $\mathbf{T}$ has rank 2 and $\mathbf{S}$ has rank 3, then

$$\mathbf{T} \otimes \mathbf{S}(\mathbf{E}, \mathbf{F}, \mathbf{G}, \mathbf{H}, \mathbf{J}) \equiv \mathbf{T}(\mathbf{E}, \mathbf{F})\mathbf{S}(\mathbf{G}, \mathbf{H}, \mathbf{J}). \tag{1.5b}$$

contraction

One last geometric (i.e., frame-independent) concept we shall need is *contraction*. We illustrate this concept first by a simple example, then give the general definition.

8. Actually, it doesn't matter which slot, since $\mathbf{I}$ is symmetric.

From two vectors $\mathbf{A}$ and $\mathbf{B}$ we can construct the tensor product $\mathbf{A} \otimes \mathbf{B}$ (a second-rank tensor), and we can also construct the scalar product $\mathbf{A} \cdot \mathbf{B}$ (a real number, i.e., a *scalar,* also known as a *rank-0 tensor*). The process of contraction is the construction of $\mathbf{A} \cdot \mathbf{B}$ from $\mathbf{A} \otimes \mathbf{B}$:

$$\boxed{\text{contraction}(\mathbf{A} \otimes \mathbf{B}) \equiv \mathbf{A} \cdot \mathbf{B}.} \tag{1.6a}$$

One can show fairly easily using component techniques (Sec. 1.5 below) that any second-rank tensor $\mathbf{T}$ can be expressed as a sum of tensor products of vectors, $\mathbf{T} = \mathbf{A} \otimes \mathbf{B} + \mathbf{C} \otimes \mathbf{D} + \ldots$. Correspondingly, it is natural to define the contraction of $\mathbf{T}$ to be $\text{contraction}(\mathbf{T}) = \mathbf{A} \cdot \mathbf{B} + \mathbf{C} \cdot \mathbf{D} + \ldots$. Note that this contraction process lowers the rank of the tensor by two, from 2 to 0. Similarly, for a tensor of rank n one can construct a tensor of rank $n - 2$ by contraction, but in this case one must specify which slots are to be contracted. For example, if $\mathbf{T}$ is a third-rank tensor, expressible as $\mathbf{T} = \mathbf{A} \otimes \mathbf{B} \otimes \mathbf{C} + \mathbf{E} \otimes \mathbf{F} \otimes \mathbf{G} + \ldots$, then the contraction of $\mathbf{T}$ on its first and third slots is the rank-1 tensor (vector)

$$1\&3\text{contraction}(\mathbf{A} \otimes \mathbf{B} \otimes \mathbf{C} + \mathbf{E} \otimes \mathbf{F} \otimes \mathbf{G} + \ldots) \equiv (\mathbf{A} \cdot \mathbf{C})\mathbf{B} + (\mathbf{E} \cdot \mathbf{G})\mathbf{F} + \ldots. \tag{1.6b}$$

Unfortunately, there is no simple index-free notation for contraction in common use.

All the concepts developed in this section (vector, tensor, metric tensor, inner product, tensor product, and contraction of a tensor) can be carried over, with no change whatsoever, into any vector space[9] that is endowed with a concept of squared length—for example, to the 4-dimensional spacetime of special relativity (next chapter).

1.4 Particle Kinetics and Lorentz Force in Geometric Language

In this section, we illustrate our geometric viewpoint by formulating Newton's laws of motion for particles.

In Newtonian physics, a classical particle moves through Euclidean 3-space as universal time t passes. At time t it is located at some point $\mathbf{x}(t)$ (its *position*). The function $\mathbf{x}(t)$ represents a curve in 3-space, the particle's *trajectory.* The particle's *velocity* $\mathbf{v}(t)$ is the time derivative of its position, its *momentum* $\mathbf{p}(t)$ is the product of its mass m and velocity, its *acceleration* $\mathbf{a}(t)$ is the time derivative of its velocity, and its *kinetic energy* $E(t)$ is half its mass times velocity squared:

trajectory, velocity, momentum, acceleration, and energy

$$\mathbf{v}(t) = \frac{d\mathbf{x}}{dt}, \quad \mathbf{p}(t) = m\mathbf{v}(t), \quad \mathbf{a}(t) = \frac{d\mathbf{v}}{dt} = \frac{d^2\mathbf{x}}{dt^2}, \quad E(t) = \frac{1}{2}m\mathbf{v}^2. \tag{1.7a}$$

9. Or, more precisely, any vector space over the real numbers. If the vector space's scalars are complex numbers, as in quantum mechanics, then slight changes are needed.

Since points in 3-space are geometric objects (defined independently of any coordinate system), so also are the trajectory $\mathbf{x}(t)$, the velocity, the momentum, the acceleration, and the energy. (Physically, of course, the velocity has an ambiguity; it depends on one's standard of rest.)

Newton's second law of motion states that the particle's momentum can change only if a force $\mathbf{F}$ acts on it, and that its change is given by

$$d\mathbf{p}/dt = m\mathbf{a} = \mathbf{F}. \tag{1.7b}$$

If the force is produced by an electric field $\mathbf{E}$ and magnetic field $\mathbf{B}$, then this law of motion in SI units takes the familiar Lorentz-force form

$$d\mathbf{p}/dt = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}). \tag{1.7c}$$

(Here we have used the vector cross product, with which the reader should be familiar, and which will be discussed formally in Sec. 1.7.)

laws of motion

The laws of motion (1.7) are geometric relationships among geometric objects. Let us illustrate this using something very familiar, planetary motion. Consider a light planet orbiting a heavy star. If there were no gravitational force, the planet would continue in a straight line with constant velocity $\mathbf{v}$ and speed $v = |\mathbf{v}|$, sweeping out area A at a rate $dA/dt = rv_t/2$, where r is the radius, and v_t is the tangential speed. Elementary geometry equates this to the constant $vb/2$, where b is the impact parameter—the smallest separation from the star. Now add a gravitational force $\mathbf{F}$ and let it cause a small radial impulse. A second application of geometry showed Newton that the product $rv_t/2$ is unchanged to first order in the impulse, and he recovered Kepler's second law ($dA/dt = \text{const}$) without introducing coordinates.[10]

Contrast this approach with one relying on coordinates. For example, one introduces an (r, ϕ) coordinate system, constructs a lagrangian and observes that the coordinate ϕ is ignorable; then the Euler-Lagrange equations immediately imply the conservation of angular momentum, which is equivalent to Kepler's second law. So, which of these two approaches is preferable? The answer is surely "both!" Newton wrote the *Principia* in the language of geometry at least partly for a reason that remains valid today: it brought him a quick understanding of fundamental laws of physics. Lagrange followed his coordinate-based path to the function that bears his name, because he wanted to solve problems in celestial mechanics that would not yield to

10. Continuing in this vein, when the force is inverse square, as it is for gravity and electrostatics, we can use Kepler's second law to argue that when the orbit turns through a succession of equal angles $d\theta$, its successive changes in velocity $d\mathbf{v} = \mathbf{a}dt$ (with $\mathbf{a}$ the gravitational acceleration) all have the same magnitude $|d\mathbf{v}|$ and have the same angles $d\theta$ from one to another. So, if we trace the head of the velocity vector in velocity space, it follows a circle. The circle is not centered on zero velocity when the eccentricity is nonzero but there exists a reference frame in which the speed of the planet is constant. This graphical representation is known as a *hodograph,* and similar geometrical approaches are used in fluid mechanics. For Richard Feynman's masterful presentation of these ideas to first-year undergraduates, see Goodstein and Goodstein (1996).

Newton's approach. So it is today. Geometry and analysis are both indispensible. In the domain of classical physics, the geometry is of greater importance in deriving and understanding fundamental laws and has arguably been underappreciated; coordinates hold sway when we apply these laws to solve real problems. Today, both old and new laws of physics are commonly expressed geometrically, using lagrangians, hamiltonians, and actions, for example Hamilton's action principle $\delta \int L dt = 0$ where L is the coordinate-independent lagrangian. Indeed, being able to do this without introducing coordinates is a powerful guide to deriving these laws and a tool for comprehending their implications.

symmetry and conservation laws

A comment is needed on the famous connection between *symmetry* and *conservation laws*. In our example above, angular momentum conservation followed from axial symmetry which was embodied in the lagrangian's independence of the angle ϕ; but we also deduced it geometrically. This is usually the case in classical physics; typically, we do not need to introduce a specific coordinate system to understand symmetry and to express the associated conservation laws. However, symmetries are sometimes well hidden, for example with a nutating top, and coordinate transformations are then usually the best approach to uncover them.

Often in classical physics, real-world factors invalidate or complicate Lagrange's and Hamilton's coordinate-based analytical dynamics, and so one is driven to geometric considerations. As an example, consider a spherical marble rolling on a flat horizontal table. The analytical dynamics approach is to express the height of the marble's center of mass and the angle of its rotation as constraints and align the basis vectors so there is a single horizontal coordinate defined by the initial condition. It is then deduced that linear and angular momenta are conserved. Of course that result is trivial and just as easily gotten without this formalism. However, this model is also used for many idealized problems where the outcome is far from obvious and the approach is brilliantly effective. But consider the real world in which tables are warped and bumpy, marbles are ellipsoidal and scratched, air imposes a resistance, and wood and glass comprise polymers that attract one another. And so on. When one includes these factors, it is to geometry that one quickly turns to understand the real marble's actual dynamics. Even ignoring these effects and just asking what happens when the marble rolls off the edge of a table introduces a *nonholonomic* constraint, and figuring out where it lands and how fast it is spinning are best addressed not by the methods of Lagrange and Hamilton, but instead by considering the geometry of the gravitational and reaction forces. In the following chapters, we shall encounter many examples where we have to deal with messy complications like these.

EXERCISES

Exercise 1.1 *Practice: Energy Change for Charged Particle*
Without introducing any coordinates or basis vectors, show that when a particle with charge q interacts with electric and magnetic fields, its kinetic energy changes at a rate

$$dE/dt = q\,\mathbf{v}\cdot\mathbf{E}. \tag{1.8}$$

Exercise 1.2 *Practice: Particle Moving in a Circular Orbit*

Consider a particle moving in a circle with uniform speed $v = |\mathbf{v}|$ and uniform magnitude $a = |\mathbf{a}|$ of acceleration. Without introducing any coordinates or basis vectors, do the following.

(a) At any moment of time, let $\mathbf{n} = \mathbf{v}/v$ be the unit vector pointing along the velocity, and let s denote distance that the particle travels in its orbit. By drawing a picture, show that $d\mathbf{n}/ds$ is a unit vector that points to the center of the particle's circular orbit, divided by the radius of the orbit.

(b) Show that the vector (not unit vector) pointing from the particle's location to the center of its orbit is $(v/a)^2\mathbf{a}$.

1.5 Component Representation of Tensor Algebra

1.5

Cartesian coordinates and orthonormal basis vectors

In the Euclidean 3-space of Newtonian physics, there is a unique set of *orthonormal basis vectors* $\{\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\} \equiv \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ associated with any *Cartesian coordinate system* $\{x, y, z\} \equiv \{x^1, x^2, x^3\} \equiv \{x_1, x_2, x_3\}$. (In Cartesian coordinates in Euclidean space, we usually place indices down, but occasionally we place them up. It doesn't matter. By definition, in Cartesian coordinates a quantity is the same whether its index is down or up.) The basis vector $\mathbf{e}_j$ points along the x_j coordinate direction, which is orthogonal to all the other coordinate directions, and it has unit length (Fig. 1.3), so

$$\boxed{\mathbf{e}_j \cdot \mathbf{e}_k = \delta_{jk},} \tag{1.9a}$$

where δ_{jk} is the Kronecker delta.

Any vector $\mathbf{A}$ in 3-space can be expanded in terms of this basis:

$$\mathbf{A} = A_j \mathbf{e}_j. \tag{1.9b}$$

Einstein summation convention

Cartesian components of a vector

Here and throughout this book, we adopt the *Einstein summation convention:* repeated indices (in this case j) are to be summed (in this 3-space case over $j = 1, 2, 3$), unless otherwise instructed. By virtue of the orthonormality of the basis, the components A_j of $\mathbf{A}$ can be computed as the scalar product

$$A_j = \mathbf{A} \cdot \mathbf{e}_j. \tag{1.9c}$$

[The proof of this is straightforward: $\mathbf{A} \cdot \mathbf{e}_j = (A_k \mathbf{e}_k) \cdot \mathbf{e}_j = A_k(\mathbf{e}_k \cdot \mathbf{e}_j) = A_k \delta_{kj} = A_j$.]

Any tensor, say, the third-rank tensor $\mathbf{T}(\_\_, \_\_, \_\_)$, can be expanded in terms of tensor products of the basis vectors:

$$\boxed{\mathbf{T} = T_{ijk} \mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k.} \tag{1.9d}$$

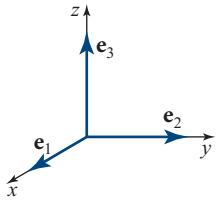

FIGURE 1.3 The orthonormal basis vectors $\mathbf{e}_j$ associated with a Euclidean coordinate system in Euclidean 3-space.

Cartesian components of a tensor

The components T_{ijk} of $\mathbf{T}$ can be computed from $\mathbf{T}$ and the basis vectors by the generalization of Eq. (1.9c):

$$\boxed{T_{ijk} = \mathbf{T}(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k).} \tag{1.9e}$$

[This equation can be derived using the orthonormality of the basis in the same way as Eq. (1.9c) was derived.] As an important example, the components of the metric tensor are $g_{jk} = \mathbf{g}(\mathbf{e}_j, \mathbf{e}_k) = \mathbf{e}_j \cdot \mathbf{e}_k = \delta_{jk}$ [where the first equality is the method (1.9e) of computing tensor components, the second is the definition (1.4b) of the metric, and the third is the orthonormality relation (1.9a)]:

$$\boxed{g_{jk} = \delta_{jk}.} \tag{1.9f}$$

The components of a tensor product [e.g., $\mathbf{T}(\_\_, \_\_, \_\_) \otimes \mathbf{S}(\_\_, \_\_)$] are easily deduced by inserting the basis vectors into the slots [Eq. (1.9e)]; they are $\mathbf{T}(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k) \otimes \mathbf{S}(\mathbf{e}_l, \mathbf{e}_m) = T_{ijk}S_{lm}$ [cf. Eq. (1.5a)]. In words, the components of a tensor product are equal to the ordinary arithmetic product of the components of the individual tensors.

In component notation, the inner product of two vectors and the value of a tensor when vectors are inserted into its slots are given by

$$\boxed{\mathbf{A} \cdot \mathbf{B} = A_j B_j, \qquad \mathbf{T}(\mathbf{A}, \mathbf{B}, \mathbf{C}) = T_{ijk}A_i B_j C_k,} \tag{1.9g}$$

as one can easily show using previous equations. Finally, the contraction of a tensor [say, the fourth-rank tensor $\mathbf{R}(\_\_, \_\_, \_\_, \_\_)$] on two of its slots (say, the first and third) has components that are easily computed from the tensor's own components:

$$\text{components of [1\&3contraction of } \mathbf{R}] = R_{ijik}. \tag{1.9h}$$

Note that R_{ijik} is summed on the i index, so it has only two free indices, j and k, and thus is the component of a second-rank tensor, as it must be if it is to represent the contraction of a fourth-rank tensor.

1.5.1 Slot-Naming Index Notation

1.5.1

We now pause in our development of the component version of tensor algebra to introduce a very important new viewpoint.

BOX 1.2. VECTORS AND TENSORS IN QUANTUM THEORY T2

The laws of quantum theory, like all other laws of Nature, can be expressed as geometric relationships among geometric objects. Most of quantum theory's geometric objects, like those of classical theory, are vectors and tensors: the quantum state $|\psi\rangle$ of a physical system (e.g., a particle in a harmonic-oscillator potential) is a Hilbert-space vector—a generalization of a Euclidean-space vector $\mathbf{A}$. There is an inner product, denoted $\langle\phi|\psi\rangle$, between any two states $|\phi\rangle$ and $|\psi\rangle$, analogous to $\mathbf{B}\cdot\mathbf{A}$; but $\mathbf{B}\cdot\mathbf{A}$ is a real number, whereas $\langle\phi|\psi\rangle$ is a complex number (and we add and subtract quantum states with complex-number coefficients). The Hermitian operators that represent observables (e.g., the hamiltonian $\hat{H}$ for the particle in the potential) are two-slotted (second-rank), complex-valued functions of vectors; $\langle\phi|\hat{H}|\psi\rangle$ is the complex number that one gets when one inserts ϕ and ψ into the first and second slots of $\hat{H}$. Just as, in Euclidean space, we get a new vector (first-rank tensor) $\mathbf{T}(\_\_, \mathbf{A})$ when we insert the vector $\mathbf{A}$ into the second slot of $\mathbf{T}$, so in quantum theory we get a new vector (physical state) $\hat{H}|\psi\rangle$ (the result of letting $\hat{H}$ "act on" $|\psi\rangle$) when we insert $|\psi\rangle$ into the second slot of $\hat{H}$. In these senses, we can regard $\mathbf{T}$ as a linear map of Euclidean vectors into Euclidean vectors and $\hat{H}$ as a linear map of states (Hilbert-space vectors) into states.

For the electron in the hydrogen atom, we can introduce a set of orthonormal basis vectors $\{|1\rangle, |2\rangle, |3\rangle, \ldots\}$, that is, the atom's energy eigenstates, with $\langle m|n\rangle = \delta_{mn}$. But by contrast with Newtonian physics, where we only need three basis vectors (because our Euclidean space is 3-dimensional), for the particle in a harmonic-oscillator potential, we need an infinite number of basis vectors (since the Hilbert space of all states is infinite-dimensional). In the particle's quantum-state basis, any observable (e.g., the particle's position $\hat{x}$ or momentum $\hat{p}$) has components computed by inserting the basis vectors into its two slots: $x_{mn} = \langle m|\hat{x}|n\rangle$, and $p_{mn} = \langle m|\hat{p}|n\rangle$. In this basis, the operator $\hat{x}\,\hat{p}$ (which maps states into states) has components $x_{jk}p_{km}$ (a matrix product), and the noncommutation of position and momentum $[\hat{x}, \hat{p}] = i\hbar$ (an important physical law) is expressible in terms of components as $x_{jk}p_{km} - p_{jk}x_{km} = i\hbar\delta_{jm}$.

Consider the rank-2 tensor $\mathbf{F}(\_\_, \_\_)$. We can define a new tensor $\mathbf{G}(\_\_, \_\_)$ to be the same as $\mathbf{F}$, but with the slots interchanged: i.e., for any two vectors $\mathbf{A}$ and $\mathbf{B}$, it is true that $\mathbf{G}(\mathbf{A}, \mathbf{B}) = \mathbf{F}(\mathbf{B}, \mathbf{A})$. We need a simple, compact way to indicate that $\mathbf{F}$ and $\mathbf{G}$ are equal except for an interchange of slots. The best way is to give the slots names, say a and b—i.e., to rewrite $\mathbf{F}(\_\_, \_\_)$ as $\mathbf{F}(\_\__a, \_\__b)$ or more conveniently as F_{ab}, and then to write the relationship between $\mathbf{G}$ and $\mathbf{F}$ as $G_{ab} = F_{ba}$. "NO!" some readers

might object. This notation is indistinguishable from our notation for components on a particular basis. "GOOD!" a more astute reader will exclaim. The relation $G_{ab} = F_{ba}$ in a particular basis is a true statement if and only if "$\mathbf{G} = \mathbf{F}$ with slots interchanged" is true, so why not use the same notation to symbolize both? In fact, we shall do this. We ask our readers to look at any "index equation," such as $G_{ab} = F_{ba}$, like they would look at an Escher drawing: momentarily think of it as a relationship between components of tensors in a specific basis; then do a quick mind-flip and regard it quite differently, as a relationship between geometric, basis-independent tensors with the indices playing the roles of slot names. This mind-flip approach to tensor algebra will pay substantial dividends.

slot-naming index notation

As an example of the power of this *slot-naming index notation,* consider the contraction of the first and third slots of a third-rank tensor $\mathbf{T}$. In any basis the components of 1&3contraction($\mathbf{T}$) are T_{aba}; cf. Eq. (1.9h). Correspondingly, in slot-naming index notation we denote 1&3contraction($\mathbf{T}$) by the simple expression T_{aba}. We can think of the first and third slots as annihilating each other by the contraction, leaving free only the second slot (named b) and therefore producing a rank-1 tensor (a vector).

We should caution that the phrase "slot-naming index notation" is unconventional. You are unlikely to find it in any other textbooks. However, we like it. It says precisely what we want it to say.

1.5.2 Particle Kinetics in Index Notation

1.5.2

As an example of slot-naming index notation, we can rewrite the equations of particle kinetics (1.7) as follows:

$$v_i = \frac{dx_i}{dt}, \quad p_i = mv_i, \quad a_i = \frac{dv_i}{dt} = \frac{d^2x_i}{dt^2},$$
$$E = \frac{1}{2}mv_jv_j, \quad \frac{dp_i}{dt} = q(E_i + \epsilon_{ijk}v_jB_k). \tag{1.10}$$

(In the last equation ϵ_{ijk} is the so-called Levi-Civita tensor, which is used to produce the cross product; we shall learn about it in Sec. 1.7. And note that the scalar energy E must not be confused with the electric field vector E_i.)

Equations (1.10) can be viewed in either of two ways: (i) as the basis-independent geometric laws $\mathbf{v} = d\mathbf{x}/dt$, $\mathbf{p} = m\mathbf{v}$, $\mathbf{a} = d\mathbf{v}/dt = d^2\mathbf{x}/dt^2$, $E = \frac{1}{2}m\mathbf{v}^2$, and $d\mathbf{p}/dt = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$ written in slot-naming index notation; or (ii) as equations for the components of $\mathbf{v}$, $\mathbf{p}$, $\mathbf{a}$, $\mathbf{E}$, and $\mathbf{B}$ in some particular Cartesian coordinate system.

EXERCISES

Exercise 1.3 *Derivation: Component Manipulation Rules*
Derive the component manipulation rules (1.9g) and (1.9h).

Exercise 1.4 *Example and Practice: Numerics of Component Manipulations*
The third-rank tensor $\mathbf{S}(\_, \_, \_)$ and vectors $\mathbf{A}$ and $\mathbf{B}$ have as their only nonzero components $S_{123} = S_{231} = S_{312} = +1$, $A_1 = 3$, $B_1 = 4$, $B_2 = 5$. What are the

components of the vector $\mathbf{C} = \mathbf{S}(\mathbf{A}, \mathbf{B}, \_\_)$, the vector $\mathbf{D} = \mathbf{S}(\mathbf{A}, \_\_, \mathbf{B})$, and the tensor $\mathbf{W} = \mathbf{A} \otimes \mathbf{B}$?

[Partial solution: In component notation, $C_k = S_{ijk}A_iB_j$, where (of course) we sum over the repeated indices i and j. This tells us that $C_1 = S_{231}A_2B_3$, because S_{231} is the only component of $\mathbf{S}$ whose last index is a 1; this in turn implies that $C_1 = 0$, since $A_2 = 0$. Similarly, $C_2 = S_{312}A_3B_1 = 0$ (because $A_3 = 0$). Finally, $C_3 = S_{123}A_1B_2 = +1 \times 3 \times 5 = 15$. Also, in component notation $W_{ij} = A_iB_j$, so $W_{11} = A_1 \times B_1 = 3 \times 4 = 12$, and $W_{12} = A_1 \times B_2 = 3 \times 5 = 15$. Here the $\times$ stands for numerical multiplication, not the vector cross product.]

Exercise 1.5 *Practice: Meaning of Slot-Naming Index Notation*

(a) The following expressions and equations are written in slot-naming index notation. Convert them to geometric, index-free notation: A_iB_{jk}, A_iB_{ji}, $S_{ijk} = S_{kji}$, $A_iB_i = A_iB_jg_{ij}$.

(b) The following expressions are written in geometric, index-free notation. Convert them to slot-naming index notation: $\mathbf{T}(\_\_, \_\_, \mathbf{A})$, $\mathbf{T}(\_\_, \mathbf{S}(\mathbf{B}, \_\_), \_\_)$.

1.6 1.6 Orthogonal Transformations of Bases

Consider two different Cartesian coordinate systems $\{x, y, z\} \equiv \{x_1, x_2, x_3\}$, and $\{\bar{x}, \bar{y}, \bar{z}\} \equiv \{x_{\bar{1}}, x_{\bar{2}}, x_{\bar{3}}\}$. Denote by $\{\mathbf{e}_i\}$ and $\{\mathbf{e}_{\bar{p}}\}$ the corresponding bases. It is possible to expand the basis vectors of one basis in terms of those of the other. We denote the expansion coefficients by the letter R and write

$$\mathbf{e}_i = \mathbf{e}_{\bar{p}}R_{\bar{p}i}, \qquad \mathbf{e}_{\bar{p}} = \mathbf{e}_iR_{i\bar{p}}. \tag{1.11}$$

The quantities $R_{\bar{p}i}$ and $R_{i\bar{p}}$ are not the components of a tensor; rather, they are the elements of transformation matrices

$$[R_{\bar{p}i}] = \begin{bmatrix} R_{\bar{1}1} & R_{\bar{1}2} & R_{\bar{1}3} \\ R_{\bar{2}1} & R_{\bar{2}2} & R_{\bar{2}3} \\ R_{\bar{3}1} & R_{\bar{3}2} & R_{\bar{3}3} \end{bmatrix}, \qquad [R_{i\bar{p}}] = \begin{bmatrix} R_{1\bar{1}} & R_{1\bar{2}} & R_{1\bar{3}} \\ R_{2\bar{1}} & R_{2\bar{2}} & R_{2\bar{3}} \\ R_{3\bar{1}} & R_{3\bar{2}} & R_{3\bar{3}} \end{bmatrix}. \tag{1.12a}$$

(Here and throughout this book we use square brackets to denote matrices.) These two matrices must be the inverse of each other, since one takes us from the barred basis to the unbarred, and the other in the reverse direction, from unbarred to barred:

$$R_{\bar{p}i}R_{i\bar{q}} = \delta_{\bar{p}\bar{q}}, \qquad R_{i\bar{p}}R_{\bar{p}j} = \delta_{ij}. \tag{1.12b}$$

The orthonormality requirement for the two bases implies that $\delta_{ij} = \mathbf{e}_i \cdot \mathbf{e}_j = (\mathbf{e}_{\bar{p}}R_{\bar{p}i}) \cdot (\mathbf{e}_{\bar{q}}R_{\bar{q}j}) = R_{\bar{p}i}R_{\bar{q}j}(\mathbf{e}_{\bar{p}} \cdot \mathbf{e}_{\bar{q}}) = R_{\bar{p}i}R_{\bar{q}j}\delta_{\bar{p}\bar{q}} = R_{\bar{p}i}R_{\bar{p}j}$. This says that the transpose of $[R_{\bar{p}i}]$ is its inverse—which we have already denoted by $[R_{i\bar{p}}]$:

$$[R_{i\bar{p}}] \equiv \text{inverse}\left([R_{\bar{p}i}]\right) = \text{transpose}\left([R_{\bar{p}i}]\right). \tag{1.12c}$$

orthogonal transformation and rotation

This property implies that the transformation matrix is orthogonal, so the transformation is a reflection or a rotation (see, e.g., Goldstein, Poole, and Safko, 2002). Thus (as should be obvious and familiar), the bases associated with any two Euclidean coordinate systems are related by a reflection or rotation, and the matrices (1.12a) are called *rotation matrices*. Note that Eq. (1.12c) does not say that $[R_{i\bar{p}}]$ is a symmetric matrix. In fact, most rotation matrices are not symmetric [see, e.g., Eq. (1.14)].

The fact that a vector $\mathbf{A}$ is a geometric, basis-independent object implies that $\mathbf{A} = A_i\mathbf{e}_i = A_i(\mathbf{e}_{\bar{p}}R_{\bar{p}i}) = (R_{\bar{p}i}A_i)\mathbf{e}_{\bar{p}} = A_{\bar{p}}\mathbf{e}_{\bar{p}}$:

$$A_{\bar{p}} = R_{\bar{p}i}A_i, \quad \text{and similarly,} \quad A_i = R_{i\bar{p}}A_{\bar{p}}; \tag{1.13a}$$

and correspondingly for the components of a tensor:

$$T_{\bar{p}\bar{q}\bar{r}} = R_{\bar{p}i}R_{\bar{q}j}R_{\bar{r}k}T_{ijk}, \quad T_{ijk} = R_{i\bar{p}}R_{j\bar{q}}R_{k\bar{r}}T_{\bar{p}\bar{q}\bar{r}}. \tag{1.13b}$$

It is instructive to compare the transformation law (1.13a) for the components of a vector with Eqs. (1.11) for the bases. To make these laws look natural, we have placed the transformation matrix on the left in the former and on the right in the latter. In Minkowski spacetime (Chap. 2), the placement of indices, up or down, will automatically tell us the order.

If we choose the origins of our two coordinate systems to coincide, then the vector $\mathbf{x}$ reaching from the common origin to some point $\mathcal{P}$, whose coordinates are x_j and $x_{\bar{p}}$, has components equal to those coordinates; and as a result, the coordinates themselves obey the same transformation law as any other vector:

$$x_{\bar{p}} = R_{\bar{p}i}x_i, \quad x_i = R_{i\bar{p}}x_{\bar{p}}. \tag{1.13c}$$

The product of two rotation matrices $[R_{i\bar{p}}R_{\bar{p}\bar{s}}]$ is another rotation matrix $[R_{i\bar{s}}]$, which transforms the Cartesian bases $\mathbf{e}_{\bar{s}}$ to $\mathbf{e}_i$. Under this product rule, the rotation matrices form a mathematical *group:* the *rotation group,* whose *group representations* play an important role in quantum theory.

rotation group

EXERCISES

Exercise 1.6 ***Example and Practice: Rotation in x-y Plane***

Consider two Cartesian coordinate systems rotated with respect to each other in the x-y plane as shown in Fig. 1.4.

(a) Show that the rotation matrix that takes the barred basis vectors to the unbarred basis vectors is

$$[R_{\bar{p}i}] = \begin{bmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{1.14}$$

and show that the inverse of this rotation matrix is, indeed, its transpose, as it must be if this is to represent a rotation.

(b) Verify that the two coordinate systems are related by Eq. (1.13c).

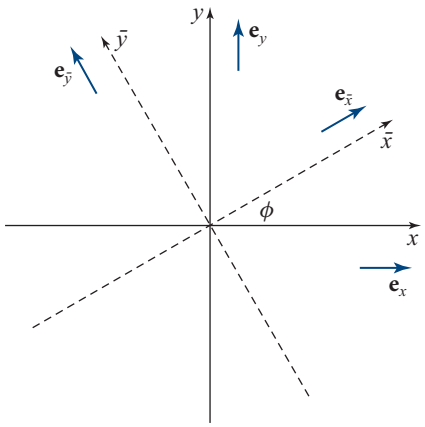

FIGURE 1.4 Two Cartesian coordinate systems $\{x, y, z\}$ and $\{\bar{x}, \bar{y}, \bar{z}\}$ and their basis vectors in Euclidean space, rotated by an angle ϕ relative to each other in the x-y plane. The z- and $\bar{z}$-axes point out of the paper or screen and are not shown.

(c) Let A_j be the components of the electromagnetic vector potential that lies in the x-y plane, so that $A_z = 0$. The two nonzero components A_x and A_y can be regarded as describing the two polarizations of an electromagnetic wave propagating in the z direction. Show that $A_{\bar{x}} + iA_{\bar{y}} = (A_x + iA_y)e^{-i\phi}$. One can show (cf. Sec. 27.3.3) that the factor $e^{-i\phi}$ implies that the quantum particle associated with the wave—the photon—has spin one [i.e., spin angular momentum $\hbar =$ (Planck's constant)$/2\pi$].

(d) Let h_{jk} be the components of a symmetric tensor that is *trace-free* (its contraction h_{jj} vanishes) and is confined to the x-y plane (so $h_{zk} = h_{kz} = 0$ for all k). Then the only nonzero components of this tensor are $h_{xx} = -h_{yy}$ and $h_{xy} = h_{yx}$. As we shall see in Sec. 27.3.1, this tensor can be regarded as describing the two polarizations of a gravitational wave propagating in the z direction. Show that $h_{\bar{x}\bar{x}} + ih_{\bar{x}\bar{y}} = (h_{xx} + ih_{xy})e^{-2i\phi}$. The factor $e^{-2i\phi}$ implies that the quantum particle associated with the gravitational wave (the graviton) has spin two (spin angular momentum $2\hbar$); cf. Eq. (27.31) and Sec. 27.3.3.

1.7 1.7 Differentiation of Scalars, Vectors, and Tensors; Cross Product and Curl

1.7

directional derivative

Consider a tensor field $\mathbf{T}(\mathcal{P})$ in Euclidean 3-space and a vector $\mathbf{A}$. We define the *directional derivative* of $\mathbf{T}$ along $\mathbf{A}$ by the obvious limiting procedure

$$\nabla_{\mathbf{A}}\mathbf{T} \equiv \lim_{\epsilon \to 0} \frac{1}{\epsilon}[\mathbf{T}(\mathbf{x}_{\mathcal{P}} + \epsilon\mathbf{A}) - \mathbf{T}(\mathbf{x}_{\mathcal{P}})] \tag{1.15a}$$

and similarly for the directional derivative of a vector field $\mathbf{B}(\mathcal{P})$ and a scalar field $\psi(\mathcal{P})$. [Here we have denoted points, e.g., $\mathcal{P}$, by the vector $\mathbf{x}_{\mathcal{P}}$ that reaches from some

arbitrary origin to the point, and $\mathbf{T}(\mathbf{x}_{\mathcal{P}})$ denotes the field's dependence on location in space; $\mathbf{T}$'s slots and dependence on what goes into the slots are suppressed; and the units of ϵ are chosen to ensure that $\epsilon\mathbf{A}$ has the same units as $\mathbf{x}_{\mathcal{P}}$. There is no other appearance of vectors in this chapter.] In definition (1.15a), the quantity in square brackets is simply the difference between two linear functions of vectors (two tensors), so the quantity on the left-hand side is also a tensor with the same rank as $\mathbf{T}$.

It should not be hard to convince oneself that this directional derivative $\nabla_{\mathbf{A}}\mathbf{T}$ of any tensor field $\mathbf{T}$ is linear in the vector $\mathbf{A}$ along which one differentiates. Correspondingly, if $\mathbf{T}$ has rank n (n slots), then there is another tensor field, denoted $\boldsymbol{\nabla}\mathbf{T}$, with rank $n+1$, such that

$$\boxed{\nabla_{\mathbf{A}}\mathbf{T} = \boldsymbol{\nabla}\mathbf{T}(\_\_, \_\_, \_\_, \mathbf{A}).} \tag{1.15b}$$

gradient

Here on the right-hand side the first n slots (3 in the case shown) are left empty, and $\mathbf{A}$ is put into the last slot (the "differentiation slot"). The quantity $\boldsymbol{\nabla}\mathbf{T}$ is called the *gradient* of $\mathbf{T}$. In slot-naming index notation, it is conventional to denote this gradient by $T_{abc;d}$, where in general the number of indices preceding the semicolon is the rank of $\mathbf{T}$. Using this notation, the directional derivative of $\mathbf{T}$ along $\mathbf{A}$ reads [cf. Eq. (1.15b)] $T_{abc;j}A_j$.

It is not hard to show that in any Cartesian coordinate system, the components of the gradient are nothing but the partial derivatives of the components of the original tensor, which we denote by a comma:

$$\boxed{T_{abc;j} = \frac{\partial T_{abc}}{\partial x_j} \equiv T_{abc,\,j}.} \tag{1.15c}$$

In a non-Cartesian basis (e.g., the spherical and cylindrical bases often used in electromagnetic theory), the components of the gradient typically are not obtained by simple partial differentiation [Eq. (1.15c) fails] because of turning and/or length changes of the basis vectors as we go from one location to another. In Sec. 11.8, we shall learn how to deal with this by using objects called *connection coefficients*. Until then, we confine ourselves to Cartesian bases, so subscript semicolons and subscript commas (partial derivatives) can be used interchangeably.

Because the gradient and the directional derivative are defined by the same standard limiting process as one uses when defining elementary derivatives, they obey the standard (Leibniz) rule for differentiating products:

$$\begin{aligned} &\nabla_{\mathbf{A}}(\mathbf{S}\otimes\mathbf{T}) = (\nabla_{\mathbf{A}}\mathbf{S})\otimes\mathbf{T} + \mathbf{S}\otimes\nabla_{\mathbf{A}}\mathbf{T}, \\ \text{or}\quad &(S_{ab}T_{cde})_{;j}A_j = (S_{ab;j}A_j)T_{cde} + S_{ab}(T_{cde;j}A_j); \end{aligned} \tag{1.16a}$$

and

$$\nabla_{\mathbf{A}}(f\mathbf{T}) = (\nabla_{\mathbf{A}}f)\mathbf{T} + f\nabla_{\mathbf{A}}\mathbf{T}, \quad \text{or} \quad (fT_{abc})_{;j}A_j = (f_{;j}A_j)T_{abc} + fT_{abc;j}A_j. \tag{1.16b}$$

In an orthonormal basis these relations should be obvious: they follow from the Leibniz rule for partial derivatives.

Because the components g_{ab} of the metric tensor are constant in any Cartesian coordinate system, Eq. (1.15c) (which is valid in such coordinates) guarantees that $g_{ab;j} = 0$; i.e., the metric has vanishing gradient:

$$\nabla \mathbf{g} = 0, \quad \text{or} \quad g_{ab;j} = 0. \tag{1.17}$$

From the gradient of any vector or tensor we can construct several other important derivatives by contracting on slots:

1. Since the gradient $\nabla \mathbf{A}$ of a vector field $\mathbf{A}$ has two slots, $\nabla \mathbf{A}(\_\_, \_\_)$, we can contract its slots on each other to obtain a scalar field. That scalar field is the *divergence* of $\mathbf{A}$ and is denoted

divergence

$$\nabla \cdot \mathbf{A} \equiv (\text{contraction of } \nabla \mathbf{A}) = A_{a;a}. \tag{1.18}$$

2. Similarly, if $\mathbf{T}$ is a tensor field of rank 3, then $T_{abc;c}$ is its divergence on its third slot, and $T_{abc;b}$ is its divergence on its second slot.
3. By taking the double gradient and then contracting on the two gradient slots we obtain, from any tensor field $\mathbf{T}$, a new tensor field with the same rank,

laplacian

$$\nabla^2 \mathbf{T} \equiv (\nabla \cdot \nabla)\mathbf{T}, \quad \text{or } T_{abc;jj}. \tag{1.19}$$

Here and henceforth, all indices following a semicolon (or comma) represent gradients (or partial derivatives): $T_{abc;jj} \equiv T_{abc;j;j}$, $T_{abc,\,jk} \equiv \partial^2 T_{abc}/\partial x_j \partial x_k$. The operator ∇^2 is called the *laplacian.*

The metric tensor is a fundamental property of the space in which it lives; it embodies the inner product and hence the space's notion of distance. In addition to the metric, there is one (and only one) other fundamental tensor that describes a piece of Euclidean space's geometry: the *Levi-Civita tensor* $\boldsymbol{\epsilon}$, which embodies the space's notion of volume.

Levi-Civita tensor

In a Euclidean space with dimension n, the Levi-Civita tensor $\boldsymbol{\epsilon}$ is a completely antisymmetric tensor with rank n (with n slots). A parallelepiped whose edges are the n vectors $\mathbf{A}, \mathbf{B}, \ldots, \mathbf{F}$ is said to have the *volume*

volume

$$\boxed{\text{volume} = \boldsymbol{\epsilon}(\mathbf{A}, \mathbf{B}, \ldots, \mathbf{F}).} \tag{1.20}$$

(We justify this definition in Sec. 1.8.) Notice that this volume can be positive or negative, and if we exchange the order of the parallelepiped's legs, the volume's sign changes: $\boldsymbol{\epsilon}(\mathbf{B}, \mathbf{A}, \ldots, \mathbf{F}) = -\boldsymbol{\epsilon}(\mathbf{A}, \mathbf{B}, \ldots, \mathbf{F})$ by antisymmetry of $\boldsymbol{\epsilon}$.

It is easy to see (Ex. 1.7) that (i) the volume vanishes unless the legs are all linearly independent, (ii) once the volume has been specified for one parallelepiped (one set of linearly independent legs), it is thereby determined for all parallelepipeds, and therefore, (iii) we require only one number plus antisymmetry to determine $\boldsymbol{\epsilon}$

fully. If the chosen parallelepiped has legs that are orthonormal (all are orthogonal to one another and all have unit length—properties determined by the metric $\mathbf{g}$), then it must have unit volume, or more precisely volume ± 1. This is a compatibility relation between $\mathbf{g}$ and $\boldsymbol{\epsilon}$. It is easy to see (Ex. 1.7) that (iv) $\boldsymbol{\epsilon}$ is fully determined by its antisymmetry, compatibility with the metric, and a single sign: the choice of which parallelepipeds have positive volume and which have negative. It is conventional in Euclidean 3-space to give right-handed parallelepipeds positive volume and left-handed ones negative volume: $\boldsymbol{\epsilon}(\mathbf{A}, \mathbf{B}, \mathbf{C})$ is positive if, when we place our right thumb along $\mathbf{C}$ and the fingers of our right hand along $\mathbf{A}$, then bend our fingers, they sweep toward $\mathbf{B}$ and not $-\mathbf{B}$.

These considerations dictate that in a right-handed orthonormal basis of Euclidean 3-space, the only nonzero components of $\boldsymbol{\epsilon}$ are

$$\epsilon_{123} = +1,$$
$$\epsilon_{abc} = \begin{cases} +1 & \text{if } a, b, c \text{ is an even permutation of 1, 2, 3} \\ -1 & \text{if } a, b, c \text{ is an odd permutation of 1, 2, 3} \\ 0 & \text{if } a, b, c \text{ are not all different;} \end{cases} \tag{1.21}$$

and in a left-handed orthonormal basis, the signs of these components are reversed.

cross product and curl

The Levi-Civita tensor is used to define the cross product and the curl:

$$\mathbf{A} \times \mathbf{B} \equiv \boldsymbol{\epsilon}(\_\_, \mathbf{A}, \mathbf{B}); \quad \text{in slot-naming index notation, } \epsilon_{ijk} A_j B_k; \tag{1.22a}$$

$$\boldsymbol{\nabla} \times \mathbf{A} \equiv (\text{the vector field whose slot-naming index form is } \epsilon_{ijk} A_{k;j}). \tag{1.22b}$$

[Equation (1.22b) is an example of an expression that is complicated if stated in index-free notation; it says that $\boldsymbol{\nabla} \times \mathbf{A}$ is the double contraction of the rank-5 tensor $\boldsymbol{\epsilon} \otimes \boldsymbol{\nabla}\mathbf{A}$ on its second and fifth slots, and on its third and fourth slots.]

Although Eqs. (1.22a) and (1.22b) look like complicated ways to deal with concepts that most readers regard as familiar and elementary, they have great power. The power comes from the following property of the Levi-Civita tensor in Euclidean 3-space [readily derivable from its components (1.21)]:

$$\boxed{\epsilon_{ijm}\epsilon_{klm} = \delta^{ij}_{kl} \equiv \delta^i_k \delta^j_l - \delta^i_l \delta^j_k.} \tag{1.23}$$

Here δ^i_k is the Kronecker delta. Examine the 4-index delta function δ^{ij}_{kl} carefully; it says that either the indices above and below each other must be the same ($i = k$ and $j = l$) with a $+$ sign, or the diagonally related indices must be the same ($i = l$ and $j = k$) with a $-$ sign. [We have put the indices ij of δ^{ij}_{kl} up solely to facilitate remembering this rule. Recall (first paragraph of Sec. 1.5) that in Euclidean space and Cartesian coordinates, it does not matter whether indices are up or down.] With the aid of Eq. (1.23) and the index-notation expressions for the cross product and curl, one can quickly and easily derive a wide variety of useful vector identities; see the very important Ex. 1.8.

EXERCISES

Exercise 1.7 *Derivation: Properties of the Levi-Civita Tensor*
From its complete antisymmetry, derive the four properties of the Levi-Civita tensor, in n-dimensional Euclidean space, that are claimed in the text following Eq. (1.20).

Exercise 1.8 ***Example and Practice: Vectorial Identities for the Cross Product and Curl*
Here is an example of how to use index notation to derive a vector identity for the double cross product $\mathbf{A} \times (\mathbf{B} \times \mathbf{C})$: in index notation this quantity is $\epsilon_{ijk} A_j (\epsilon_{klm} B_l C_m)$. By permuting the indices on the second ϵ and then invoking Eq. (1.23), we can write this as $\epsilon_{ijk}\epsilon_{lmk} A_j B_l C_m = \delta^{lm}_{ij} A_j B_l C_m$. By then invoking the meaning of the 4-index delta function [Eq. (1.23)], we bring this into the form $A_j B_i C_j - A_j B_j C_i$, which is the slot-naming index-notation form of $(\mathbf{A} \cdot \mathbf{C})\mathbf{B} - (\mathbf{A} \cdot \mathbf{B})\mathbf{C}$. Thus, it must be that $\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = (\mathbf{A} \cdot \mathbf{C})\mathbf{B} - (\mathbf{A} \cdot \mathbf{B})\mathbf{C}$. Use similar techniques to evaluate the following quantities.

(a) $\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \mathbf{A})$.

(b) $(\mathbf{A} \times \mathbf{B}) \cdot (\mathbf{C} \times \mathbf{D})$.

(c) $(\mathbf{A} \times \mathbf{B}) \times (\mathbf{C} \times \mathbf{D})$.

Exercise 1.9 ***Example and Practice: Levi-Civita Tensor in 2-Dimensional Euclidean Space*
In Euclidean 2-space, let $\{\mathbf{e}_1, \mathbf{e}_2\}$ be an orthonormal basis with positive volume.

(a) Show that the components of $\boldsymbol{\epsilon}$ in this basis are

$$\epsilon_{12} = +1, \qquad \epsilon_{21} = -1, \qquad \epsilon_{11} = \epsilon_{22} = 0. \tag{1.24a}$$

(b) Show that

$$\epsilon_{ik}\epsilon_{jk} = \delta_{ij}. \tag{1.24b}$$

1.8 1.8 Volumes, Integration, and Integral Conservation Laws

In Cartesian coordinates of 2-dimensional Euclidean space, the basis vectors are orthonormal, so (with a conventional choice of sign) the components of the Levi-Civita tensor are given by Eqs. (1.24a). Correspondingly, the area (i.e., 2-dimensional volume) of a parallelogram whose sides are $\mathbf{A}$ and $\mathbf{B}$ is

$$\text{2-volume} = \boldsymbol{\epsilon}(\mathbf{A}, \mathbf{B}) = \epsilon_{ab} A_a B_b = A_1 B_2 - A_2 B_1 = \det \begin{bmatrix} A_1 & B_1 \\ A_2 & B_2 \end{bmatrix}, \tag{1.25}$$

a relation that should be familiar from elementary geometry. Equally familiar should be the following expression for the 3-dimensional volume of a parallelepiped with legs

A, **B**, and **C** [which follows from the components (1.21) of the Levi-Civita tensor]:

$$\text{3-volume} = \boldsymbol{\epsilon}(\mathbf{A}, \mathbf{B}, \mathbf{C}) = \epsilon_{ijk} A_i B_j C_k = \mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = \det \begin{bmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \\ A_3 & B_3 & C_3 \end{bmatrix}. \quad (1.26)$$

3-volume

Our formal definition (1.20) of volume is justified because it gives rise to these familiar equations.

Equations (1.25) and (1.26) are foundations from which one can derive the usual formulas $dA = dx\, dy$ and $dV = dx\, dy\, dz$ for the area and volume of elementary surface and volume elements with Cartesian side lengths dx, dy, and dz (Ex. 1.10).

In Euclidean 3-space, we define the vectorial surface area of a 2-dimensional parallelogram with legs **A** and **B** to be

$$\boxed{\boldsymbol{\Sigma} = \mathbf{A} \times \mathbf{B} = \boldsymbol{\epsilon}(\_\_, \mathbf{A}, \mathbf{B}).} \quad (1.27)$$

vectorial surface area

This vectorial surface area has a magnitude equal to the area of the parallelogram and a direction perpendicular to it. Notice that this surface area $\boldsymbol{\epsilon}(\_\_, \mathbf{A}, \mathbf{B})$ can be thought of as an object that is waiting for us to insert a third leg, **C**, so as to compute a 3-volume $\boldsymbol{\epsilon}(\mathbf{C}, \mathbf{A}, \mathbf{B})$—the volume of the parallelepiped with legs **C**, **A**, and **B**.

A parallelogram's surface has two faces (two sides), called the *positive face* and the *negative face*. If the vector **C** sticks out of the positive face, then $\boldsymbol{\Sigma}(\mathbf{C}) = \boldsymbol{\epsilon}(\mathbf{C}, \mathbf{A}, \mathbf{B})$ is positive; if **C** sticks out of the negative face, then $\boldsymbol{\Sigma}(\mathbf{C})$ is negative.

1.8.1 Gauss's and Stokes' Theorems

1.8.1

Gauss's and Stokes' theorems

Such vectorial surface areas are the foundation for surface integrals in 3-dimensional space and for the familiar *Gauss's theorem,*

$$\boxed{\int_{\mathcal{V}_3} (\boldsymbol{\nabla} \cdot \mathbf{A}) dV = \int_{\partial \mathcal{V}_3} \mathbf{A} \cdot d\boldsymbol{\Sigma}} \quad (1.28a)$$

(where $\mathcal{V}_3$ is a compact 3-dimensional region, and $\partial \mathcal{V}_3$ is its closed 2-dimensional boundary) and *Stokes' theorem,*

$$\boxed{\int_{\mathcal{V}_2} \boldsymbol{\nabla} \times \mathbf{A} \cdot d\boldsymbol{\Sigma} = \int_{\partial \mathcal{V}_2} \mathbf{A} \cdot d\mathbf{l}} \quad (1.28b)$$

(where $\mathcal{V}_2$ is a compact 2-dimensional region, $\partial \mathcal{V}_2$ is the 1-dimensional closed curve that bounds it, and the last integral is a line integral around that curve); see, e.g., Arfken, Weber, and Harris (2013).

This mathematics is illustrated by the integral and differential conservation laws for electric charge and for particles: The total charge and the total number of particles inside a 3-dimensional region of space $\mathcal{V}_3$ are $\int_{\mathcal{V}_3} \rho_e\, dV$ and $\int_{\mathcal{V}_3} n dV$, where ρ_e is the charge density and n the number density of particles. The rates that charge and particles flow out of $\mathcal{V}_3$ are the integrals of the current density **j** and the particle flux

vector $\mathbf{S}$ over its boundary $\partial\mathcal{V}_3$. Therefore, the *integral laws of charge conservation and particle conservation* are

integral conservation laws

$$\boxed{\frac{d}{dt}\int_{\mathcal{V}_3}\rho_e\,dV+\int_{\partial\mathcal{V}_3}\mathbf{j}\cdot d\mathbf{\Sigma}=0,}\qquad\boxed{\frac{d}{dt}\int_{\mathcal{V}_3}n dV+\int_{\partial\mathcal{V}_3}\mathbf{S}\cdot d\mathbf{\Sigma}=0.}\tag{1.29}$$

Pull the time derivative inside each volume integral (where it becomes a partial derivative), and apply Gauss's law to each surface integral; the results are $\int_{\mathcal{V}_3}(\partial\rho_e/\partial t+\mathbf{\nabla}\cdot\mathbf{j})dV=0$ and similarly for particles. The only way these equations can be true for all choices of $\mathcal{V}_3$ is for the integrands to vanish:

differential conservation laws

$$\boxed{\partial\rho_e/\partial t+\mathbf{\nabla}\cdot\mathbf{j}=0,}\qquad\boxed{\partial n/\partial t+\mathbf{\nabla}\cdot\mathbf{S}=0.}\tag{1.30}$$

These are the *differential conservation laws for charge and for particles.* They have a standard, universal form: the time derivative of the density of a quantity plus the divergence of its flux vanishes.

Note that the integral conservation laws (1.29) and the differential conservation laws (1.30) require no coordinate system or basis for their description, and no coordinate system or basis was used in deriving the differential laws from the integral laws. This is an example of the fundamental principle that *the Newtonian physical laws are all expressible as geometric relationships among geometric objects.*

EXERCISES

Exercise 1.10 *Derivation and Practice: Volume Elements in Cartesian Coordinates*
Use Eqs. (1.25) and (1.26) to derive the usual formulas $dA=dxdy$ and $dV=dxdydz$ for the 2-dimensional and 3-dimensional integration elements, respectively, in right-handed Cartesian coordinates. [Hint: Use as the edges of the integration volumes $dx\,\mathbf{e}_x$, $dy\,\mathbf{e}_y$, and $dz\,\mathbf{e}_z$.]

Exercise 1.11 *Example and Practice: Integral of a Vector Field over a Sphere*
Integrate the vector field $\mathbf{A}=z\mathbf{e}_z$ over a sphere with radius a, centered at the origin of the Cartesian coordinate system (i.e., compute $\int\mathbf{A}\cdot d\mathbf{\Sigma}$). Hints:

(a) Introduce spherical polar coordinates on the sphere, and construct the vectorial integration element $d\mathbf{\Sigma}$ from the two legs $ad\theta\,\mathbf{e}_{\hat\theta}$ and $a\sin\theta d\phi\,\mathbf{e}_{\hat\phi}$. Here $\mathbf{e}_{\hat\theta}$ and $\mathbf{e}_{\hat\phi}$ are unit-length vectors along the θ and ϕ directions. (Here as in Sec. 1.6 and throughout this book, we use accents on indices to indicate which basis the index is associated with: hats here for the spherical orthonormal basis, bars in Sec. 1.6 for the barred Cartesian basis.) Explain the factors $ad\theta$ and $a\sin\theta d\phi$ in the definitions of the legs. Show that

$$d\mathbf{\Sigma}=\boldsymbol{\epsilon}(\_\_,\mathbf{e}_{\hat\theta},\mathbf{e}_{\hat\phi})a^2\sin\theta d\theta d\phi.\tag{1.31}$$

(b) Using $z=a\cos\theta$ and $\mathbf{e}_z=\cos\theta\mathbf{e}_{\hat r}-\sin\theta\mathbf{e}_{\hat\theta}$ on the sphere (where $\mathbf{e}_{\hat r}$ is the unit vector pointing in the radial direction), show that

$$\mathbf{A}\cdot d\mathbf{\Sigma}=a\cos^2\theta\,\boldsymbol{\epsilon}(\mathbf{e}_{\hat r},\mathbf{e}_{\hat\theta},\mathbf{e}_{\hat\phi})\,a^2\sin\theta d\theta d\phi.$$

(c) Explain why $\boldsymbol{\epsilon}(\mathbf{e}_{\hat{r}}, \mathbf{e}_{\hat{\theta}}, \mathbf{e}_{\hat{\phi}}) = 1$.

(d) Perform the integral $\int \mathbf{A} \cdot d\boldsymbol{\Sigma}$ over the sphere's surface to obtain your final answer $(4\pi/3)a^3$. This, of course, is the volume of the sphere. Explain pictorially why this had to be the answer.

Exercise 1.12 *Example: Faraday's Law of Induction*

One of Maxwell's equations says that $\boldsymbol{\nabla} \times \mathbf{E} = -\partial \mathbf{B}/\partial t$ (in SI units), where $\mathbf{E}$ and $\mathbf{B}$ are the electric and magnetic fields. This is a geometric relationship between geometric objects; it requires no coordinates or basis for its statement. By integrating this equation over a 2-dimensional surface $\mathcal{V}_2$ with boundary curve $\partial\mathcal{V}_2$ and applying Stokes' theorem, derive Faraday's law of induction—again, a geometric relationship between geometric objects.

1.9 The Stress Tensor and Momentum Conservation

Press your hands together in the y-z plane and feel the force that one hand exerts on the other across a tiny area A—say, one square millimeter of your hands' palms (Fig. 1.5). That force, of course, is a vector $\mathbf{F}$. It has a normal component (along the x direction). It also has a tangential component: if you try to slide your hands past each other, you feel a component of force along their surface, a "shear" force in the y and z directions. Not only is the force $\mathbf{F}$ vectorial; so is the 2-surface across which it acts, $\boldsymbol{\Sigma} = A\,\mathbf{e}_x$. (Here $\mathbf{e}_x$ is the unit vector orthogonal to the tiny area A, and we have chosen the negative side of the surface to be the $-x$ side and the positive side to be $+x$. With this choice, the force $\mathbf{F}$ is that which the negative hand, on the $-x$ side, exerts on the positive hand.)

force vector

Now, it should be obvious that the force $\mathbf{F}$ is a linear function of our chosen surface $\boldsymbol{\Sigma}$. Therefore, there must be a tensor, the *stress tensor*, that reports the force to us when we insert the surface into its second slot:

stress tensor

$$\boxed{\mathbf{F}(\_\_) = \mathbf{T}(\_\_, \boldsymbol{\Sigma}), \quad \text{or} \quad F_i = T_{ij}\Sigma_j.} \tag{1.32}$$

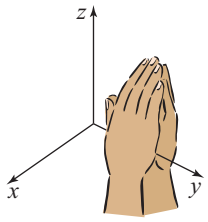

FIGURE 1.5 Hands, pressed together, exert a force on each other.

Newton's law of action and reaction tells us that the force that the positive hand exerts on the negative hand must be equal and opposite to that which the negative hand exerts on the positive. This shows up trivially in Eq. (1.32): by changing the sign of $\mathbf{\Sigma}$, one reverses which hand is regarded as negative and which positive, and since $\mathbf{T}$ is linear in $\mathbf{\Sigma}$, one also reverses the sign of the force.

The definition (1.32) of the stress tensor gives rise to the following physical meaning of its components:

meaning of components of stress tensor

$$T_{jk} = \begin{pmatrix} j \text{ component of force per unit area} \\ \text{across a surface perpendicular to } \mathbf{e}_k \end{pmatrix} = \begin{pmatrix} j \text{ component of momentum that crosses a unit} \\ \text{area that is perpendicular to } \mathbf{e}_k, \text{ per unit time,} \\ \text{with the crossing being from } -x_k \text{ to } +x_k \end{pmatrix}. \tag{1.33}$$

The stresses inside a table with a heavy weight on it are described by the stress tensor $\mathbf{T}$, as are the stresses in a flowing fluid or plasma, in the electromagnetic field, and in any other physical medium. Accordingly, we shall use the stress tensor as an important mathematical tool in our study of force balance in kinetic theory (Chap. 3), elasticity (Part IV), fluid dynamics (Part V), and plasma physics (Part VI).

symmetry of stress tensor

It is not obvious from its definition, but the stress tensor $\mathbf{T}$ is always symmetric in its two slots. To see this, consider a small cube with side L in any medium (or field) (Fig. 1.6). The medium outside the cube exerts forces, and hence also torques, on the cube's faces. The z-component of the torque is produced by the shear forces on the front and back faces and on the left and right. As shown in the figure, the shear forces on the front and back faces have magnitudes $T_{xy}L^2$ and point in opposite directions, so they exert identical torques on the cube, $N_z = T_{xy}L^2(L/2)$ (where $L/2$ is the distance of each face from the cube's center). Similarly, the shear forces on the left and right faces have magnitudes $T_{yx}L^2$ and point in opposite directions, thereby exerting identical torques on the cube, $N_z = -T_{yx}L^2(L/2)$. Adding the torques from all four faces and equating them to the rate of change of angular momentum, $\frac{1}{6}\rho L^5 d\Omega_z/dt$ (where ρ is the mass density, $\frac{1}{6}\rho L^5$ is the cube's moment of inertia, and Ω_z is the z component of its angular velocity), we obtain $(T_{xy} - T_{yx})L^3 = \frac{1}{6}\rho L^5 d\Omega_z/dt$. Now, let the cube's edge length become arbitrarily small, $L \to 0$. If $T_{xy} - T_{yx}$ does not vanish, then the cube will be set into rotation with an infinitely large angular acceleration, $d\Omega_z/dt \propto 1/L^2 \to \infty$—an obviously unphysical behavior. Therefore, $T_{yx} = T_{xy}$, and similarly for all other components: *the stress tensor is always symmetric under interchange of its two slots.*

1.9.1 Examples: Electromagnetic Field and Perfect Fluid

Two examples will make the concept of the stress tensor more concrete.

- **Electromagnetic field:** See Ex. 1.14.
- **Perfect fluid:** A *perfect fluid* is a medium that can exert an isotropic pressure P but no shear stresses, so the only nonzero components of its stress tensor

perfect fluid

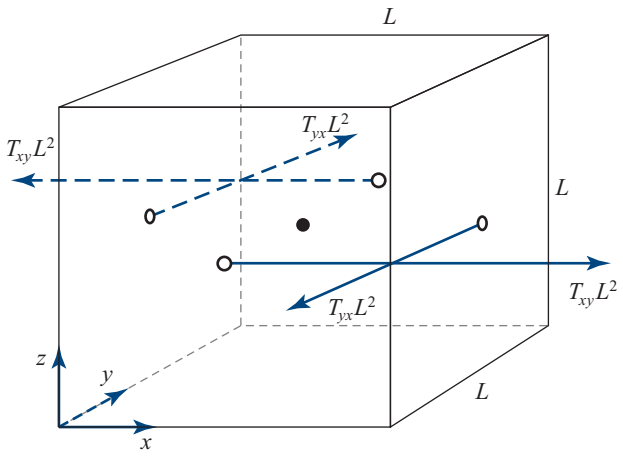

FIGURE 1.6 The shear forces exerted on the left, right, front, and back faces of a vanishingly small cube of side length L. The resulting torque about the z direction will set the cube into rotation with an arbitrarily large angular acceleration unless the stress tensor is symmetric.

in a Cartesian basis are $T_{xx} = T_{yy} = T_{zz} = P$. (Examples of nearly perfect fluids are air and water, but not molasses.) We can summarize this property by $T_{ij} = P\delta_{ij}$ or equivalently, since δ_{ij} are the components of the Euclidean metric, $T_{ij} = Pg_{ij}$. The frame-independent version of this is

$$\mathbf{T} = P\mathbf{g} \quad \text{or, in slot-naming index notation,} \quad T_{ij} = Pg_{ij}. \tag{1.34}$$

Note that, as always, the formula in slot-naming index notation looks identical to the formula $T_{ij} = Pg_{ij}$ for the components in our chosen Cartesian coordinate system. To check Eq. (1.34), consider a 2-surface $\mathbf{\Sigma} = A\mathbf{n}$ with area A oriented perpendicular to some arbitrary unit vector $\mathbf{n}$. The vectorial force that the fluid exerts across $\mathbf{\Sigma}$ is, in index notation, $F_j = T_{jk}\Sigma_k = Pg_{jk}An_k = PAn_j$ (i.e., it is a normal force with magnitude equal to the fluid pressure P times the surface area A). This is what it should be.

1.9.2 Conservation of Momentum 1.9.2

The stress tensor plays a central role in the Newtonian law of momentum conservation because (by definition) the force acting across a surface is the same as the rate of flow of momentum, per unit area, across the surface: *the stress tensor is the flux of momentum.*

Consider the 3-dimensional region of space $\mathcal{V}_3$ used above in formulating the integral laws of charge and particle conservation (1.29). The total momentum in $\mathcal{V}_3$ is $\int_{\mathcal{V}_3} \mathbf{G}\,dV$, where $\mathbf{G}$ is the momentum density. This quantity changes as a result of momentum flowing into and out of $\mathcal{V}_3$. The net rate at which momentum flows outward is the integral of the stress tensor over the surface $\partial\mathcal{V}_3$ of $\mathcal{V}_3$. Therefore, by

analogy with charge and particle conservation (1.29), *the integral law of momentum conservation* says

integral conservation of momentum

$$\boxed{\frac{d}{dt}\int_{\mathcal{V}_3}\mathbf{G}\,dV+\int_{\partial\mathcal{V}_3}\mathbf{T}\cdot d\mathbf{\Sigma}=0.} \tag{1.35}$$

By pulling the time derivative inside the volume integral (where it becomes a partial derivative) and applying the vectorial version of Gauss's law to the surface integral, we obtain $\int_{\mathcal{V}_3}(\partial\mathbf{G}/\partial t+\mathbf{\nabla}\cdot\mathbf{T})\,dV=0$. This can be true for all choices of $\mathcal{V}_3$ only if the integrand vanishes:

differential conservation of momentum

$$\boxed{\frac{\partial\mathbf{G}}{\partial t}+\mathbf{\nabla}\cdot\mathbf{T}=0,\quad\text{or}\quad\frac{\partial G_j}{\partial t}+T_{jk;k}=0.} \tag{1.36}$$

(Because $\mathbf{T}$ is symmetric, it does not matter which of its slots the divergence acts on.) This is *the differential law of momentum conservation*. It has the standard form for any local conservation law: the time derivative of the density of some quantity (here momentum), plus the divergence of the flux of that quantity (here the momentum flux is the stress tensor), is zero. We shall make extensive use of this Newtonian law of momentum conservation in Part IV (elasticity), Part V (fluid dynamics), and Part VI (plasma physics).

EXERCISES

Exercise 1.13 ***Example: Equations of Motion for a Perfect Fluid*

(a) Consider a perfect fluid with density ρ, pressure P, and velocity $\mathbf{v}$ that vary in time and space. Explain why the fluid's momentum density is $\mathbf{G}=\rho\mathbf{v}$, and explain why its momentum flux (stress tensor) is

$$\boxed{\mathbf{T}=P\mathbf{g}+\rho\mathbf{v}\otimes\mathbf{v},\quad\text{or, in slot-naming index notation,}\quad T_{ij}=Pg_{ij}+\rho v_iv_j.} \tag{1.37a}$$

(b) Explain why the law of mass conservation for this fluid is

$$\frac{\partial\rho}{\partial t}+\mathbf{\nabla}\cdot(\rho\mathbf{v})=0. \tag{1.37b}$$

(c) Explain why the derivative operator

$$\frac{d}{dt}\equiv\frac{\partial}{\partial t}+\mathbf{v}\cdot\mathbf{\nabla} \tag{1.37c}$$

describes the rate of change as measured by somebody who moves locally with the fluid (i.e., with velocity $\mathbf{v}$). This is sometimes called the fluid's *advective time derivative* or *convective time derivative* or *material derivative*.

(d) Show that the fluid's law of mass conservation (1.37b) can be rewritten as

$$\frac{1}{\rho}\frac{d\rho}{dt} = -\boldsymbol{\nabla}\cdot\mathbf{v}, \tag{1.37d}$$

which says that the divergence of the fluid's velocity field is minus the fractional rate of change of its density, as measured in the fluid's local rest frame.

(e) Show that the differential law of momentum conservation (1.36) for the fluid can be written as

$$\frac{d\mathbf{v}}{dt} = -\frac{\boldsymbol{\nabla}P}{\rho}. \tag{1.37e}$$

This is called the fluid's *Euler equation*. Explain why this Euler equation is Newton's second law of motion, $\mathbf{F} = m\mathbf{a}$, written on a per unit mass basis.

In Part V of this book, we use Eqs. (1.37) to study the dynamical behaviors of fluids. For many applications, the Euler equation will need to be augmented by the force per unit mass exerted by the fluid's internal viscosity.

Exercise 1.14 ***Problem: Electromagnetic Stress Tensor*

(a) An electric field $\mathbf{E}$ exerts (in SI units) a pressure $\epsilon_o\mathbf{E}^2/2$ orthogonal to itself and a tension of this same magnitude along itself. Similarly, a magnetic field $\mathbf{B}$ exerts a pressure $\mathbf{B}^2/2\mu_o = \epsilon_o c^2\mathbf{B}^2/2$ orthogonal to itself and a tension of this same magnitude along itself. Verify that the following stress tensor embodies these stresses:

$$\boxed{\mathbf{T} = \frac{\epsilon_o}{2}\Big[(\mathbf{E}^2 + c^2\mathbf{B}^2)\mathbf{g} - 2(\mathbf{E}\otimes\mathbf{E} + c^2\mathbf{B}\otimes\mathbf{B})\Big].} \tag{1.38}$$

(b) Consider an electromagnetic field interacting with a material that has a charge density ρ_e and a current density $\mathbf{j}$. Compute the divergence of the electromagnetic stress tensor (1.38) and evaluate the derivatives using Maxwell's equations. Show that the result is the negative of the force density that the electromagnetic field exerts on the material. Use momentum conservation to explain why this has to be so.

1.10 Geometrized Units and Relativistic Particles for Newtonian Readers 1.10

Readers who are skipping the relativistic parts of this book will need to know two important pieces of relativity: (i) geometrized units and (ii) the relativistic energy and momentum of a moving particle.

1.10.1 Geometrized Units 1.10.1

The speed of light is independent of one's reference frame (i.e., independent of how fast one moves). This is a fundamental tenet of special relativity, and in the era before 1983, when the meter and the second were defined independently, it was tested and

confirmed experimentally with very high precision. By 1983, this constancy had become so universally accepted that it was used to redefine the meter (which is hard to measure precisely) in terms of the second (which is much easier to measure with modern technology).[11] The meter is now related to the second in such a way that the speed of light is precisely $c = 299{,}792{,}458\ \mathrm{m\,s^{-1}}$ (i.e., 1 meter is the distance traveled by light in 1/299,792,458 seconds). Because of this constancy of the light speed, it is permissible when studying special relativity to set c to unity. Doing so is equivalent to the relationship

$$c = 2.99792458 \times 10^8\ \mathrm{m\,s^{-1}} = 1 \tag{1.39a}$$

between seconds and centimeters; i.e., equivalent to

$$1\ \mathrm{s} = 2.99792458 \times 10^8\ \mathrm{m}. \tag{1.39b}$$

geometrized units

We refer to units in which $c = 1$ as *geometrized units,* and we adopt them throughout this book when dealing with relativistic physics, since they make equations look much simpler. Occasionally it will be useful to restore the factors of c to an equation, thereby converting it to ordinary (SI or cgs) units. This restoration is achieved easily using dimensional considerations. For example, the equivalence of mass m and relativistic energy $\mathcal{E}$ is written in geometrized units as $\mathcal{E} = m$. In SI units $\mathcal{E}$ has dimensions of joule = kg m^2 s^{-2}, while m has dimensions of kg, so to make $\mathcal{E} = m$ dimensionally correct we must multiply the right side by a power of c that has dimensions m^2 s^{-2} (i.e., by c^2); thereby we obtain $\mathcal{E} = mc^2$.

1.10.2

1.10.2 Energy and Momentum of a Moving Particle

relativistic energy and momentum

A particle with rest mass m, moving with velocity $\mathbf{v} = d\mathbf{x}/dt$ and speed $v = |\mathbf{v}|$, has a relativistic energy $\mathcal{E}$ (including its rest mass), relativistic kinetic energy E (excluding its rest mass), and relativistic momentum $\mathbf{p}$ given by

$$\boxed{\mathcal{E} = \frac{m}{\sqrt{1-v^2}} \equiv \frac{m}{\sqrt{1-v^2/c^2}} \equiv E + m,} \qquad \boxed{\mathbf{p} = \mathcal{E}\mathbf{v} = \frac{m\mathbf{v}}{\sqrt{1-v^2}};} \tag{1.40}$$

$$\text{so } \boxed{\mathcal{E} = \sqrt{m^2 + \mathbf{p}^2}.}$$

In the low-velocity (Newtonian) limit, the energy E with rest mass removed (kinetic energy) and the momentum $\mathbf{p}$ take their familiar Newtonian forms:

$$\text{When } v \ll c \equiv 1, \quad E \to \frac{1}{2}mv^2 \quad \text{and } \mathbf{p} \to m\mathbf{v}. \tag{1.41}$$

11. The second is defined as the duration of 9,192,631,770 periods of the radiation produced by a certain hyperfine transition in the ground state of a ^{133}Cs atom that is at rest in empty space. Today (2016) all fundamental physical units except mass units (e.g., the kilogram) are defined similarly in terms of fundamental constants of Nature.

A particle with zero rest mass (a photon or a graviton)[12] always moves with the speed of light $v = c = 1$, and like other particles it has momentum $\mathbf{p} = \mathcal{E}\mathbf{v}$, so the magnitude of its momentum is equal to its energy: $|\mathbf{p}| = \mathcal{E}v = \mathcal{E}c = \mathcal{E}$.

When particles interact (e.g., in chemical reactions, nuclear reactions, and elementary-particle collisions) the sum of the particle energies $\mathcal{E}$ is conserved, as is the sum of the particle momenta $\mathbf{p}$.

For further details and explanations, see Chap. 2.

EXERCISES

Exercise 1.15 *Practice: Geometrized Units*

Convert the following equations from the geometrized units in which they are written to SI units.

(a) The "Planck time" t_P expressed in terms of Newton's gravitation constant G and Planck's reduced constant $\hbar$, $t_P = \sqrt{G\hbar}$. What is the numerical value of t_P in seconds? in meters?

(b) The energy $\mathcal{E} = 2m$ obtained from the annihilation of an electron and a positron, each with rest mass m.

(c) The Lorentz force law $m d\mathbf{v}/dt = e(\mathbf{E} + \mathbf{v} \times \mathbf{B})$.

(d) The expression $\mathbf{p} = \hbar\omega\mathbf{n}$ for the momentum $\mathbf{p}$ of a photon in terms of its angular frequency ω and direction $\mathbf{n}$ of propagation.

How tall are you, in seconds? How old are you, in meters?

Bibliographic Note

Most of the concepts developed in this chapter are treated, though from rather different viewpoints, in intermediate and advanced textbooks on classical mechanics or electrodynamics, such as Marion and Thornton (1995); Jackson (1999); Griffiths (1999); Goldstein, Poole, and Safko (2002).

Landau and Lifshitz's (1976) advanced text *Mechanics* is famous for its concise and precise formulations; it lays heavy emphasis on symmetry principles and their implications. A similar approach is followed in the next volume in their Course of Theoretical Physics series, *The Classical Theory of Fields* (Landau and Lifshitz, 1975), which is rooted in special relativity and goes on to cover general relativity. We refer to other volumes in this remarkable series in subsequent chapters.

The three-volume *Feynman Lectures on Physics* (Feynman, Leighton, and Sands, 2013) had a big influence on several generations of physicists, and even more so on their teachers. Both of us (Blandford and Thorne) are immensely indebted to Richard Feynman for shaping our own approaches to physics. His insights on the foundations

12. We do not know for sure that photons and gravitons are massless, but the laws of physics as currently understood require them to be massless, and there are tight experimental limits on their rest masses.

of classical physics and its relationship to quantum mechanics, and on calculational techniques, are as relevant today as in 1963, when his course was first delivered.

The geometric viewpoint on the laws of physics, which we present and advocate in this chapter, is not common (but it should be because of its great power). For example, the vast majority of mechanics and electrodynamics textbooks, including all those listed above, define a tensor as a matrix-like entity whose components transform under rotations in the manner described by Eq. (1.13b). This is a complicated definition that hides the great simplicity of a tensor as nothing more than a linear function of vectors; it obscures thinking about tensors geometrically, without the aid of any coordinate system or basis.

The geometric viewpoint comes to the physics community from mathematicians, largely by way of relativity theory. By now, most relativity textbooks espouse it. See the Bibliographic Note to Chap. 2. Fortunately, this viewpoint is gradually seeping into the nonrelativistic physics curriculum (e.g., Kleppner and Kolenkow, 2013). We hope this chapter will accelerate that seepage.

NAME INDEX

Page numbers for entries in boxes are followed by "b," those for epigraphs at the beginning of a chapter by "e," those for figures by "f," and those for notes by "n."

SUBJECT INDEX

Second and third level entries are not ordered alphabetically. Instead, the most important or general entries come first, followed by less important or less general ones, with specific applications last.

Page numbers for entries in boxes are followed by "b," those for epigraphs at the beginning of a chapter by "e," those for figures by "f," for notes by "n," and for tables by "t."