CHAPTER 1

GEOMETRIC SETTING

Physicists recognize two branches of mechanics: classical and quantum. For the past century, the quantum view, emphasizing the corpuscular nature of matter at atomic and finer scales, has played a dominant role in most universities' physics curricula. In those settings, classical mechanics enjoys a distinguished mathematical pedigree, being based on ideas developed by Galileo Galilei, Johannes Kepler, Isaac Newton, Leonhard Euler, Joseph Louis Lagrange, William Rowan Hamilton, and others. Nevertheless, many academic physics departments treat classical mechanics as a mathematical training ground for undergraduates preparing to study the principles of subatomic particles and quantum fields, subjects commonly regarded as more fundamental.

Applied mathematicians and engineers tend to view classical mechanics from a different perspective. While the most elementary formulations of Newton's laws and the Lagrangian and Hamiltonian formalisms focus on idealized particles with mass, many natural phenomena appear to macroscopic observers—those whose scales of observation are significantly larger than 10^{-10} meters—as continuous in space and

Continuum Mechanics: The Birthplace of Mathematical Models, First Edition. M.B. Allen.

time. For these phenomena, fruitful mathematical descriptions typically arise from extensions of classical mechanics pioneered, most notably, by Leonhard Euler and Augustin-Louis Cauchy and refined during the last half of the twentieth century by a large community of scientists, some of whom are mentioned in the preface. In these extensions, matter appears to be continuous, in a sense to be made more precise in the next chapter.

Continuum mechanics embodies these extensions, furnishing useful mathematical models of fluids, elastic solids, and viscoelastic materials. These models describe phenomena that we see and feel in our everyday interactions with the world: rocks in the Earth's crust, water on and beneath its surface, weather, the structures that humans build, and the biological tissues that we occupy. The models typically take the form of partial differential equations describing rates of change with respect to spatial position and time. Advances in our ability to understand and solve these types of equations—especially using high-performance computers—have made continuum mechanics one of the most powerful tools in applied mathematics and engineering. For this reason, in developing the elements of the subject, this book frequently draws connections between its core concepts and the qualitative theory of partial differential equations.

1.1 VECTORS AND EUCLIDEAN POINT SPACE

From a mathematical perspective, continuum mechanics has roots in geometry. In the most natural geometric setting, basic principles do not depend on any observer's particular frame of reference or choice of coordinate systems. One aim of this book is to develop the rudiments of continuum mechanics in a manner that minimizes reliance on particular coordinates, recognizing that using these concepts in specific problems often requires the adoption of a well-chosen coordinate system.

The geometric setting here is relatively simple, relying on ideas familiar to anyone who has studied multivariable calculus and linear algebra. For a more sophisticated approach, refer to [38].

1.1.1 Vectors

Fundamental to continuum mechanics is the three-dimensional Euclidean vector space over the field $\mathbb{R}$ of real numbers. This space, which we denote as $\mathbb{E}$, has features that do not depend on any system for assigning numbers to the vectors in it. In particular, $\mathbb{E}$ has three attributes beyond those common to all vector spaces.

Although the attributes are elementary, it is useful to review them in coordinate-free language and to show how coordinates arise.

• **Inner product.** $\mathbb{E}$ possesses an **inner product**, that is, a binary operation that maps each pair of vectors $\mathbf{x}, \mathbf{y} \in \mathbb{E}$ onto a real number $\mathbf{x} \cdot \mathbf{y}$ with the following properties:

$\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$	symmetry
$(\mathbf{x} + \mathbf{y}) \cdot \mathbf{z} = \mathbf{x} \cdot \mathbf{z} + \mathbf{y} \cdot \mathbf{z}$	additivity
$(a\,\mathbf{x}) \cdot \mathbf{y} = a(\mathbf{x} \cdot \mathbf{y})$ for all $a \in \mathbb{R}$	homogeneity
$\mathbf{x} \cdot \mathbf{x} \geqslant 0;\ \mathbf{x} \cdot \mathbf{x} = 0$ if and only if $\mathbf{x} = \mathbf{0}$.	positive definiteness

Geometry in $\mathbb{E}$ arises from the inner product. It allows us to associate with each vector in $\mathbb{E}$ a **length**,

$$\|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}},$$

and, when $\mathbf{x} \neq \mathbf{0}$, a **direction** $\mathbf{x}/\|\mathbf{x}\|$. Two nonzero vectors $\mathbf{x}$ and $\mathbf{y}$ have **angle** θ given by

$$\cos\theta = \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|}.$$

Two vectors $\mathbf{x}$ and $\mathbf{y}$ are **orthogonal** if $\mathbf{x} \cdot \mathbf{y} = 0$. For two arbitrary vectors $\mathbf{x}, \mathbf{y} \in \mathbb{E}$ with $\mathbf{x} \neq 0$, the **orthogonal projection** of $\mathbf{y}$ onto $\mathbf{x}$ is

$$\mathrm{Proj}_{\mathbf{x}}(\mathbf{y}) = \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\|^2}\mathbf{x}. \tag{1.1.1}$$

See Figure 1.1.

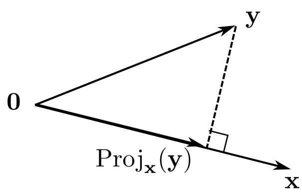

Figure 1.1. The orthogonal projection of $\mathbf{y}$ onto $\mathbf{x}$.

• **Orthonormal basis.** There is an **orthonormal basis** $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ for $\mathbb{E}$, that is, a basis such that $\mathbf{e}_i \cdot \mathbf{e}_i = 1$ for $i = 1, 2, 3$ and $\mathbf{e}_i \cdot \mathbf{e}_j = 0$ when $i \neq j$.

The basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$, shown in Figure 1.2, establishes a Cartesian coordinate system on $\mathbb{E}$. Using this system, we represent any vector $\mathbf{x} \in \mathbb{E}$ as a point in the vector space $\mathbb{R}^3$ of ordered triples of real numbers: if $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3$, then

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

denotes its representation in $\mathbb{R}^3$ with respect to the basis. A different choice of basis vectors for $\mathbb{E}$—even a different choice of orthonormal basis—yields a different representation in $\mathbb{R}^3$, but the vector in $\mathbb{E}$ remains fixed in magnitude and direction. For this reason we distinguish $\mathbb{E}$ from $\mathbb{R}^3$.

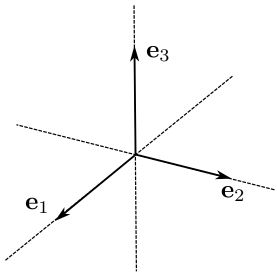

Figure 1.2. Standard orthonormal basis vectors defining a Cartesian coordinate system.

EXERCISE 1 *For a given orthonormal basis* $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$, *it is possible to determine the coefficients* x_i *knowing the vector* $\mathbf{x}$. *Show that* $x_i = \mathbf{x} \cdot \mathbf{e}_i$.

For consistency with the conventions of matrix multiplication, discussed later, we write representations of vectors in $\mathbb{R}^3$ as column arrays. Under this convention, when the basis for $\mathbb{E}$ is understood, we sometimes abuse notation by writing as if the vector equals its representation:

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \qquad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \qquad \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

For typesetting convenience we sometimes denote column vectors as formal transposes of row vectors, for example,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = (x_1, x_2, x_3)^\top.$$

With respect to the orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$, the inner product of two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{E}$ has the value

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^{3} x_i y_i.$$

The following exercise gives a coordinate-free expression for the inner product in terms of lengths.

EXERCISE 2 *Prove the* **polarization identity***:* $\mathbf{x} \cdot \mathbf{y} = \frac{1}{4}(\|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2)$.

• **Cross product.** $\mathbb{E}$ admits a second form of vector multiplication. If $\mathbf{x}, \mathbf{y} \in \mathbb{E}$ have angle θ, then their **cross product** is the vector $\mathbf{x} \times \mathbf{y} \in \mathbb{E}$ that has length $\|\mathbf{x}\|\,\|\mathbf{y}\|\,|\sin\theta|$ and direction orthogonal to both $\mathbf{x}$ and $\mathbf{y}$, with sense given by the right-hand rule, as illustrated in Figure 1.3.

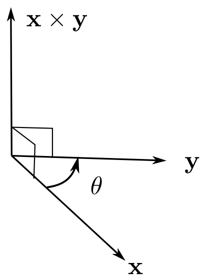

Figure 1.3. The cross product $\mathbf{x} \times \mathbf{y}$, showing the right-hand rule.

As a binary operation on vectors, the cross product is endemic to three space dimensions, as discussed further in Section 3.3. In all that follows, we assume that the orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ for $\mathbb{E}$ has **positive orientation**, meaning that $\mathbf{e}_i \times \mathbf{e}_j = \mathbf{e}_k$ whenever $(i, j, k) = (1, 2, 3)$ or $(2, 3, 1)$ or $(3, 1, 2)$, that is, whenever (i, j, k) is an even permutation of $(1, 2, 3)$. Under this convention, with respect to the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$, the cross product has the value

$$\mathbf{x} \times \mathbf{y} = (x_2 y_3 - x_3 y_2)\mathbf{e}_1 + (x_3 y_1 - x_1 y_3)\mathbf{e}_2 + (x_1 y_2 - x_2 y_1)\mathbf{e}_3. \qquad (1.1.2)$$

Later sections explore additional algebraic and geometric interpretations of the cross product.

Representations with respect to the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ can be useful for calculations, but two caveats are in order. First, they are not the only numerical representations available for vectors $\mathbf{x} \in \mathbb{E}$. Infinitely many orthonormal bases exist for $\mathbb{E}$, and it is possible to construct infinitely many nonorthonormal, non-Cartesian bases. Second, the principles of continuum mechanics do not require *any* choice of basis or associated coordinate system. Nevertheless, this book frequently uses $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ and the Cartesian coordinate system it defines to discuss examples, since this basis furnishes a computationally familiar setting. The remainder of this chapter reviews further aspects of the algebra and geometry of $\mathbb{E}$, with an attempt to minimize unnecessary references to coordinates.

Subsequent chapters refer to vectors that have a variety of physical dimensions. For example, the dimension of position vectors is length, denoted by [L], while velocity vectors have dimension length/time, or [LT^{-1}]. Fastidious readers may anticipate some apparent anomalies associated with algebraic operations involving pairs of vectors having different physical dimensions. Section 1.2 proposes a resolution.

1.1.2 Euclidean Point Space

As fundamental as the Euclidean vector space $\mathbb{E}$ may be to the mathematics of continuum mechanics, the objects of interest—sets of material points defined in Chapter 2—do not reside there. Instead, consistent with experience, these objects occupy points P in a type of space $\mathbb{X}$ that has no intrinsic algebraic structure. To apply the tools of algebra and calculus, we attach the vector space $\mathbb{E}$ to $\mathbb{X}$, choosing a point $O \in \mathbb{X}$ to serve as the origin and assigning to every point $P \in \mathbb{X}$ a vector in $\mathbb{E}$ that translates O into P. This approach allows us to refer to points $P \in \mathbb{X}$ in a way that facilitates the mathematical analysis available for vectors $\mathbf{x} \in \mathbb{E}$. This subsection furnishes details of the association between $\mathbb{X}$ and the space $\mathbb{E}$.

Recognizing the distinction between the set $\mathbb{X}$ of points and the algebraic structure $\mathbb{E}$ may seem pedantic. But without doing so, we cannot correctly account for disparate descriptions made by different observers, who assign vectors to points in different ways. Sections 3.5 and 6.3 examine the effects of such differences.

DEFINITION. *A set* $\mathbb{X}$ *of points is a* **Euclidean point space** *over* $\mathbb{E}$ *if there is a* **translation mapping** $\boldsymbol{d}\colon \mathbb{X} \times \mathbb{X} \to \mathbb{E}$ *having the following properties:*

1. $\boldsymbol{d}(P, P) = \mathbf{0} \in \mathbb{E}$ *for every* $P \in \mathbb{X}$.
2. $\boldsymbol{d}(P, R) = \boldsymbol{d}(P, Q) + \boldsymbol{d}(Q, R)$ *for all* $P, Q, R \in \mathbb{X}$.
3. *For every fixed* $P \in \mathbb{X}$ *the mapping* $\boldsymbol{d}(P, \cdot)$ *is one-to-one and onto. In other words, for every point* $P \in \mathbb{X}$ *and every vector* $\mathbf{t} \in \mathbb{E}$ *there exists a unique point* $Q \in \mathbb{X}$ *such that* $\boldsymbol{d}(P, Q) = \mathbf{t}$.

EXERCISE 3 *From these properties show that* $\boldsymbol{d}(P, Q) = -\boldsymbol{d}(Q, P)$.

We interpret this definition as follows: to each pair (P, Q) of points in $\mathbb{X}$, the mapping $\boldsymbol{d}$ assigns the vector $\mathbf{t} \in \mathbb{E}$ that translates P to Q. With the set $\mathbb{X}$ of points and the vector space $\mathbb{E}$ related in this way, we call $\mathbb{E}$ the **translation space** for $\mathbb{X}$.

Using $\boldsymbol{d}$ it is possible to assign to each point in $\mathbb{X}$ a unique vector $\mathbf{x} \in \mathbb{E}$. To construct such an assignment $\boldsymbol{f}\colon \mathbb{X} \to \mathbb{E}$, first pick a point $O \in \mathbb{X}$ and define $\boldsymbol{f}(O) = \mathbf{0}$. The point O is the **origin** of $\mathbb{X}$ under $\boldsymbol{f}$. Then for any other point $P \in \mathbb{X}$, define $\boldsymbol{f}(P) = \boldsymbol{d}(O, P)$, that is, the vector that translates the origin to P. As shown in Figure 1.4, the translation vector connecting two points $P, Q \in \mathbb{X}$ is

$$\mathbf{t} = \boldsymbol{d}(P, Q) = \boldsymbol{d}(P, O) + \boldsymbol{d}(O, Q) = \boldsymbol{f}(Q) - \boldsymbol{f}(P).$$

In this way, $\boldsymbol{f}$ furnishes a distance function on $\mathbb{X}$:

$$\text{distance}\,(P, Q) = \|\boldsymbol{f}(P) - \boldsymbol{f}(Q)\|.$$

The mapping $\boldsymbol{f}$ also provides access to the algebraic tools available in the Euclidean vector space $\mathbb{E}$. In this way, $\mathbb{E}$ constitutes a space of **position vectors** for $\mathbb{X}$ under $\boldsymbol{f}$.

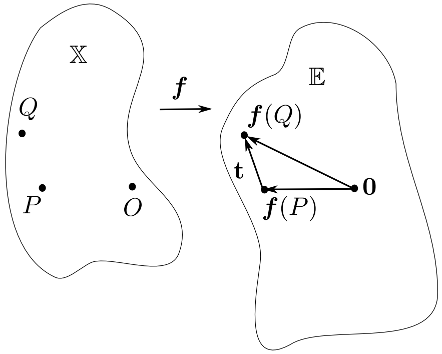

Figure 1.4. Points in $\mathbb{X}$ and vectors in $\mathbb{E}$, showing the translation vector $\mathbf{t} = \boldsymbol{d}(P, Q)$.

Through the mapping $\boldsymbol{f}\colon \mathbb{X} \to \mathbb{E}$, any coordinate system on $\mathbb{E}$ induces a coordinate system on the Euclidean point space $\mathbb{X}$. In particular, the orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ introduced earlier induces a Cartesian coordinate system on $\mathbb{X}$.

The mapping $\boldsymbol{f} = \boldsymbol{d}(O, \cdot)$ is not unique. Infinitely many translation maps $\boldsymbol{d}$ exist, and for each there are infinitely many possible choices of the origin $O \in \mathbb{X}$. In addition, the analysis of continua involves time t as a variable, so we may as well regard the objects of interest as residing in **space-time**, $\mathbb{X} \times \mathbb{T}$, where $\mathbb{T}$ is the time interval of interest. There are infinitely many ways to assign a temporal instant $t_0 \in \mathbb{T}$ to the origin of the time axis $\mathbb{R}$.

DEFINITION. *A* **frame of reference** *is a choice* $(\boldsymbol{f}, t_0)$ *of one-to-one correspondence* $\boldsymbol{f}\colon \mathbb{X} \to \mathbb{E}$*, having the properties listed above, together with a choice* t_0 *of instant to be assigned to the origin of the time axis* $\mathbb{R}$*.*

Because any frame of reference amounts to a choice made by an observer, we cannot treat it as an essential aspect of the physics. For this reason, it will be necessary, eventually, to examine (1) which quantities associated with the physics remain invariant under changes in frame of reference and (2) how noninvariant quantities vary under such changes. We return to these questions in Section 3.5.

1.1.3 Summary

Vectors in this text belong to a three-dimensional Euclidean vector space $\mathbb{E}$. This space possesses an inner product, which gives rise to such geometric concepts as length, angle, and orthogonality. It also possesses a standard orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ and a cross product. The orthonormal basis enables us to represent vectors in $\mathbb{E}$ as elements of $\mathbb{R}^3$, which facilitates calculations in specific problems.

The objects of interest in continuum mechanics reside in a Euclidean point space. The vector space $\mathbb{E}$ serves as a translation space for this space, through a mapping $\boldsymbol{f}$ that assigns to a unique point O the vector $\mathbf{0}$ and assigns to every other point P the vector $\mathbf{x}$ that translates O to P. Any choice of (1) a mapping $\boldsymbol{f}$ and (2) a time instant t_0 to assign to the origin 0 of the time axis $\mathbb{R}$ constitutes a frame of reference, which different observers may choose in different ways.

1.2 TENSORS

The analysis of motion involves linear mappings that act on vectors as geometric objects. Coordinate systems play an important role in calculations, but they are ancillary in the sense that the mappings' geometric actions remain invariant under a change of coordinates. Linear mappings having this property are tensors. In this text we encounter two orders of tensors.

1.2.1 First-Order Tensors and Vectors

Many types of vectors have representations as elements of $\mathbb{R}^3$. Examples include position vectors in $\mathbb{E}$, velocities, forces, and other entities that have both direction and magnitude. Logically, these different types of vectors belong to different vector spaces, all of which are inner-product spaces having the same algebraic structure as $\mathbb{R}^3$.

DEFINITION. *A* **Euclidean 3-space** *is a three-dimensional vector space* $\mathbb{V}$ *for which there is a one-to-one, onto mapping* $\Phi\colon \mathbb{V} \to \mathbb{R}^3$ *such that*

$$\begin{aligned}
&\Phi(\mathbf{x}+\mathbf{y}) = \Phi(\mathbf{x}) + \Phi(\mathbf{y}),\\
&\Phi(s\mathbf{x}) = s\Phi(\mathbf{x}),\\
&\Phi(\mathbf{x})\cdot\Phi(\mathbf{y}) = \mathbf{x}\cdot\mathbf{y},
\end{aligned}$$

for all $\mathbf{x}, \mathbf{y} \in \mathbb{V}$ *and for all* $s \in \mathbb{R}$.

In words, the algebra of $\mathbb{V}$ mirrors that of $\mathbb{R}^3$. $\mathbb{E}$ is such a space. Because this definition admits spaces associated with different physical dimensions, it gives rise to potential anomalies in vector algebra alluded to in Section 1.1. For example, if $\mathbf{v}$ is a velocity vector with dimension [LT^{-1}], computing its component in the direction of a unit vector $\mathbf{n}$ having dimension [L] requires calculating $\mathbf{v}\cdot\mathbf{n}$. Yet because they have different physical dimensions, $\mathbf{v}$ and $\mathbf{n}$ belong to different Euclidean 3-spaces, and in a fussy sense their inner product *per se* appears to make little sense.

In Euclidean spaces, we can resolve the matter by viewing vectors not only as geometric objects but also as mappings acting on other vectors.

DEFINITION. *A* **first-order tensor** *on a Euclidean 3-space* $\mathbb{V}$ *is a linear functional on* $\mathbb{V}$*, that is, a mapping* $\ell\colon \mathbb{V} \to \mathbb{R}$ *that satisfies the following properties:*

$$\begin{aligned}
&\ell(\mathbf{x}+\mathbf{y}) = \ell(\mathbf{x}) + \ell(\mathbf{y}),\\
&\ell(a\mathbf{x}) = a\ell(\mathbf{x}) \quad \text{for any } a \in \mathbb{R}.
\end{aligned}$$

It is possible to represent any first-order tensor on an inner-product space $\mathbb{V}$ as a vector in $\mathbb{R}^3$. In particular, if $\mathbf{x} \in \mathbb{V}$ has representation $x_1\mathbf{w}_1 + x_2\mathbf{w}_2 + x_3\mathbf{w}_3$ with respect to a basis $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ for $\mathbb{V}$, then any linear functional ℓ defined on $\mathbb{V}$ has an action that decomposes as follows:

$$\ell(\mathbf{x}) = \sum_{i=1}^{3} x_i \ell(\mathbf{w}_i) = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{c}\cdot\mathbf{x},$$

where each $c_i = \ell(\mathbf{w}_i)$. In this sense, first-order tensors are vectors.

EXERCISE 4 *Let* $\mathbf{n} \in \mathbb{E}$ *be a fixed vector, and let* $\mathbb{V}$ *be the Euclidean 3-space of velocities. Define* $\ell_{\mathbf{n}}\colon \mathbb{V} \to \mathbb{R}$ *by* $\ell_{\mathbf{n}}(\mathbf{v}) = \mathbf{v}\cdot\mathbf{n}$*. Show that* $\ell_{\mathbf{n}}$ *is a first-order tensor on* $\mathbb{V}$*.*

The apparent anomaly in the expression $\mathbf{v} \cdot \mathbf{n}$ now disappears. The inner product simply furnishes a computational device for representing the linear functional defined by $\ell_{\mathbf{n}}(\mathbf{v}) = \mathbf{v} \cdot \mathbf{n}$. This first-order tensor maps $\mathbf{v}$ onto the quantity that results when we formally compute the inner product $\mathbf{v} \cdot \mathbf{n}$ and assign the appropriate physical dimension, $[\mathrm{LT}^{-1}][\mathrm{L}] = [\mathrm{L}^2\mathrm{T}^{-1}]$. With this observation in mind, we henceforth denote linear functionals on Euclidean 3-spaces as vectors in $\mathbb{E}$.

Representing linear functionals as vectors in $\mathbb{E}$ involves no serious risks of confusion in this book. In contrast, representing vectors as elements of $\mathbb{R}^3$ requires a strong caveat. Any numerical representation of a first-order tensor depends on the choice of basis. When we change to a new basis—for example, by changing Cartesian coordinate systems—we leave the geometric action of the linear functional unchanged. But its numerical representation as an ordered triple in $\mathbb{R}^3$ changes.

EXERCISE 5 *In terms of the orthonormal basis* $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$*, a typical vector* $\mathbf{x} \in \mathbb{E}$ *has an expansion of the form* $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3$. *Change to a coordinate system having the same origin but a new orthonormal basis* $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3\}$*, where* $\hat{\mathbf{e}}_1 = (\mathbf{e}_1 + \mathbf{e}_3)/\sqrt{2}$, $\hat{\mathbf{e}}_2 = \mathbf{e}_2$, $\hat{\mathbf{e}}_3 = (-\mathbf{e}_1 + \mathbf{e}_3)/\sqrt{2}$. *We now have two expansions for a typical vector* $\mathbf{x}$*:*

$$\mathbf{x} = \sum_{j=1}^{3} x_j \mathbf{e}_j = \sum_{j=1}^{3} \hat{x}_j \hat{\mathbf{e}}_j .$$

Find expressions for the coefficients $\hat{x}_j$ *in terms of the coefficients* x_i.

Exercises 6 and 7 establish the transformation rules that representations of vectors in $\mathbb{R}^3$ must obey to keep their lengths and directions—and their actions as first-order tensors—invariant under coordinate changes.

EXERCISE 6 *For a more general change of orthonormal basis from* $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ *to* $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3\}$*, leaving the origin fixed, show that the following relationship holds between the three coefficients* x_j *and the three coefficients* $\hat{x}_i$*:*

$$\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{e}}_1 \cdot \mathbf{e}_1 & \hat{\mathbf{e}}_1 \cdot \mathbf{e}_2 & \hat{\mathbf{e}}_1 \cdot \mathbf{e}_3 \\ \hat{\mathbf{e}}_2 \cdot \mathbf{e}_1 & \hat{\mathbf{e}}_2 \cdot \mathbf{e}_2 & \hat{\mathbf{e}}_2 \cdot \mathbf{e}_3 \\ \hat{\mathbf{e}}_3 \cdot \mathbf{e}_1 & \hat{\mathbf{e}}_3 \cdot \mathbf{e}_2 & \hat{\mathbf{e}}_3 \cdot \mathbf{e}_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} . \qquad (1.2.1)$$

Thus we can account for this type of change in coordinates through matrix multiplication. We call the entries $Q_{ij} = \hat{\mathbf{e}}_i \cdot \mathbf{e}_j$ *of the transformation matrix the* **direction cosines** *of the transformation. Why?*

EXERCISE 7 *Write the matrix multiplication in Equation (1.2.1) as follows:*

$$\hat{x}_m = \sum_{i=1}^{3} Q_{mi} x_i.$$

Show that

$$x_j = \sum_{l=1}^{3} Q_{lj} \hat{x}_l = \sum_{l=1}^{3} Q_{jl}^{\top} \hat{x}_l,$$

where $Q_{jl}^{\top}$ *denotes the* 3×3 *matrix obtained by reflecting the entries of the matrix* Q_{jl} *in Equation (1.2.1) across the diagonal. Thus the expression giving each* x_j *in terms of the coordinates* $\hat{x}_l$ *also takes the form of matrix multiplication by direction cosines.*

1.2.2 Second-Order Tensors

The idea of mappings on Euclidean 3-spaces extends further. Let $\mathbb{V}$ and $\mathbb{W}$ be two such spaces.

DEFINITION. *A* **second-order tensor** *from* $\mathbb{V}$ *to* $\mathbb{W}$ *is a linear transformation* $\mathsf{A}\colon \mathbb{V} \to \mathbb{W}$. *We denote the set of all second-order tensors from* $\mathbb{V}$ *to* $\mathbb{W}$ *as* $\mathrm{L}(\mathbb{V}, \mathbb{W})$ *or, when* $\mathbb{W} = \mathbb{V}$*, as* $\mathrm{L}(\mathbb{V})$.

The second-order tensors of chief interest in this text are those in $\mathrm{L}(\mathbb{E})$, where we focus attention from now on. These entities need not be exotic. For example, the identity tensor I, defined by the action $\mathsf{I}\mathbf{x} = \mathbf{x}$ for all $\mathbf{x} \in \mathbb{E}$, is a second-order tensor. Less trivially, for fixed $\mathbf{x} \in \mathbb{E}$, let $\mathrm{Proj}_{\mathbf{x}}\colon \mathbb{E} \to \mathbb{E}$ be the standard orthogonal projection operator onto $\mathbf{x}$, described in Equation (1.1.1).

EXERCISE 8 *Show that* $\mathrm{Proj}_{\mathbf{x}}$ *is a second-order tensor.*

The following example furnishes many more possibilities:

DEFINITION. *The* **dyadic product** *of two vectors* $\mathbf{a}, \mathbf{b} \in \mathbb{E}$ *is the mapping* $\mathbf{a} \otimes \mathbf{b}\colon \mathbb{E} \to \mathbb{E}$ *defined by*

$$(\mathbf{a} \otimes \mathbf{b})(\mathbf{x}) = \mathbf{a}(\mathbf{b} \cdot \mathbf{x}).$$

EXERCISE 9 *Show that* $\mathbf{a} \otimes \mathbf{b}$ *is a second-order tensor. Find a representation for* $\mathrm{Proj}_{\mathbf{x}}$ *as a dyadic product.*

With respect to a fixed choice of basis for $\mathbb{E}$, any second-order tensor $\mathsf{A} \in \mathrm{L}(\mathbb{E})$ has a numerical representation. If $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ and $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ are the bases for the domain and target space of $\mathsf{A} \in \mathrm{L}(\mathbb{E})$, then the corresponding **matrix representation** for A is the 3×3 array of real numbers in which the entry in row i, column j is the coefficient A_{ij} multiplying $\mathbf{w}_i$ in the expansion

$$\mathsf{A}\mathbf{v}_j = \sum_{i=1}^{3} A_{ij}\mathbf{w}_i \in \mathbb{E}.$$

EXERCISE 10 *With respect to the basis* $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ *for* $\mathbb{E}$ *defined in Section 1.1, show that the matrix representation of* $\mathsf{A} \in \mathrm{L}(\mathbb{E})$ *has entries* $A_{ij} = \mathbf{e}_i \cdot \mathsf{A}\mathbf{e}_j$. *Hence, the matrix representation of* $\mathbf{e}_1 \otimes \mathbf{e}_3$ *is*

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

With respect to this same basis, show that $\mathbf{a} \otimes \mathbf{b}$ *has matrix representation*

$$\begin{bmatrix} a_1b_1 & a_1b_2 & a_1b_3 \\ a_2b_1 & a_2b_2 & a_2b_3 \\ a_3b_1 & a_3b_2 & a_3b_3 \end{bmatrix}.$$

Thus any second-order tensor $\mathsf{A} \in \mathrm{L}(\mathbb{E})$ has an expansion in terms of the elementary dyadic products $\mathbf{e}_i \otimes \mathbf{e}_j$:

$$\mathsf{A} = \sum_{i=1}^{3}\sum_{j=1}^{3} A_{ij}\mathbf{e}_i \otimes \mathbf{e}_j.$$

Sums over coordinate indices appear so often in continuum mechanics that we frequently use a common shorthand for them. Consider the expressions

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^{3} x_i y_i, \qquad \mathsf{A} = \sum_{i=1}^{3}\sum_{j=1}^{3} A_{ij}\mathbf{e}_i \otimes \mathbf{e}_j. \qquad (1.2.2)$$

In these equations the notations $\sum_{i=1}^{3}$ and $\sum_{i=1}^{3}\sum_{j=1}^{3}$ become superfluous if we adopt the **Einstein summation convention**:

1. If a letter index such as i or j appears exactly once in a term, it stands for any of the values $1, 2, 3$.

2. If a letter index appears exactly twice in a term, we treat the index as a dummy index and sum the terms over the index values $1, 2, 3$.

3. If a term requires a letter index to appear more than twice, we agree to write any required summation signs explicitly.

Under this convention, repetition of the dummy indices i (in the term $x_i y_i$) and i, j (in the term $A_{ij}\mathbf{e}_i \otimes \mathbf{e}_j$) implies summation from 1 to 3 over the repeated indices. Thus we write $x_i y_i$ and $A_{ij}\mathbf{e}_i \otimes \mathbf{e}_j$ as shorthand for the more explicit summation notations in Equation 1.2.2.

Matrix representations and the Einstein summation convention yield a sleek way to represent compositions of second-order tensors. If $\mathsf{A}, \mathsf{B} \in \mathrm{L}(\mathbb{E})$, we denote their composition as a product: $(\mathsf{A} \circ \mathsf{B})(\mathbf{x}) = \mathsf{A}(\mathsf{B}(\mathbf{x})) = (\mathsf{AB})\mathbf{x}$. The notation is suggestive.

EXERCISE 11 *Show that* $\mathsf{AB} \in \mathrm{L}(\mathbb{E})$. *Show that*

$$\mathsf{AB} = \sum_{j=1}^{3}\sum_{l=1}^{3}\sum_{k=1}^{3} A_{jk}B_{kl}\mathbf{e}_j \otimes \mathbf{e}_l = A_{jk}B_{kl}\mathbf{e}_j \otimes \mathbf{e}_l$$

Standard notation Summation convention

Thus the matrix representation of the composition AB is the ordinary matrix product of the 3×3 representations for A and B. In shorthand, the (j, l) entry in the matrix representation for AB is $A_{jk}B_{kl}$.

As with first-order tensors, the representation of a second-order tensor as an array of numbers depends on the choice of bases for the Euclidean 3-spaces involved. When we change from one orthonormal basis to another, we leave the geometric action of the linear transformation unchanged but change its numerical representation as a matrix. To see how, consider the vector $\mathbf{y} = \mathsf{A}\mathbf{x}$ resulting from the action of a second-order tensor A on a vector $\mathbf{x}$. The coordinates of $\mathbf{y}$ with respect to the orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ are $y_i = \mathbf{y} \cdot \mathbf{e}_i$. Substituting for $\mathbf{y}$ in terms of a new orthonormal basis $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3\}$ and for $\hat{x}_m$ using the results of Exercise 6 yields (using the summation convention):

$$\begin{aligned} y_i = \hat{y}_l\hat{\mathbf{e}}_l \cdot \mathbf{e}_i &= \hat{A}_{lm}\hat{x}_m\hat{\mathbf{e}}_l \cdot \mathbf{e}_i \\ &= \hat{A}_{lm}\hat{x}_m Q_{li} \\ &= \hat{A}_{lm}Q_{mj}x_j Q_{li} = (Q_{il}^{\top}\hat{A}_{lm}Q_{mj})x_j. \end{aligned} \tag{1.2.3}$$

Here, Q_{il} denotes entries of the direction cosine matrix introduced in Exercise 6. The expression (1.2.3) for $y_i = A_{ij}x_j$ holds for all possible vectors $\mathbf{x}$. It follows that the entries of A_{ij} transform according to the following matrix multiplication:

$$A_{ij} = Q_{il}^{\top}\hat{A}_{lm}Q_{mj}.$$

EXERCISE 12 *Show that* $\hat{A}_{lm} = Q_{li} A_{ij} Q^{\top}_{jm}$.

Caution is in order here. We do not regard the matrix Q_{il} of direction cosines as a representation of a second-order tensor. Second-order tensors act on vectors in $\mathbb{E}$ to produce (possibly new) vectors in $\mathbb{E}$. The matrix Q_{il} acts on *numerical representations* of vectors in $\mathbb{E}$ and tensors in $\mathrm{L}(\mathbb{E})$ to produce (possibly new) representations of these same entities.

The concept of a tensor generalizes to orders higher than one and two. Although higher-order tensors do not figure prominently in the mechanics covered in this book, there is utility in the approach. Recall that a first-order tensor (which we regard as a vector in $\mathbb{E}$) is a linear functional $\ell \colon \mathbb{E} \to \mathbb{R}$. By extension, we regard a second-order tensor $\mathsf{A} \in \mathrm{L}(\mathbb{E})$ as a **bilinear functional** $\mathsf{A} \colon \mathbb{E} \times \mathbb{E} \to \mathbb{R}$, that is, a mapping $\mathsf{A} \colon (\mathbf{a}_1, \mathbf{a}_2) \mapsto \mathbf{a}_1 \cdot \mathsf{A}\mathbf{a}_2$ that is linear in both of the vector arguments $\mathbf{a}_1, \mathbf{a}_2$. To generalize this idea, define an **nth-order tensor** as an n-linear mapping, that is, a function $\mathsf{A} \colon \mathbb{E}^n \to \mathbb{R}$ that is linear in each of its n arguments.

For example, recall that any second-order tensor is a linear combination of dyadic products $\mathbf{a}_1 \otimes \mathbf{a}_2$, and observe that the actions of these products as bilinear functionals take the following form:

$$\mathbf{a}_1 \otimes \mathbf{a}_2 \colon (\mathbf{b}_1, \mathbf{b}_2) \mapsto \mathbf{b}_1 \cdot (\mathbf{a}_1 \otimes \mathbf{a}_2)\mathbf{b}_2 = (\mathbf{a}_1 \cdot \mathbf{b}_1)(\mathbf{a}_2 \cdot \mathbf{b}_2).$$

More generally, the n-linear functional $\mathbf{a}_1 \otimes \mathbf{a}_2 \otimes \cdots \otimes \mathbf{a}_n$ acts on an n-tuple $(\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n) \in \mathbb{E}^n$ as follows:

$$(\mathbf{a}_1 \otimes \mathbf{a}_2 \otimes \cdots \otimes \mathbf{a}_n)(\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n) = \prod_{i=1}^{n} \mathbf{a}_i \cdot \mathbf{b}_i.$$

As in the case $n = 2$, any nth-order tensor is a linear combination of products of this form.

This view of second-order tensors as bilinear functionals gives rise to two additional concepts from elementary linear algebra.

DEFINITION. *If* $\mathsf{A} \in \mathrm{L}(\mathbb{E})$, *then its* **transpose** *is the second-order tensor* $\mathsf{A}^{\top} \in \mathrm{L}(\mathbb{E})$ *defined by the equation* $\mathbf{a} \cdot \mathsf{A}^{\top}\mathbf{b} = \mathsf{A}\mathbf{a} \cdot \mathbf{b}$ *for all* $\mathbf{a}, \mathbf{b} \in \mathbb{E}$.

One easily verifies that $\mathbf{b} \otimes \mathbf{a} = (\mathbf{a} \otimes \mathbf{b})^{\top}$. Corresponding to this observation is the fact that one obtains the matrix representation of $\mathsf{A}^{\top}$ with respect to any orthonormal basis simply by reflecting the matrix for A across its main diagonal: $A^{\top}_{ij} = A_{ji}$.

DEFINITION. $\mathsf{A} \in \mathrm{L}(\mathbb{E})$ *is* **invertible** *if there exists a tensor* $\mathsf{A}^{-1} \in \mathrm{L}(\mathbb{E})$ *such that* $\mathsf{A}\mathsf{A}^{-1} = \mathsf{A}^{-1}\mathsf{A} = \mathsf{I}$.

Each of the following conditions is necessary and sufficient for A to be invertible:

1. $\mathsf{A}\colon \mathbb{E} \to \mathbb{E}$ is one-to-one, that is, $\mathsf{A}\mathbf{x} = \mathsf{A}\mathbf{y}$ only if $\mathbf{x} = \mathbf{y}$.

2. $\mathsf{A}\mathbf{x} = \mathbf{0}$ only if $\mathbf{x} = \mathbf{0}$.

EXERCISE 13 *Show that each of these conditions implies the other.*

1.2.3 Cross Products, Triple Products, and Determinants

Three concepts—cross products, scalar triple products, and determinants—connect the algebra of $\mathbb{E}$ and $\mathrm{L}(\mathbb{E})$ to the geometry of volumes. As introduced in Section 1.1, the cross product $\mathbf{x} \times \mathbf{y}$ is orthogonal to the plane spanned by $\mathbf{x}$ and $\mathbf{y}$, as illustrated in Figure 1.3. It has length $\|\mathbf{x}\|\,\|\mathbf{y}\|\,|\sin\theta|$, where θ is the angle between $\mathbf{x}$ and $\mathbf{y}$. Equivalently, the length of $\mathbf{x} \times \mathbf{y}$ is the area of the parallelogram with sides $\mathbf{x}$ and $\mathbf{y}$. Figure 1.5 also shows this relationship.

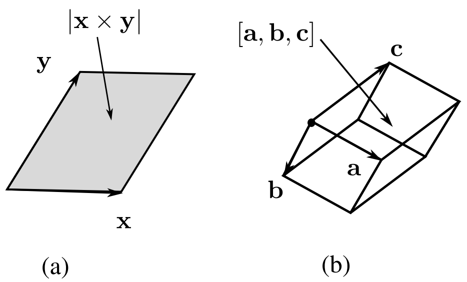

Figure 1.5. (a) Relationship between $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{x} \times \mathbf{y}$. (b) The parallelepiped having adjacent edges $\mathbf{a}, \mathbf{b}, \mathbf{c}$.

Equation (1.1.2) gives the components of the cross product $\mathbf{x} \times \mathbf{y}$ with respect to the orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. Occasionally an indicial form for $\mathbf{x} \times \mathbf{y}$ is useful:

$$\mathbf{x} \times \mathbf{y} = \varepsilon_{ijk} x_j y_k \mathbf{e}_i.$$

Here, ε_{ijk} is the **Levi-Civita symbol**, defined as follows:

$$\varepsilon_{ijk} = \begin{cases} 1, & \text{if} \quad (i,j,k) = (1,2,3), (3,1,2), (2,3,1); \\ -1, & \text{if} \quad (i,j,k) = (1,3,2), (2,1,3), (3,2,1); \\ 0, & \text{otherwise.} \end{cases}$$

EXERCISE 14 *Verify that*

$$\varepsilon_{ijk}\varepsilon_{lmk} = \begin{cases} 1, & \text{if} \quad i = l \text{ and } j = m; \\ -1, & \text{if} \quad i = m \text{ and } j = l; \\ 0, & \text{otherwise.} \end{cases}$$

The connection between the cross product and volumes hinges on the following algebraic operation.

DEFINITION. *The* **scalar triple product** *of* $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{E}$ *is* $\left[\mathbf{a}, \mathbf{b}, \mathbf{c}\right] = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$.

Geometrically, $\left[\mathbf{a}, \mathbf{b}, \mathbf{c}\right]$ gives the signed volume of the parallelepiped having $\mathbf{a}, \mathbf{b}, \mathbf{c}$ as adjacent edges, as shown in Figure 1.5.

From an algebraic perspective, the scalar triple product is an **alternating 3-linear form**, that is, the mapping $[\cdot,\cdot,\cdot]\colon \mathbb{E}^3 \to \mathbb{R}$ has the following properties:

1. $\left[\cdot,\cdot,\cdot\right]$ is linear in each argument.
2. For all ordered triples $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) \in \mathbb{E}^3$, $\left[\mathbf{v}_i, \mathbf{v}_j, \mathbf{v}_k\right] = \varepsilon_{ijk}\left[\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\right]$.

EXERCISE 15 *Show that* $\left[\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\right] = 0$ *whenever* $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} \subset \mathbb{E}$ *is a linearly dependent set. Show that* $\left[\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\right] = \pm 1$ *whenever* $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ *is an orthonormal basis for* $\mathbb{E}$.

When the numbering of an orthonormal basis $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ satisfies the condition $[\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3] = 1$, we say that the basis has **positive orientation**.

Alternating 3-linear forms stand in an interesting relationship to each other:

THEOREM 1.2.1 (PROPORTIONALITY OF ALTERNATING 3-LINEAR FORMS). *If* $f\colon \mathbb{E}^3 \to \mathbb{R}$ *is a nonzero alternating 3-linear form and* g *is any other alternating 3-linear form, then there exists a scalar* $D \in \mathbb{R}$ *such that*

$$g(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) = Df(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) \tag{1.2.4}$$

for all ordered triples $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) \in \mathbb{E}^3$.

Crucially, the scalar D is independent of the vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$.

PROOF: Since f is nonzero, it is possible to pick $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3) \in \mathbb{E}^3$ such that $f(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3) \neq 0$. By Exercise 15, $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ cannot be linearly dependent, so it constitutes a basis for $\mathbb{E}$. Define

$$D = \frac{g(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)}{f(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)}.$$

We must show that Equation (1.2.4) holds using this choice of D, for all ordered triples $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) \in \mathbb{E}^3$. Each vector in such an ordered triple has an expansion of the form $\mathbf{v}_i = c_{ij}\mathbf{u}_j$. By 3-linearity,

$$\begin{aligned} f(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) &= f(c_{1j}\mathbf{u}_j, c_{2k}\mathbf{u}_k, c_{3l}\mathbf{u}_l) \\ &= c_{1j}c_{2k}c_{3l} f(\mathbf{u}_j, \mathbf{u}_k, \mathbf{u}_l) \\ &= \varepsilon_{jkl}c_{1j}c_{2k}c_{3l} f(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3) \\ &= C\, f(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3) \neq 0, \end{aligned}$$

where $C = \varepsilon_{jkl}c_{1j}c_{2k}c_{3l}$. Similarly, $g(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) = C\, g(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$, for the same constant C. The identity (1.2.4) follows. ∎

Of special interest are alternating 3-linear forms generated when second-order tensors act on ordered triples of vectors. If $\mathsf{A} \in \mathrm{L}(\mathbb{E})$ and f is any nonzero alternating 3-linear form, then the mapping defined by

$$(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) \mapsto f(\mathsf{A}\mathbf{v}_1, \mathsf{A}\mathbf{v}_2, \mathsf{A}\mathbf{v}_3) \tag{1.2.5}$$

is also an alternating 3-linear form. Hence, by Theorem 1.2.1, there exists a scalar $D_A \in \mathbb{R}$ such that

$$f(\mathsf{A}\mathbf{v}_1, \mathsf{A}\mathbf{v}_2, \mathsf{A}\mathbf{v}_3) = D_A f(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3), \tag{1.2.6}$$

for all ordered triples $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) \in \mathbb{E}^3$. More remarkable is the following fact:

THEOREM 1.2.2 (A SCALAR INVARIANT). *The scalar factor D_A, defined by Equation (1.2.6) for the tensor $\mathsf{A} \in \mathrm{L}(\mathbb{E})$ and the alternating 3-linear form f, is independent of f.*

PROOF: If g is any other nonzero alternating 3-linear form, we must show that $g(\mathsf{A}\mathbf{v}_1, \mathsf{A}\mathbf{v}_2, \mathsf{A}\mathbf{v}_3) = D_A g(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ for the scalar D_A defined by Equation (1.2.6). By Theorem 1.2.1, $g = Df$ for some nonzero scalar D. For the alternating 3-linear form defined by the mapping

$$(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) \mapsto g(\mathsf{A}\mathbf{v}_1, \mathsf{A}\mathbf{v}_2, \mathsf{A}\mathbf{v}_3),$$

we have

$$\begin{aligned} g(\mathsf{A}\mathbf{v}_1, \mathsf{A}\mathbf{v}_2, \mathsf{A}\mathbf{v}_3) &= D\, f(\mathsf{A}\mathbf{v}_1, \mathsf{A}\mathbf{v}_2, \mathsf{A}\mathbf{v}_3) \\ &= D D_A f(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) \\ &= D_A D\, f(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) = D_A\, g(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3). \end{aligned}$$

Thus the factor D_A is a characteristic of A, independent of the alternating 3-linear form f. ■

This fact justifies a special name and notation for the invariant factor D_A:

DEFINITION. *The* **determinant** *of* $\mathsf{A} \in \mathrm{L}(\mathbb{E})$ *is the number* $\det \mathsf{A} \in \mathbb{R}$ *such that*

$$(\det \mathsf{A}) f(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) = f(\mathsf{A}\mathbf{v}_1, \mathsf{A}\mathbf{v}_2, \mathsf{A}\mathbf{v}_3),$$

for every ordered triple $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) \in \mathbb{E}^3$*, for every alternating 3-linear form* f.

An immediate consequence is the following observation about the effect of a second-order tensor $\mathsf{A} \in \mathrm{L}(\mathbb{E})$ on volumes. Given three vectors $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{E}$, A maps their scalar triple product as follows:

$$\left[\mathbf{a}, \mathbf{b}, \mathbf{c}\right] \longmapsto \left[\mathsf{A}\mathbf{a}, \mathsf{A}\mathbf{b}, \mathsf{A}\mathbf{c}\right].$$

This mapping carries the parallelepiped corresponding to $\mathbf{a}, \mathbf{b}, \mathbf{c}$ into a new parallelepiped, as shown in Figure 1.6. The scalar quantity $\det \mathsf{A}$ gives the ratio of the transformed volume to the original volume: for all $(\mathbf{a}, \mathbf{b}, \mathbf{c}) \in \mathbb{E}^3$,

$$\left[\mathbf{a}, \mathbf{b}, \mathbf{c}\right] \det \mathsf{A} = \left[\mathsf{A}\mathbf{a}, \mathsf{A}\mathbf{b}, \mathsf{A}\mathbf{c}\right]. \qquad (1.2.7)$$

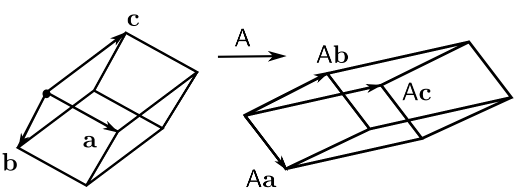

Figure 1.6. Effect of the linear transformation A on the parallelepiped having edges $\mathbf{a}, \mathbf{b}, \mathbf{c}$.

The definition of the determinant adopted here makes no reference to a basis for $\mathbb{E}$, and as a consequence the numerical value of $\det \mathsf{A}$ is independent of any matrix representation used for A. However, a calculation similar to that used in Theorem 1.2.2 yields a familiar formula for $\det \mathsf{A}$ in terms of its matrix representation A_{ij}

with respect to a basis $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ for $\mathbb{E}$:

$$
\begin{aligned}
(\det \mathsf{A})\big[\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\big] &= \big[\mathsf{A}\mathbf{u}_1, \mathsf{A}\mathbf{u}_2, \mathsf{A}\mathbf{u}_3\big] \\
&= \big[A_{i1}\mathbf{u}_i, A_{j2}\mathbf{u}_j, A_{k3}\mathbf{u}_k\big] \\
&= A_{i1}A_{j2}A_{k3}\big[\mathbf{u}_i, \mathbf{u}_j, \mathbf{u}_k\big] \\
&= \varepsilon_{ijk}A_{i1}A_{j2}A_{k3}\big[\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\big].
\end{aligned}
$$

It follows that

$$\det \mathsf{A} = \varepsilon_{ijk}A_{i1}A_{j2}A_{k3}. \tag{1.2.8}$$

The following theorem summarizes additional facts about determinants.

THEOREM 1.2.3 (PROPERTIES OF DETERMINANTS). *Let* $\mathbf{a}, \mathbf{b} \in \mathbb{E}$ *and* $\mathsf{A}, \mathsf{B} \in \mathrm{L}(\mathbb{E})$.

1. $\det \mathbf{a} \otimes \mathbf{b} = 0$.
2. $\det \mathsf{A} \neq 0$ *if and only if* A *is invertible.*
3. $\det(\mathsf{AB}) = \det \mathsf{A} \det \mathsf{B}$.
4. $\det \mathsf{A}^\top = \det \mathsf{A}$.
5. $\det(c\mathsf{A}) = c^3 \det \mathsf{A}$.

EXERCISE 16 *Prove Theorem 1.2.3.*

The connection between determinants and invertibility merits a brief digression. Let

$$\mathrm{GL}(\mathbb{E}) = \Big\{\mathsf{A} \in \mathrm{L}(\mathbb{E}) \;\Big|\; \det \mathsf{A} \neq 0\Big\}.$$

EXERCISE 17 *Show that* $\mathrm{GL}(\mathbb{E})$ *is a* **group** *under composition. That is:*

1. $\mathrm{GL}(\mathbb{E})$ *is closed under composition: if* $\mathsf{A}, \mathsf{B} \in \mathrm{GL}(\mathbb{E})$, *then* $\mathsf{AB} \in \mathrm{GL}(\mathbb{E})$.
2. *Composition is associative:* $(\mathsf{AB})\mathsf{C} = \mathsf{A}(\mathsf{BC})$.
3. $\mathrm{GL}(\mathbb{E})$ *possesses an identity* I.
4. *Every* $\mathsf{A} \in \mathrm{GL}(\mathbb{E})$ *possesses an inverse* $\mathsf{A}^{-1} \in \mathrm{GL}(\mathbb{E})$ *satisfying the equations* $\mathsf{AA}^{-1} = \mathsf{A}^{-1}\mathsf{A} = \mathsf{I}$.

$\mathrm{GL}(\mathbb{E})$ *is the* **general linear group** *over* $\mathbb{E}$.

EXERCISE 18 *Show that, with respect to any orthonormal basis for* $\mathbb{E}$, *the identity tensor* $\mathsf{I} \in \mathrm{GL}(\mathbb{E})$ *has a matrix representation in which* $I_{ij} = \delta_{ij}$, *where* δ_{ij} *is the* **Kronecker symbol**,

$$\delta_{ij} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{otherwise.} \end{cases}$$

EXERCISE 19 *Verify the* $\boldsymbol{\varepsilon\delta}$ **identity**,

$$\varepsilon_{ijk}\varepsilon_{lmk} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}. \tag{1.2.9}$$

EXERCISE 20 *Using index notation and the orthonormal basis* $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$, *prove the following identity for the* **vector triple product***:*

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \mathbf{b}(\mathbf{a} \cdot \mathbf{c}) - \mathbf{c}(\mathbf{a} \cdot \mathbf{b}). \tag{1.2.10}$$

EXERCISE 21 *Prove that*

$$(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{a} \times \mathbf{b}) = \|\mathbf{a}\|^2\|\mathbf{b}\|^2 - (\mathbf{a} \cdot \mathbf{b})^2. \tag{1.2.11}$$

1.2.4 Orthogonal Tensors

The following definition identifies a special class of invertible tensors encountered repeatedly in mechanics.

DEFINITION. *A second-order tensor* $\mathsf{Q} \in \mathrm{L}(\mathbb{E})$ *is* **orthogonal** *if* Q *is invertible and* $\mathsf{Q}^{-1} = \mathsf{Q}^\top$.

Clearly the identity tensor I is orthogonal—a fact that leads to a more refined algebraic observation:

EXERCISE 22 *Denote*

$$\mathrm{O}(\mathbb{E}) = \Big\{ \mathsf{Q} \in \mathrm{L}(\mathbb{E}) \ \Big|\ \mathsf{Q} \text{ is orthogonal} \Big\}.$$

Show that $\mathrm{O}(\mathbb{E})$ *is a* **subgroup** *of* $\mathrm{GL}(\mathbb{E})$, *that is,* $\mathrm{O}(\mathbb{E}) \subset \mathrm{GL}(\mathbb{E})$ *and is a group. Show that* $\det \mathsf{Q} = \det \mathsf{Q}^\top$ *whenever* $\mathsf{Q} \in \mathrm{O}(\mathbb{E})$, *and hence* $\det \mathsf{Q} = \pm 1$.

EXERCISE 23 *Show that the set*

$$\mathrm{O}^+(\mathbb{E}) = \left\{ \mathsf{Q} \in \mathrm{O}(\mathbb{E}) \;\middle|\; \det \mathsf{Q} = 1 \right\}$$

of **proper orthogonal transformations** *is a subgroup of* $\mathrm{O}(\mathbb{E})$.

Orthogonal tensors share with I an important geometric property:

EXERCISE 24 *Show that, if* $\mathsf{Q} \in \mathrm{O}(\mathbb{E})$*, then* $\mathsf{Q}\mathbf{u} \cdot \mathsf{Q}\mathbf{v} = \mathbf{u} \cdot \mathbf{v}$ *for all* $\mathbf{u}, \mathbf{v} \in \mathbb{E}$*. In particular,* $\|\mathsf{Q}\mathbf{u}\| = \|\mathbf{u}\|$ *for every* $\mathbf{u} \in \mathbb{E}$*. We call any length-preserving tensor an* **isometry**.

From the geometric point of view, proper orthogonal tensors act by rotating vectors, leaving their lengths unchanged. Section 2.1 explores this notion in more detail. An orthogonal tensor having negative determinant effects, in addition, a reflection across an axis. Exercise 24 shows that all orthogonal tensors preserve inner products. The following exercise shows that cross products behave somewhat differently when the transformation involves a reflection.

EXERCISE 25 *Let* $\mathsf{Q} \in \mathrm{O}(\mathbb{E})$ *and* $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{E}$*. Show that* $(\mathsf{Q}\mathbf{a} \times \mathsf{Q}\mathbf{b}) \cdot \mathsf{Q}\mathbf{c} = (\det \mathsf{Q})\,(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$*. Show also that* $\mathsf{Q}(\mathbf{a} \times \mathbf{b}) \cdot \mathsf{Q}\mathbf{c} = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$*. Conclude that* $\mathsf{Q}\mathbf{a} \times \mathsf{Q}\mathbf{b} = (\det \mathsf{Q})\,\mathsf{Q}(\mathbf{a} \times \mathbf{b})$*. Thus, only proper orthogonal tensors preserve cross products, even after accounting for the rotation.*

EXERCISE 26 *Let* $\mathsf{Q} \in \mathrm{O}(\mathbb{E})$*, and suppose that* $\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$ *is an orthonormal basis for* $\mathbb{E}$*. Show that* $\{\mathsf{Q}\mathbf{p}_1, \mathsf{Q}\mathbf{p}_2, \mathsf{Q}\mathbf{p}_3\}$ *is also an orthonormal basis for* $\mathbb{E}$*.*

EXERCISE 27 *The converse of Exercise 26 also holds: suppose* $\{\mathsf{Q}\mathbf{p}_1, \mathsf{Q}\mathbf{p}_2, \mathsf{Q}\mathbf{p}_3\}$ *is an orthonormal basis for* $\mathbb{E}$ *whenever* $\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$ *is an orthonormal basis. Prove that* $\mathsf{Q} \in \mathrm{O}(\mathbb{E})$.

In light of Exercises 26 and 27, to define an orthogonal tensor Q, it suffices to define $\mathsf{Q}\mathbf{p}_1, \mathsf{Q}\mathbf{p}_2, \mathsf{Q}\mathbf{p}_3$ for some orthonormal basis $\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$, in such a way that $\{\mathsf{Q}\mathbf{p}_1, \mathsf{Q}\mathbf{p}_2, \mathsf{Q}\mathbf{p}_3\}$ also constitutes an orthonormal basis.

1.2.5 Invariants of a Tensor

The determinant of a tensor $\mathsf{A} \in \mathrm{L}(\mathbb{E})$ is just one of many scalars associated with A that do not depend on the choice of bases or coordinate systems. We call such scalars **invariants** of A.

As we establish shortly, the following quantity associated with A is also invariant:

DEFINITION. *The* **trace** *of a dyadic product* $\mathbf{a} \otimes \mathbf{b}$ *is* $\operatorname{tr}(\mathbf{a} \otimes \mathbf{b}) = \mathbf{a} \cdot \mathbf{b}$.

Since every second-order tensor A is a linear combination of dyadic products, the definition of the trace extends by linearity to all of $\mathrm{L}(\mathbb{E})$:

$$\operatorname{tr} \sum_i c_i \, \mathbf{a}_i \otimes \mathbf{b}_i = \sum_i c_i \operatorname{tr}(\mathbf{a}_i \otimes \mathbf{b}_i).$$

EXERCISE 28 *Show that, with respect to any orthonormal basis,*

$$\operatorname{tr} \mathsf{A} = A_{11} + A_{22} + A_{33} = A_{ii}.$$

EXERCISE 29 *Show that* $\operatorname{tr}(\mathsf{AB}) = \operatorname{tr}(\mathsf{BA})$. *Conclude that* $\operatorname{tr}(\mathsf{BAB}^{-1}) = \operatorname{tr}(\mathsf{A})$ *whenever* B *is invertible.*

Using the trace we define, for later chapters, a scalar-valued product of second-order tensors as follows:

DEFINITION. *The* **double-dot product** *of* $\mathsf{A}, \mathsf{B} \in \mathrm{L}(\mathbb{E})$ *is*

$$\mathsf{A} : \mathsf{B} = \operatorname{tr}(\mathsf{A}^\top \mathsf{B}). \tag{1.2.12}$$

Clearly $\operatorname{tr}(\mathsf{A}) = \mathsf{I} : \mathsf{A}$. Additional properties follow almost as easily:

EXERCISE 30 *(a) Show that, with respect to any orthonormal basis,* $\mathsf{A} : \mathsf{B} = A_{ij} B_{ij}$. *(b) Show that the operation* $:$ *possesses the properties of an inner product, as listed in Section 1.1.*

To show that $\operatorname{tr} \mathsf{A}$ is an invariant and to identify other invariants require some basic facts about characteristic polynomials and eigenvalues.

DEFINITION. *If* $\mathsf{A} \in \mathrm{L}(\mathbb{E})$, *its* **characteristic polynomial** *is the cubic expression*

$$p_A(\lambda) = \det(\lambda \mathsf{I} - \mathsf{A}). \tag{1.2.13}$$

The **eigenvalues** *of* A *are the roots* $\lambda_1, \lambda_2, \lambda_3$ *of* p_A.

The condition $p_A(\lambda) = 0$ for a root implies the existence of at least one nonzero vector $\mathbf{p}$ such that $(\lambda \mathsf{I} - \mathsf{A})\mathbf{p} = \mathbf{0}$, that is, $\mathsf{A}\mathbf{p} = \lambda \mathbf{p}$. Geometrically, the action of A in the direction of $\mathbf{p}$ reduces to scalar multiplication. Since A is linear, whenever $\mathbf{p}$ is an eigenvector, so is any scalar multiple of $\mathbf{p}$. Hence, we can choose eigenvectors that have unit length whenever doing so is convenient.

DEFINITION. *If* $\mathsf{A} \in \mathrm{L}(\mathbb{E})$, *a nonzero vector* $\mathbf{p}$ *for which* $\mathsf{A}\mathbf{p} = \lambda\mathbf{p}$ *is an* **eigenvector** *of* A *associated with* λ.

A priori, a root λ of the characteristic polynomial p_A can be either real or complex with nonzero imaginary part. The fact that $p_A(\lambda)$ is cubic with real coefficients guarantees that at least one eigenvalue is real. The other two eigenvalues may be either (1) both real or (2) complex conjugates of each other. If $\lambda = d + iw$ is a complex eigenvalue with nonzero imaginary part w, then any eigenvector that corresponds to λ may have the complex form $\mathbf{p} + i\mathbf{q}$, where $\mathbf{p}, \mathbf{q} \in \mathbb{E}$. We call an eigenvector $\mathbf{p} + i\mathbf{q}$ of A for which $\mathbf{q} = \mathbf{0}$ a **real eigenvector**.

The eigenvalues of A are arguably its most important invariants. But computing them requires finding the roots of a cubic polynomial, which can be inconvenient. To identify another set of invariants for A, observe that $p_A(\lambda) = \lambda^3 + \cdots$ is the only cubic polynomial having leading coefficient 1 that has the roots $\lambda_1, \lambda_2, \lambda_3$. Therefore, A uniquely determines the remaining three coefficients in $p_A(\lambda)$. It follows that these coefficients must be invariants for A.

DEFINITION. *The coefficients* $\mathrm{I}_A, \mathrm{II}_A, \mathrm{III}_A$ *in the expansion*

$$p_A(\lambda) = \lambda^3 - \mathrm{I}_A\lambda^2 + \mathrm{II}_A\lambda - \mathrm{III}_A \tag{1.2.14}$$

are the **principal invariants** *of* A.

EXERCISE 31 *Use the property (1.2.7) of* $\det \mathsf{A}$ *to show that* $p_A(\lambda) = 0$ *if and only if* λ *satisfies the equation*

$$\big[\lambda\mathbf{a} - \mathsf{A}\mathbf{a}, \lambda\mathbf{b} - \mathsf{A}\mathbf{b}, \lambda\mathbf{c} - \mathsf{A}\mathbf{c}\big] = 0 \tag{1.2.15}$$

for all $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{E}$.

Straightforward expansion of the vector triple product in Equation (1.2.15) and dividing through by the common factor $\big[\mathbf{a}, \mathbf{b}, \mathbf{c}\big]$ (which we can always arrange to be nonzero by choosing $\mathbf{a}, \mathbf{b}, \mathbf{c}$ appropriately) yield:

$$\begin{aligned} 0 = \lambda^3 &- \frac{\big(\big[\mathsf{A}\mathbf{a}, \mathbf{b}, \mathbf{c}\big] + \big[\mathbf{a}, \mathsf{A}\mathbf{b}, \mathbf{c}\big] + \big[\mathbf{a}, \mathbf{b}, \mathsf{A}\mathbf{c}\big]\big)}{\big[\mathbf{a}, \mathbf{b}, \mathbf{c}\big]}\lambda^2 \\ &+ \frac{\big(\big[\mathbf{a}, \mathsf{A}\mathbf{b}, \mathsf{A}\mathbf{c}\big] + \big[\mathsf{A}\mathbf{a}, \mathbf{b}, \mathsf{A}\mathbf{c}\big] + \big[\mathsf{A}\mathbf{a}, \mathsf{A}\mathbf{b}, \mathbf{c}\big]\big)}{\big[\mathbf{a}, \mathbf{b}, \mathbf{c}\big]}\lambda - \det \mathsf{A} = 0. \end{aligned} \tag{1.2.16}$$

Since it has leading coefficient 1, the right side of this equation is $p_A(\lambda)$. It follows from Equation (1.2.14) that

$$\mathrm{I}_A = \frac{[\mathsf{A}\mathbf{a}, \mathbf{b}, \mathbf{c}] + [\mathbf{a}, \mathsf{A}\mathbf{b}, \mathbf{c}] + [\mathbf{a}, \mathbf{b}, \mathsf{A}\mathbf{c}]}{[\mathbf{a}, \mathbf{b}, \mathbf{c}]},$$

$$\mathrm{II}_A = \frac{[\mathbf{a}, \mathsf{A}\mathbf{b}, \mathsf{A}\mathbf{c}] + [\mathsf{A}\mathbf{a}, \mathbf{b}, \mathsf{A}\mathbf{c}] + [\mathsf{A}\mathbf{a}, \mathsf{A}\mathbf{b}, \mathbf{c}]}{[\mathbf{a}, \mathbf{b}, \mathbf{c}]},$$

$$\mathrm{III}_A = \det \mathsf{A}, \tag{1.2.17}$$

for any $\mathbf{a}, \mathbf{b}, \mathbf{c}$ having nonzero scalar triple product.

Equations (1.2.17) confirm, in particular, that $\det \mathsf{A}$ is an invariant for A. But directly reducing the more opaque expressions for I_A and II_A requires some effort. The fact that these quantities are invariant under changes in basis for $\mathbb{E}$ affords an alternative approach, as the following exercise illustrates.

EXERCISE 32 *Expand $p_A(\lambda)$ with respect to the orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ to prove the following identities:*

$$\mathrm{I}_A = \operatorname{tr} \mathsf{A}, \qquad \mathrm{II}_A = \tfrac{1}{2}[(\operatorname{tr} \mathsf{A})^2 - \operatorname{tr}(\mathsf{A}^2)], \qquad \mathrm{III}_A = \det \mathsf{A}. \tag{1.2.18}$$

From Equations (1.2.17) and (1.2.18) it follows that, for any $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{E}$,

$$[\mathbf{a}, \mathbf{b}, \mathbf{c}] \operatorname{tr} \mathsf{A} = [\mathsf{A}\mathbf{a}, \mathbf{b}, \mathbf{c}] + [\mathbf{a}, \mathsf{A}\mathbf{b}, \mathbf{c}] + [\mathbf{a}, \mathbf{b}, \mathsf{A}\mathbf{c}] \tag{1.2.19}$$

whenever $\mathsf{A} \in \mathrm{L}(\mathbb{E})$.

1.2.6 Derivatives of Tensor-Valued Functions

The remainder of this chapter briefly reviews concepts from the differential calculus of tensor-valued functions. Section 2.2 presents further aspects of differential calculus specific to continuum mechanics, and concepts from integral calculus appear in Section 2.4 and Appendix B.

First, consider vector- and tensor-valued functions of time, regarded as a single parameter $t \in \mathbb{R}$. For such functions, the concept of differentiability extends the concept for scalar-valued functions in a straightforward manner:

DEFINITION. *A vector-valued function $\mathbf{a}\colon \mathbb{R} \to \mathbb{E}$ is* **differentiable** *at $t \in \mathbb{R}$ if the real-valued function $\mathbf{a}(t) \cdot \mathbf{e}$ is differentiable at t for every unit vector $\mathbf{e} \in \mathbb{E}$. A tensor-valued function $\mathsf{A}\colon \mathbb{R} \to \mathrm{L}(\mathbb{E})$ is differentiable at $t \in \mathbb{R}$ if the real-valued function $\mathbf{d} \cdot \mathsf{A}(t)\mathbf{e}$ is differentiable at t for all unit vectors $\mathbf{d}, \mathbf{e} \in \mathbb{E}$.*

The following theorem proves useful in the next chapter:

THEOREM 1.2.4 (DERIVATIVE OF A DETERMINANT). *If* $\mathsf{A}\colon \mathbb{R} \to \mathrm{GL}(\mathbb{E})$ *is differentiable at every* $t \in \mathbb{R}$*, then*

$$\frac{d}{dt}\det\mathsf{A} = \operatorname{tr}\left(\frac{d\mathsf{A}}{dt}\mathsf{A}^{-1}\right)\det\mathsf{A}. \tag{1.2.20}$$

PROOF: Applying the product rule to Equation (1.2.7) yields

$$\begin{aligned}
\left[\mathbf{a},\mathbf{b},\mathbf{c}\right]\frac{d}{dt}\det\mathsf{A} &= \left[\frac{d\mathsf{A}}{dt}\mathbf{a},\mathsf{A}\mathbf{b},\mathsf{A}\mathbf{c}\right] + \left[\mathsf{A}\mathbf{a},\frac{d\mathsf{A}}{dt}\mathbf{b},\mathsf{A}\mathbf{c}\right] + \left[\mathsf{A}\mathbf{a},\mathsf{A}\mathbf{b},\frac{d\mathsf{A}}{dt}\mathbf{c}\right] \\
&= \left[\frac{d\mathsf{A}}{dt}\mathsf{A}^{-1}\mathsf{A}\mathbf{a},\mathsf{A}\mathbf{b},\mathsf{A}\mathbf{c}\right] + \left[\mathsf{A}\mathbf{a},\frac{d\mathsf{A}}{dt}\mathsf{A}^{-1}\mathsf{A}\mathbf{b},\mathsf{A}\mathbf{c}\right] + \left[\mathsf{A}\mathbf{a},\mathsf{A}\mathbf{b},\frac{d\mathsf{A}}{dt}\mathsf{A}^{-1}\mathsf{A}\mathbf{c}\right] \\
&= \operatorname{tr}\left(\frac{d\mathsf{A}}{dt}\mathsf{A}^{-1}\right)\left[\mathsf{A}\mathbf{a},\mathsf{A}\mathbf{b},\mathsf{A}\mathbf{c}\right] = \left[\mathbf{a},\mathbf{b},\mathbf{c}\right]\operatorname{tr}\left(\frac{d\mathsf{A}}{dt}\mathsf{A}^{-1}\right)\det\mathsf{A},
\end{aligned}$$

the last two steps following from the identities (1.2.19) and (1.2.7). Since $\mathbf{a}, \mathbf{b}, \mathbf{c}$ are arbitrary, Equation (1.2.20) follows. ■

Now consider vector- and tensor-valued functions of position $\mathbf{x} \in \mathbb{E}$. The concepts of differentiability and the derivatives of real-valued functions of real variables generalize to the following types of functions that arise throughout continuum mechanics:

1. **scalar fields**, that is, real-valued functions $f\colon \mathbb{E} \to \mathbb{R}$ of position;
2. **vector fields**, that is, vector-valued functions $\mathbf{a}\colon \mathbb{E} \to \mathbb{E}$ of position; and
3. **tensor fields**, that is, tensor-valued functions $\mathsf{A}\colon \mathbb{E} \to \mathrm{L}(\mathbb{E})$ of position.

DEFINITION. *A scalar field* $f\colon \mathbb{E} \to \mathbb{R}$ *is* **differentiable** *at* $\mathbf{x}$ *if there exists a unique vector* $\mathbf{g} \in \mathbb{E}$ *such that*

$$\mathbf{g}\cdot\mathbf{e} = \lim_{h\to 0}\frac{f(\mathbf{x}+h\mathbf{e}) - f(\mathbf{x})}{h} \tag{1.2.21}$$

for all unit vectors $\mathbf{e} \in \mathbb{E}$*. In this case, the vector* $\mathbf{g}$ *is the* **derivative** *of* f *at* $\mathbf{x}$*, denoted by* $\mathbf{g} = \operatorname{grad} f(\mathbf{x})$ *and often called the* **gradient** *of* f *at* $\mathbf{x}$*.*

The quantity $\mathbf{g}\cdot\mathbf{e} = \operatorname{grad} f(\mathbf{x})\cdot\mathbf{e}$ is the **directional derivative** of f at $\mathbf{x}$ in the direction $\mathbf{e}$. When f is differentiable at every point $\mathbf{x}$ in a region of $\mathbb{E}$, we treat $\operatorname{grad} f$ as a function of $\mathbf{x}$. In Equation (1.2.21), substituting for $\mathbf{e}$ the basis vectors

$\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ gives the partial derivatives $\partial f/\partial x_i$ associated with a Cartesian coordinate system for $\mathbb{E}$ and yields the following representation for $\operatorname{grad} f$ in $\mathbb{R}^3$:

$$\begin{bmatrix} \partial f/\partial x_1 \\ \partial f/\partial x_2 \\ \partial f/\partial x_3 \end{bmatrix}.$$

In using coordinate representations of vector and tensor fields and their derivatives, the following index notation meshes well with the Einstein summation convention: for any differentiable scalar field f,

$$\frac{\partial f}{\partial x_i} = f_{,i}.$$

Thus, for example, $\operatorname{grad} f = f_{,i}\mathbf{e}_i$.

These concepts extend to vector and tensor fields as follows:

DEFINITION. *A vector field* $\mathbf{a}\colon \mathbb{E} \to \mathbb{E}$ *is* **differentiable** *at* $\mathbf{x}$ *if the scalar field* $\mathbf{a}(\mathbf{x})\cdot\mathbf{e}$ *is differentiable at* $\mathbf{x}$ *for every unit vector* $\mathbf{e} \in \mathbb{E}$*. A tensor field* $\mathsf{A}\colon \mathbb{E} \to \mathrm{L}(\mathbb{E})$ *is* **differentiable** *at* $\mathbf{x}$ *if the scalar field* $\mathbf{d} \cdot \mathsf{A}(\mathbf{x})\mathbf{e}$ *is differentiable at* $\mathbf{x}$ *for all unit vectors* $\mathbf{d}, \mathbf{e} \in \mathbb{E}$*. We denote the derivatives of* $\mathbf{a}$ *and* A *at* $\mathbf{x}$ *by* $\operatorname{grad} \mathbf{a}(\mathbf{x})$ *and* $\operatorname{grad} \mathsf{A}(\mathbf{x})$*, respectively.*

As a function of $\mathbf{x}$, the derivative $\operatorname{grad} \mathbf{a}$ is a tensor field, defined by its action: for any vector $\mathbf{b} \in \mathbb{E}$:

$$(\operatorname{grad} \mathbf{a})^\top \mathbf{b} = \operatorname{grad}(\mathbf{a} \cdot \mathbf{b}).$$

By extending the reasoning given for $\operatorname{grad} f$, substituting the basis vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ for the general unit vector $\mathbf{e}$ yields a coordinate representation for $\operatorname{grad} \mathbf{a}(\mathbf{x})$:

$$\operatorname{grad} \mathbf{a} = \frac{\partial a_i}{\partial x_j}\mathbf{e}_i \otimes \mathbf{e}_j = a_{i,j}\mathbf{e}_i \otimes \mathbf{e}_j,$$

which has the following 3×3 matrix representation:

$$\begin{bmatrix} \partial a_1/\partial x_1 & \partial a_1/\partial x_2 & \partial a_1/\partial x_3 \\ \partial a_2/\partial x_1 & \partial a_2/\partial x_2 & \partial a_2/\partial x_3 \\ \partial a_3/\partial x_1 & \partial a_3/\partial x_2 & \partial a_3/\partial x_3 \end{bmatrix}.$$

Several related differential operators arise in mechanics:

DEFINITION. *The* **divergence** *of a differentiable vector field* $\mathbf{a}$ *is*

$$\operatorname{div} \mathbf{a} = \operatorname{tr}(\operatorname{grad} \mathbf{a}). \tag{1.2.22}$$

The **curl** *of* $\mathbf{a}$ *is a vector field having the following action: for any* $\mathbf{b} \in \mathbb{E}$*,*

$$(\operatorname{curl} \mathbf{a}) \cdot \mathbf{b} = \operatorname{div} (\mathbf{a} \times \mathbf{b}). \tag{1.2.23}$$

Although the direct use of $\operatorname{grad} \mathsf{A}$ arises far more rarely, we have occasion to use the following quantity:

DEFINITION. *The* **divergence**[1] *of a differentiable tensor field* A *is a vector field having the following action: for all* $\mathbf{b} \in \mathbb{E}$*,*

$$(\operatorname{div} \mathsf{A}) \cdot \mathbf{b} = \operatorname{div} (\mathsf{A}^\top \mathbf{b}). \tag{1.2.24}$$

EXERCISE 33 *Suppose that* $\mathbf{a}(\mathbf{x})$ *and* $\mathbf{b}(\mathbf{x})$ *are differentiable vector fields. Show that*

$$\operatorname{div} (\mathbf{a} \otimes \mathbf{b}) = \begin{bmatrix} \operatorname{div} (a_1 \mathbf{b}) \\ \operatorname{div} (a_2 \mathbf{b}) \\ \operatorname{div} (a_3 \mathbf{b}) \end{bmatrix}.$$

EXERCISE 34 *Using Cartesian coordinates, show that if* $\mathbf{a} \colon \mathbb{E} \to \mathbb{E}$ *and* $\mathsf{A} \colon \mathbb{E} \to \mathrm{L}(\mathbb{E})$ *are differentiable,*

$$\operatorname{div} \mathbf{a} = \frac{\partial a_i}{\partial x_i} \quad \textit{and} \quad \operatorname{curl} \mathbf{a} = \varepsilon_{ijk} \frac{\partial a_i}{\partial x_j} \mathbf{e}_k = \varepsilon_{ijk} a_{i,j} \mathbf{e}_k. \tag{1.2.25}$$

Show that

$$\operatorname{div} \mathsf{A} = \frac{\partial A_{ij}}{\partial x_j} \mathbf{e}_i = A_{ij,j} \mathbf{e}_i.$$

1.2.7 Summary

Tensors arise in continuum mechanics as linear transformations of vectors in $\mathbb{E}$. We denote the space of all such transformations as $\mathrm{L}(\mathbb{E})$. Although matrix representations of tensors are useful for some purposes, their essential properties—including their actions on vectors and such scalar invariants as the trace and determinant—have geometric and algebraic significance independent of any choice of basis for $\mathbb{E}$. Differentiation of vector- and tensor-valued functions simply extends the concept of differentiation of scalar-valued functions.

[1] Some but not all authors adopt the transpose of this definition, along with corresponding changes in Theorem B.2.2 and its consequences. See, for example, [12, p. 40]. The notation adopted here is consistent with that used, for example, in [35, p. 278] and [9, p. 247].