Chapter 1

Preliminaries

1.1 The Vector Concept Revisited

The concept of a vector has been one of the most fruitful ideas in all of mathematics, and it is not surprising that we receive repeated exposure to the idea throughout our education. Students in elementary mathematics deal with vectors in component form — with quantities such as

$$\mathbf{x} = (2, 1, 3)$$

for example. But let us examine this situation more closely. Do the components 2, 1, 3 determine the vector $\mathbf{x}$? They surely do if we specify the basis vectors of the coordinate frame. In elementary mathematics these are supposed to be mutually orthogonal and of unit length; even then they are not fully characterized, however, because such a frame can be rotated. In the description of many common phenomena we deal with vectorial quantities like forces that have definite directions and magnitudes. An example is the force your body exerts on a chair as you sit in front of the television set. This force does not depend on the coordinate frame employed by someone writing a textbook on vectors somewhere in Russia or China. Because the vector $\mathbf{f}$ representing a particular force is something objective, we should be able to write it in such a form that it ceases to depend on the details of the coordinate frame. The simplest way is to incorporate the frame vectors $\mathbf{e}_i$, $i = 1, 2, 3$, explicitly into the notation: if $\mathbf{x}$ is a vector we may write

$$\mathbf{x} = \sum_{i=1}^{3} x_i \mathbf{e}_i. \tag{1.1}$$

Then if we wish to change the frame, we should do so in such a way that $\mathbf{x}$ remains the same. This of course means that we cannot change only the frame vectors $\mathbf{e}_i$: we must change the components x_i correspondingly. So the components of a vector $\mathbf{x}$ in a new frame are not independent of those in the old frame.

1.2 A First Look at Tensors

In what follows we shall discuss how to work with vectors using different coordinate frames. Let us note that in mechanics there are objects of another nature. For example, there is a so-called tensor of inertia. This is an objective characteristic of a solid body, determining how the body rotates when torques act upon it. If the body is considered in a Cartesian frame, the tensor of inertia is described by a 3×3 matrix. If we change the frame, the matrix elements change according to certain rules. In textbooks on mechanics the reader can find lengthy discussions on how to change the matrix elements to maintain the same objective characteristic of the body when the new frame is also Cartesian. Although the tensor of inertia is objective (i.e., frame-independent), it is not a vector: it belongs to another class of mathematical objects. Many such *tensors of the second rank* arise in continuum mechanics: tensors of stress, strain, etc. They characterize certain properties of a body at each point; again, their "components" should transform in such a way that the tensors themselves do not depend on the frame.

For both vectors and tensors we can introduce various operations. Of course, the introduction of any new operation should be done in such a way that the results agree with known special cases when such familiar cases are met. If we introduce, say, dot multiplication of a tensor by a vector, then in a Cartesian frame the operation should resemble the multiplication of a matrix by a column vector. Similarly, the multiplication of two tensors should be defined so that in a Cartesian frame the operation involves matrix multiplication. To this end we consider *dyads* of vectors. These are quantities of the form[1]

$$\mathbf{e}_i \mathbf{e}_j .$$

[1]The quantity $\mathbf{e}_i\mathbf{e}_j$ is also called the *tensor product* of the vectors $\mathbf{e}_i$ and $\mathbf{e}_j$, and is sometimes denoted $\mathbf{e}_i \otimes \mathbf{e}_j$. Our notation (without the symbol $\otimes$) emphasizes that, for example, $\mathbf{e}_1\mathbf{e}_2$ is an elemental object belonging to the set of second-rank tensors, in the same way that $\mathbf{e}_1$ is an elemental object belonging to the set of vectors. Note that $\mathbf{e}_2\mathbf{e}_1$ and $\mathbf{e}_1\mathbf{e}_2$ are different objects, however. The term "tensor product" indicates that the operation shares certain properties with the product we know from elementary algebra.

A tensor may then be represented as

$$\sum_{i,j} a_{ij}\, \mathbf{e}_i \mathbf{e}_j$$

where the a_{ij} are the components of the tensor. We compare with equation (1.1) and notice the similarity in notation.

Natural objects can possess characteristics described by tensors of higher rank. For example, the elastic properties of a body are described by a tensor of the fourth rank (i.e., a tensor whose elemental parts are of the form $\mathbf{abcd}$, where $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$ are vectors). This means that in general the properties of a body are given by a "table" consisting of $3 \times 3 \times 3 \times 3 = 81$ elements. The elements change according to certain rules if we change the frame.

Tensors also occur in electrodynamics, the general theory of relativity, and many other sciences that deal with objects situated or distributed in space.

1.3 Assumed Background

In what follows we suppose a familiarity with the dot and cross products and their expression in Cartesian frames. Recall that if $\mathbf{a}$ and $\mathbf{b}$ are vectors, then by definition

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos\theta,$$

where $|\mathbf{a}|$ and $|\mathbf{b}|$ are the magnitudes of $\mathbf{a}$ and $\mathbf{b}$ and θ is the (smaller) angle between $\mathbf{a}$ and $\mathbf{b}$. In a Cartesian frame with basis vectors[2] $\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3$ where $\mathbf{a}$ and $\mathbf{b}$ are expressed as

$$\mathbf{a} = a_1\mathbf{i}_1 + a_2\mathbf{i}_2 + a_3\mathbf{i}_3, \qquad \mathbf{b} = b_1\mathbf{i}_1 + b_2\mathbf{i}_2 + b_3\mathbf{i}_3,$$

we have

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + a_3 b_3.$$

Also recall that

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{i}_1 & \mathbf{i}_2 & \mathbf{i}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

[2]From now on we reserve the symbol $\mathbf{i}$ for the basis vectors of a Cartesian system.

The dot product will play a role in our discussion from the very beginning. The cross product will be used as needed, and a fuller discussion will appear in § 2.6.

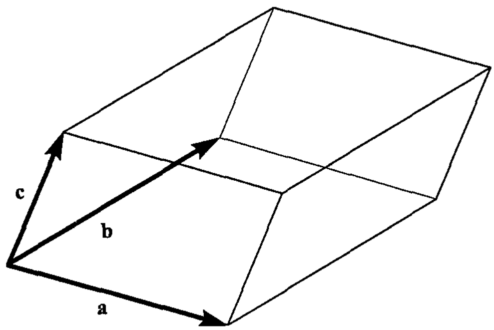

Fig. 1.1 Geometrical meaning of the scalar triple product.

Given three vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ we can form the *scalar triple product*

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}).$$

This may be interpreted as the volume of the parallelepiped having $\mathbf{a}, \mathbf{b}, \mathbf{c}$ as three of its co-terminal edges (Fig. 1.1). In rectangular components we have, according to the expressions above,

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

Permissible manipulations with the scalar triple product include cyclic interchange of the vectors:

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

Exercise 1.1 What does the condition $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) \neq 0$ say about $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$? (Hints for this and many other exercises appear in Appendix B beginning on page 165.)

So far we have made reference to vectors in a three-dimensional space, and shall continue this practice throughout most of the book. It is also possible (and sometimes useful) to introduce vectors in a more general space of $n > 3$ dimensions, e.g.,

$$\mathbf{a} = \mathbf{i}_1 a_1 + \mathbf{i}_2 a_2 + \cdots + \mathbf{i}_n a_n.$$

It turns out that many (but not all) of the principles and techniques we shall learn have direct extensions to such higher-dimensional spaces. It is also true that many three-dimensional notions can be reconsidered in *two* dimensions. The reader should take the time to reduce each three-dimensional formula to its two-dimensional analogue to understand what happens with the corresponding assertion.

1.4 More on the Notion of a Vector

Before closing out this chapter we should mention the notions of a vector as a "directed line segment" or "quantity having magnitude and direction." We find these in elementary math and physics textbooks. But it is easy to point to a situation in which a quantity of interest has magnitude and direction but is not a vector. The total electric current flowing in a thin wire is one example: to describe the current we must specify the rate of flow of electrons, the orientation of the wire, and the sense of electron movement along the wire. However, if two wires lie in a plane and carry equal currents running perpendicular to each other, then we cannot duplicate their physical effects by replacing them with a third wire directed at a 45° angle with respect to the first two. Total electric currents cannot be considered as vector quantities since they do not combine according to the rule for vector addition.

Another problem concerns the notion of an n-dimensional Euclidean space. We sometimes hear it defined as the set each point of which is uniquely determined by n parameters. However, it is not reasonable to regard every list of n things as a possible vector. A tailor may take measurements for her client and compose an ordered list of lengths, widths, etc., and by the above "definition" the set of all such lists is an n-dimensional space. But in $\mathbb{R}^n$ any two points are joined by a vector whose components equal the differences of the corresponding coordinates of the points. In the tailor's case this notion is completely senseless. To make matters worse, in $\mathbb{R}^n$ one can multiply any vector by any real number and obtain another vector in the space. A tailor's list showing that someone is six meters tall would be pretty hard to find.

Such simplistic definitions can be dangerous. The notion of a vector was eventually elaborated in physics, more precisely in mechanics where the behavior of forces was used as the model for a general vector. However, forces have some rather strange features: for example, if we shift the point of application of a force acting on a solid body, then we must introduce a moment acting on the body or the motion of the body will change. So the mechanical forces that gave birth to the notion of a vector possess features not covered by the mathematical definition of a vector. In the classical mechanics of a rigid body we are allowed to move a force along its line of action but cannot simply shift it off that line. With a deformable body we cannot move a force anywhere because in doing so we immediately change the state of the body. We can vectorially add two forces acting on a rigid body (not forgetting about the moments arising during shift of the forces). On the other hand, if two forces act on two different material points then we can add the vectors that represent the forces, but will not necessarily obtain a new vector that is relevant to the physics of the situation. So we should understand that the idea of a vector in mathematics reflects only some of the features of the real object it describes.

We would like to mention something else about vectors. In elementary mathematics students use vectors and points quite interchangeably. However, these are objects of different natures: there are vectors in space and there are the points of a space. We can, for instance, associate with a vector in an n-dimensional Euclidean vector space a point in an n-dimensional Euclidean point space. We can then consider a vector $\mathbf{x}$ as a shift of all points of the point space by an amount specified by $\mathbf{x}$. The result of this map is the same space of points; each pair of points that correspond under the mapping define a vector $\mathbf{x}$ under which a point shifts into its image. This vector is what we find when we subtract the Cartesian coordinates of the initial point from those of the final point. If we add the fact that the composition of two such maps obeys the rules of vector addition, then we get a strict definition of the space introduced intuitively in elementary mathematics. Engineers might object to the imposition of such formality on an apparently simple situation, but mathematical rigor has proved its worth from many practical points of view. For example, a computer that processes information in the complete absence of intuition can deal properly with objects for which rigorous definitions and manipulation rules have been formulated.

This brings us to our final word, regarding the expression of vectors in component-free notation. The simple and compact notation $\mathbf{x}$ for a vector leads to powerful ways of exhibiting relationships between the many vector (and tensor) quantities that occur in mathematical physics. It permits us to accomplish manipulations that would take many pages and become quite confusing if done in component form. This is typical for problems of nonlinear physics, and for those where change of coordinate frame becomes necessary. The resulting formal nature of the manipulations means that computers can be expected to take on more and more of this sort of work, even at the level of original research.