Chapter 1

Vectors and Scalars

1.1 Introduction

There are two basic entities in the mathematical language of fundamental physics: scalars and vectors. A scalar is a pure number or a quantity such as volume, temperature, and speed, that is completely specified by a numerical value in standard units. For example, temperature is most commonly specified with respect to the Celsius, Fahrenheit, Kelvin, or Rankin scales. Scalars can refer to individual systems or to particles, such as the charge of an electron or the volume of a gas, or to variations within a continuous region, within which a scalar is assigned to every point, such as the temperature in a geographic region on a weather map. In the latter case, we speak of a scalar function.

Scalars do not always provide enough information to completely determine a physical quantity. The speed of a vehicle on a highway indicates the rate at which distance is travelled, as measured by a speedometer, but not the direction of travel, as tracking by a satellite navigation system would indicate. Direction is the additional information that vectors provide. Physical laws in subjects as diverse as mechanics, electromagnetism, and fluid mechanics can be expressed in an intuitive and concise form by using vectors. Moreover, abstract extensions of vectors and their operations provide the mathematical framework for quantum mechanics.

The widespread use of vectors in science and engineering is due in large part to the notation and exposition put forward in the late 1880s by Josiah Willard Gibbs, an American scientist, mathematician, and engineer, who is perhaps best known for establishing the foundation of thermodynamics. In fact, the notion of vectors is implicit in Isaac Newton’s Principia Mathematica, which was published in the late 1600s. Although Newton did not use vectors as such, his diagrams indicate
that he thought about forces in these terms. For example, Newton postulated two forces that act simultaneously can be treated as acting sequentially, thereby introducing the idea of the addition of vectors.

The foregoing discussion about the importance of vectors can be summarized as follows:

- Vectors provide a natural graphical representation of physical quantities that have a magnitude and direction.
- Vectors simplify the statement of many physical laws, including Newton’s laws of motion, Maxwell’s equations of electromagnetism, and the Navier–Stokes equations of fluid mechanics.
- Vectors have intuitive mathematical properties that simplify the formulation and solution of many physical problems.

We begin this chapter by defining Cartesian coordinate systems, which provide the foundation for discussing vectors. We then discuss the graphical and algebraic representations of vectors. Many aspects of vectors can be demonstrated in the plane, unencumbered by the presence of higher dimensions. Much of our introductory discussion will focus on such vectors, with excursions into higher dimensions taken as they are warranted.

### 1.2 Cartesian Coordinate Systems

The French mathematician and philosopher René Descartes introduced the notion of coordinate systems that bear his name (also known as a rectangular coordinate systems) in the 1630s to enable the position of a point in space to be specified. The definition of such coordinate systems is as follows:

**Definition 1.1.** A coordinate system for which the position of a point is marked by distances along a set of perpendicularly directed lines (axes) that intersect at the origin of the system is called a **Cartesian** or **rectangular coordinate system**. The displacement along each axis is the **coordinate** along that axis.

The simplest example of Cartesian coordinates is the one-dimensional space of real numbers, also known as the real line, and denoted by \( \mathbb{R} \). The coordinate of a point on this line (Fig 1.1) is a real number which is taken as positive (resp. negative) if the point lies to the right (resp. left) of the origin, which is assigned the number 0.
Vectors and Scalars

Figure 1.1: One-dimensional Cartesian coordinate system, showing the origin, labelled as 0, which separates positive and negative coordinates. A general coordinate is signified by $x$.

The two-dimensional Cartesian coordinate system, $\mathbb{R} \times \mathbb{R} \equiv \mathbb{R}^2$ is the coordinate system of the plane. Coordinates are defined by two perpendicular lines, whose intersection defines the origin (Fig 1.2). The origin separates positive from negative regions along each coordinate axis, with the result that there are four quadrants whose coordinates are indicated by $(+, +), (+, -), (-, +)$, and $(-, -)$, as indicated in Fig. 1.2. The standard convention is that positive values of $x$ are to the right of the origin, and positive values of $y$ are upward from the origin. This is called a “right-handed” coordinate system because if the thumb of your right hand points upward from the plane of the paper, your fingers curl around from the positive $x$- to the positive $y$-axis. Left-hand coordinate systems are also encountered, although they are less common, with coordinates defined accordingly.

The common convention is to label the horizontal axis by $x$ and the vertical axis by $y$. The coordinates of every point in the plane are then uniquely specified by the ordered pair $(x, y)$. The first entry is the position on the $x$-axis and the second entry is the position along the $y$ axis.

Figure 1.2: Right-handed two-dimensional Cartesian coordinate system showing the two coordinate axes, labelled $x$ and $y$, together with several points that show how the signs change with respect to the origin in the four quadrants of this coordinate system.
Three-dimensional Cartesian coordinates are similarly defined as the product $\mathbb{R} \times \mathbb{R} \times \mathbb{R} \equiv \mathbb{R}^3$. Three coordinate axes are three mutually perpendicular straight lines that intersect at the origin. In a right-handed system, as shown in Fig. 1.3, if your hand curls from the positive $x$-axis to the positive $y$-axis, your thumb will point along the positive $z$-axis. The coordinates $(x, y, z)$ of a typical point are obtained from the positions along the $x$-, $y$, and $z$-axis, respectively. Every point in $\mathbb{R}^3$ has a unique set of coordinates.

The extension of Cartesian coordinate systems to $n$ dimensions is now easy to understand, even though explicit diagrams are not available. This illustrates the power of mathematical abstraction. The coordinate axes are defined by the mutual intersection of $n$ straight lines, whose intersection point defines the origin of the coordinate system. The coordinates $(s_1, s_2, \ldots, s_n)$ of a point in $\mathbb{R}^n$ are obtained from the positions on the $n$th entry $x_n$ on the $n$ coordinate axis. Every point in $\mathbb{R}^n$ has a unique set of such coordinates.

### 1.3 Scalars and Vectors

#### 1.3.1 Scalars

We have already discussed scalars in the Introduction, but we will formalize that discussion by giving the definition of a scalar.

**Definition 1.2.** A **scalar** is a quantity that is completely specified by a real
Vectors and Scalars

Scalars can be positive or negative, and can be dimensionless or be expressed in a set of physical units.

Example 1.1. We have already given some basic examples, but here we provide other examples to illustrate additional properties of scalars. Temperature is a scalar and can take negative values, depending on the scale used. In the Celsius scale, a temperature of $-10^\circ$ C is perfectly acceptable, but not in the (absolute) Kelvin scale. Volume is always a positive quantity, but charge can be negative or positive. Changes to other scalars can also be positive or negative.

1.3.2 Vectors

Vectors were also discussed in the Introduction, so we again begin with a definition:

Definition 1.3. A vector is a quantity that has both magnitude and direction. The magnitude is positive, except for the zero vector. Vectors can have units or be dimensionless.

A vector is indicated graphically by an arrow directed from its tail, labelled by $A$ in Fig. 1.4(a), to its head, indicated by $B$. We denote this vector as $\mathbf{a} = \overrightarrow{AB}$. The ordering of the labels is important: $\overrightarrow{BA}$ is a vector whose tail is at $B$ and head at $A$, which has the reverse direction of $\overrightarrow{AB}$. The length or magnitude of a vector is the distance between the head and the tail. Two vectors are equivalent if they have the same magnitude and direction. Thus, rigid translations of the vector in Fig. 1.4(a),

Figure 1.4: (a) A vector represented by its tail $A$ and head $B$. (b) Three vectors that have the same magnitude and direction as that in (a) and are, therefore, equivalent to that vector.
that is, the same translation of the head and tail of the vector, produce equivalent vectors, as shown in Fig. 1.4(b). For example, on a wind map, equivalent vectors correspond to the same wind speed and direction. They are distinguished only by their geographical location, that is, by their latitude and longitude.

Vectors are inherently geometrical objects that need not be referred to any particular coordinate system. However, there are advantages to assigning coordinates to the heads and tails of all vectors in a calculation. The head and tail of a vector are points in an $n$-dimensional Cartesian coordinate system, which specifies the dimension of that vector. As indicated in Fig. 1.4(b), all the vector in a set can be translated so that their tails lie at the same point. A convenient choice for this point – and the one we will usually make in this course – is the origin. Figure 1.5(a) shows a two-dimensional vector. The tail of the vector is at the origin $(0,0)$ and the head is at $(a_1,a_2)$. The analogous representation of a three-dimensional vector is shown in Fig. 1.5(b). The tail of the vector is at $(0,0,0)$ and the head is at $(a_1,a_2,a_3)$. When its tail is at the origin, a vector is completely specified by the coordinates of its head, in which case, we write vectors as

$$a = (a_1,a_2), \quad a = (a_1,a_2,a_3),$$

(1.1)

for two- and three-dimensional vectors, respectively. We will confine ourselves mostly to two- and three-dimensional vectors, but will point out where our presentation has immediate generalizations to higher dimensions. For example, a vector in $n$ dimensions (whose tail is at the origin) can be written as an array of $n$ real numbers:

$$a = (a_1,a_2,\ldots,a_n),$$

(1.2)
Vectors and Scalars

where the ellipsis (...) mean that the omitted items are continued in the pattern initiated by the first several entries.

1.4 The Length and Direction of a Vector

According to Definition 1.3, a vector has both magnitude and a direction. The representation of a vector in Cartesian coordinates enables both of these quantities to be calculated for a vector in any dimension.

1.4.1 The Length of a Vector

Once a vector has been assigned coordinates, we can compute its length by using Pythagoras’ theorem. Referring to Fig. 1.5(a), the length of a two-dimensional vector is the hypotenuse of a right triangle whose other two sides lie along the x- and y-axes and have lengths $a_1$ and $a_2$, respectively. Hence, the magnitude of $\mathbf{a} = (a_1, a_2)$, which is denoted by $|\mathbf{a}|$, is

$$|\mathbf{a}| = (a_1^2 + a_2^2)^{\frac{1}{2}}. \quad (1.3)$$

The magnitude of a vector is a scalar in the units of the vector and is always a positive quantity, except for the special case of the zero vector, which will be discussed in the next chapter. The extension of this calculation to higher dimensions is based on the recursive application of Pythagoras’ theorem. To calculate the length of a vector $\mathbf{a} = (a_1, a_2, a_3)$, we start with the length of the two-dimensional vector $(a_1, a_2)$ in the x-y plane, which is regarded as the base of a right triangle whose height is the z-component of the vector, as shown in Fig. 1.5(b). The length of the three-dimensional vector is the hypotenuse of the triangle, which is

$$|\mathbf{a}| = (a_1^2 + a_2^2 + a_3^2)^{\frac{1}{2}}. \quad (1.4)$$

Proceeding to an n-dimensional vector $\mathbf{a} = (a_1, a_2, \ldots, a_n)$, the length of a vector in $n-1$ dimensions, $(a_1, a_2, \ldots, a_{n-1})$ forms the base of a right triangle whose height is $n$th component $a_n$ of the vector, which forms the hypotenuse of this triangle. The length of this vector is, therefore,

$$|\mathbf{a}| = (a_1^2 + a_2^2 + \cdots + a_n^2)^{\frac{1}{2}}. \quad (1.5)$$
1.4.2 The Direction of a Vector

There are several ways of specifying the direction of a vector. One way that is suggested by two-dimensional vectors and easily generalized to higher dimensions is in terms of direction cosines, that is, the cosine of the angle between the vector and each of the coordinate axes. Consider the construction in Fig. 1.6(a).

The shaded region indicates the right triangle formed by the origin and the points \((a_1,0)\) and \((a_1,a_2)\), with the vector as the hypotenuse. The cosine of the angle \(\phi\) is \(a_1/|\mathbf{a}|\). An analogous construction yields the cosine of the angle \(\psi\) between the vector and the \(y\)-axis: \(a_2/|\mathbf{a}|\). But these two direction cosines cannot be specified independently, because

\[
\cos^2 \phi + \cos^2 \psi = \frac{a_1^2}{|\mathbf{a}|^2} + \frac{a_2^2}{|\mathbf{a}|^2} = 1, \tag{1.6}
\]

where we have used (1.3) for the length of a two-dimensional vector.

In three dimensions, the direction cosines of \(\mathbf{a} = (a_1,a_2,a_3)\) are [Fig. 1.6(b)]:

\[
\cos \alpha = \frac{a_1}{|\mathbf{a}|}, \quad \cos \beta = \frac{a_2}{|\mathbf{a}|}, \quad \cos \gamma = \frac{a_3}{|\mathbf{a}|}. \tag{1.7}
\]

These angles cannot be specified independently because

\[
\cos^2 \alpha + \cos^2 \beta + \cos^2 \gamma = \frac{a_1^2 + a_2^2 + a_3^2}{|\mathbf{a}|^2} = 1, \tag{1.8}
\]

where we have used (1.4) for the calculation of the length of a three-dimensional vector.

Figure 1.6: The orientation of vectors in (a) two dimensions and (b) three dimensions in terms of angles with respect to Cartesian axes. The shaded regions indicate the right triangles from which the orientation angles are determined.
Vectors and Scalars

A more economical way of specifying the direction of a vector is obtained by keeping the total number of variables the same as the dimension of the vector. After all, in \( n \) dimensions, we have \( n \) degrees of freedom for specifying the vector in terms of \( n \) coordinates. No matter how we specify a vector we cannot exceed this number of degrees of freedom. The magnitude of a vector is one degree of freedom, leaving \( n - 1 \) degrees of freedom to specify the direction. Consider as an example, a two-dimensional vector \( \mathbf{a} = (a_1, a_2) \). The direction of this vector is given by an angle with respect to any fixed axis. By convention, this angle \( \phi \) is taken with respect to the positive \( x \)-axis, with positive angles measured in the counterclockwise direction. Thus, from Figs. 1.5(a) and 1.5(a), the length \( |\mathbf{a}| \) and angle \( \phi \) of \( \mathbf{a} = (a_1, a_2) \),

\[
|\mathbf{a}| = (a_1^2 + a_2^2)^{1/2}, \quad \cos \phi = \frac{a_1}{|\mathbf{a}|},
\]

completely specify this vector. This provides the motivation for defining a coordinate system in terms of the length and orientation of every point in the plane. Figure 1.7(a) shows the Cartesian coordinate system we have been using. Every point lies at the intersection of two grid lines that are parallel to the coordinate axes, which determines the coordinates of the point. In **circular polar coordinates** (Fig. 1.7(b)), every point lies at the intersection of a circle, which gives the radial coordinate of the point, and and one of the straight lines through the origin, which gives the orientation of the point.

In three dimensions we must specify two angles. The first is a angle \( \phi \) with respect to the positive \( x \)-axis, with positive angles measured in the counterclockwise direction when viewed from the positive \( z \)-axis. The second is an angle \( \theta \)

![Figure 1.7: Representation of the points in the plane in terms of (a) Cartesian coordinates and (b) circular polar coordinates.](image-url)
measured with respect to the $z$-axis. Elementary trigonometry yields

\[
|\mathbf{a}| = (a_1^2 + a_2^2 + a_3^2)^{\frac{1}{2}}, \quad \cos \phi = \frac{a_1}{|\mathbf{a}|}, \quad \cos \theta = \frac{a_3}{|\mathbf{a}|}.
\]

Here, too, a coordinate system can be defined in terms of a radial coordinate and two angular coordinates, which are known as **spherical polar coordinates**. The mathematics of using different coordinates are covered in the course on vector calculus.

### 1.5 Summary

The main result of this chapter are:

- The definitions of scalars and vectors.
- Cartesian coordinate systems in two and three dimensions
- The difference between right- and left-handed coordinate systems.
- The definition of a vector in terms of the coordinates of its tail and head.
- How to determine the length and direction of a vector in two and three dimensions