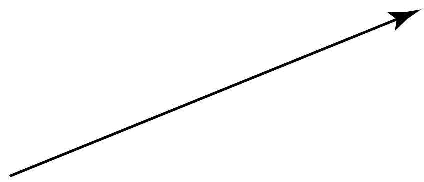
1
Vector algebra

1.1 Preliminaries

In introductory physics, we often deal with physical quantities that can be described by a single number. The temperature of a heated body, the mass of an object, and the electric potential of an insulated metal sphere are all examples of such *scalar* quantities.

Descriptions of physical phenomena are not always (indeed, rarely) that simple, however, and often we must use multiple, but related, numbers to offer a complete description of an effect. The next level of complexity is the introduction of *vector* quantities.

A vector may be described as a conceptual object having both *magnitude* and *direction*. Graphically, vectors can be represented by an arrow:



The length of the arrow is the *magnitude* of the vector, and the direction of the arrow indicates the *direction* of the vector.

Examples of vectors in elementary physics include displacement, velocity, force, momentum, and angular momentum, though the concept can be extended to more complicated and abstract systems. Algebraically, we will usually represent vectors by boldface characters, i.e. **F** for force, **v** for velocity, and so on.

It is worth noting at this point that the word "vector" is used in mathematics with somewhat broader meaning. In mathematics, a *vector space* is defined quite generally as a set of elements (called vectors) together with rules relating to their addition and scalar multiplication of vectors. In this sense, the set of real numbers form a vector space, as does any ordered set of numbers, including matrices, to be discussed in Chapter 4, and complex numbers, to be discussed in Chapter 9. For most of this chapter we reserve the term "vector" for quantities which possess magnitude and direction in three-dimensional space, and are independent of the specific choice of coordinate system in a manner to be discussed in

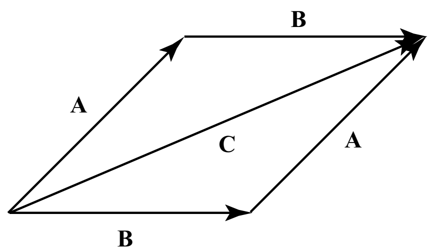


Figure 1.1 The parallelogram law of vector addition. Adding **B** to **A** (the addition above the **C**-line) is equivalent to adding **A** to **B** (the addition below the **C**-line).

Section 1.2. We briefly describe vector spaces at the end of this chapter, in Section 1.5. The interested reader can also consult Ref. [Kre78, Sec. 2.1].

Vector addition is *commutative* and *associative*; commutativity refers to the observation that the addition of vectors is order independent, i.e.

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A} = \mathbf{C}. \tag{1.1}$$

This can be depicted graphically by the parallelogram law of vector addition, illustrated in Fig. 1.1. A pair of vectors are added "tip-to-tail"; that is, the second vector is added to the first by putting its tail at the end of the tip of the first vector. The resultant vector is found by drawing an arrow from the origin of the first vector to the tip of the second vector. Associativity refers to the observation that the addition of multiple vectors is independent of the way the vectors are grouped for addition, i.e.

$$(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C}). \tag{1.2}$$

This may also be demonstrated graphically if we first define the following vector additions:

$$\mathbf{E} \equiv \mathbf{A} + \mathbf{B}, \tag{1.3}$$

$$\mathbf{D} \equiv \mathbf{E} + \mathbf{C}, \tag{1.4}$$

$$\mathbf{F} \equiv \mathbf{B} + \mathbf{C}. \tag{1.5}$$

The vectors and their additions are illustrated in Fig. 1.2. It can be immediately seen that

$$\mathbf{E} + \mathbf{C} = \mathbf{A} + \mathbf{F}. \tag{1.6}$$

So far, we have introduced vectors as purely geometrical objects which are independent of any specific coordinate system. Intuitively, this is an obvious requirement: where I am standing in a room (my "position vector") is independent of whether I choose to describe it by measuring it from the rear left corner of the room or the front right corner. In other words, the vector has a physical significance which does not change when I change my method of describing it.

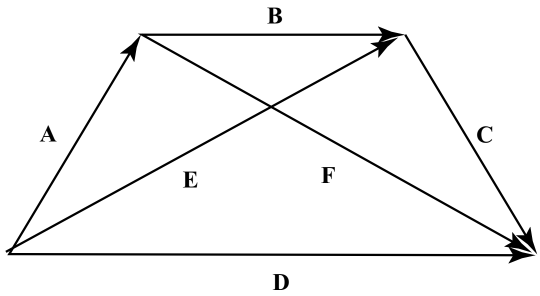


Figure 1.2 The trapezoid rule of vector addition. It makes no difference if we first add **A** and **B**, and then **C**, or first add **B** and **C**, and then **A**.

By choosing a coordinate system, however, we may create a *representation* of the vector in terms of these coordinates. We start by considering a Cartesian coordinate system with coordinates x, y, z which are all mutually perpendicular and form a right-handed coordinate system.[1] For a given Cartesian coordinate system, the vector **A**, which starts at the origin and ends at the point with coordinates (A_x, A_y, A_z), is completely described by the coordinates of the end point.

It is highly convenient to express a vector in terms of these components by use of unit vectors $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, $\hat{\mathbf{z}}$, vectors of unit magnitude pointing in the directions of the positive coordinate axes,

$$\mathbf{A} = A_x\hat{\mathbf{x}} + A_y\hat{\mathbf{y}} + A_z\hat{\mathbf{z}}. \tag{1.7}$$

This equation indicates that a vector equals the vector sum of its components. In three dimensions, the *position vector* **r** which measures the distance from a chosen origin is written as

$$\mathbf{r} = x\hat{\mathbf{x}} + y\hat{\mathbf{y}} + z\hat{\mathbf{z}}, \tag{1.8}$$

where x, y, and z are the lengths along the different coordinate axes.

The sum of two vectors can be found by taking the sum of their individual components. This means that the sum of two vectors **A** and **B** can be written as

$$\mathbf{A} + \mathbf{B} = (A_x + B_x)\hat{\mathbf{x}} + (A_y + B_y)\hat{\mathbf{y}} + (A_z + B_z)\hat{\mathbf{z}}. \tag{1.9}$$

The magnitude (length) of a vector in terms of its components can be found by two successive applications of the Pythagorean theorem. The magnitude A of the complete vector, also written as $|\mathbf{A}|$, is found to be

$$A = \sqrt{A_x^2 + A_y^2 + A_z^2}. \tag{1.10}$$

Another way to represent the vector in a particular coordinate system is by its magnitude A and the angles α, β, γ that the vector makes with each of the positive coordinate axes.

[1] If x is the outward-pointing index finger of the right hand, y is the folded-in ring finger and z is the thumb, pointing straight up.

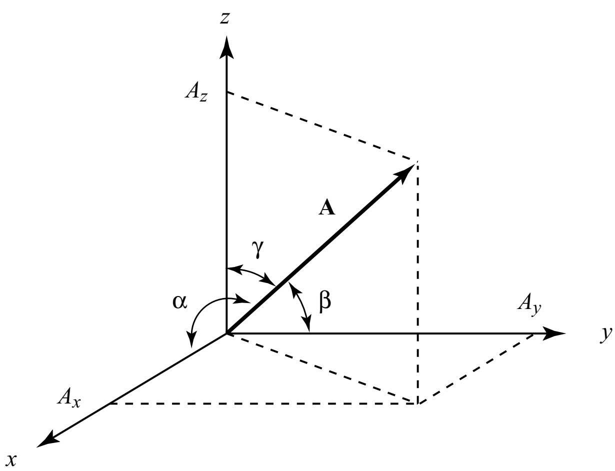


Figure 1.3 Illustration of the vector **A**, its components (A_x, A_y, A_z), and the angles α, β, γ.

These angles and their relationship to the vector and its components are illustrated in Fig. 1.3. The quantities $\cos\alpha$, $\cos\beta$, and $\cos\gamma$ are called *direction cosines*. It might seem that there is an inconsistency with this representation, since we now evidently need four numbers $(A, \alpha, \beta, \gamma)$ to describe the vector, where we needed only three (A_x, A_y, A_z) before. This seeming contradiction is resolved by the observation that α, β, and γ are not independent quantities; they are related by the equation,

$$\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1. \tag{1.11}$$

In the spherical coordinate system to be discussed in Chapter 3, we will see that we may completely specify the position vector by its magnitude r and two angles θ and ϕ.

It is to be noted that we usually see vectors in physics in two distinct classes:

1. Vectors associated with the property of a single, localized object, such as the velocity of a car, or the force of gravity acting on a moving projectile.
2. Vectors associated with the property of a nonlocalized "object" or system, such as the electric field of a light wave, or the velocity of a fluid. In such a case, the vector quantity is a function of position and possibly time and we may do calculus with respect to the position and time variables. This vector quantity is usually referred to as a *vector field*.

Vector fields are extremely important quantities in physics and we will return to them often.

1.2 Coordinate system invariance

We have said that a vector is independent of any specific coordinate system – in other words, that a vector is independent of how we choose to characterize it. This seems like an obvious criterion, but there are physical quantities which have magnitude and direction but are not vectors; an example of this in optics is the set of principle indices of refraction of an anisotropic crystal. Thus, to define a vector properly, we need to formulate mathematically this concept of *coordinate system invariance*. Furthermore, it is not uncommon to require, in the solution of a physical problem, the transformation from one coordinate system to another. We therefore take some time to study the mathematics relating to the behavior of a vector under a change of coordinates.

The simplest coordinate transformation is a change of origin, leaving the orientation of the axes unchanged. The only vector that depends explicitly upon the origin is the position vector $\mathbf{r}$, which is a measure of the vector distance from the origin. If the new origin of a new coordinate system, described by position vector $\mathbf{r}'$, is located at the position $\mathbf{r}_0$ from the old origin, the coordinates are related by the formula

$$\mathbf{r}' = \mathbf{r} - \mathbf{r}_0. \tag{1.12}$$

Most other basic vectors depend upon the *displacement vector* $\mathbf{R} = \mathbf{r}_2 - \mathbf{r}_1$, i.e. the change in position, and therefore are unaffected by a change in origin. Examples include the velocity, momentum, and force upon an object.

A less trivial example of a change of coordinate system is a change of the orientation of coordinate axes, and its effect on a position vector $\mathbf{r}$. For simplicity, we first consider the two dimensional case. The vector $\mathbf{r}$ may be written in one coordinate system as $\mathbf{r} = x\hat{\mathbf{x}} + y\hat{\mathbf{y}}$, while in a second coordinate system this vector may be written as $\mathbf{r}' = x'\hat{\mathbf{x}}' + y'\hat{\mathbf{y}}'$. The (x, y) coordinate axes are rotated to a new location to become the (x', y') axes, while leaving the vector $\mathbf{r}$ (in particular, the location of the tip of $\mathbf{r}$) fixed. The question we ask: what are the components of the vector $\mathbf{r}$ in the new coordinate system, which makes an angle ϕ with the old system? The relation between the two systems is illustrated in Fig. 1.4.

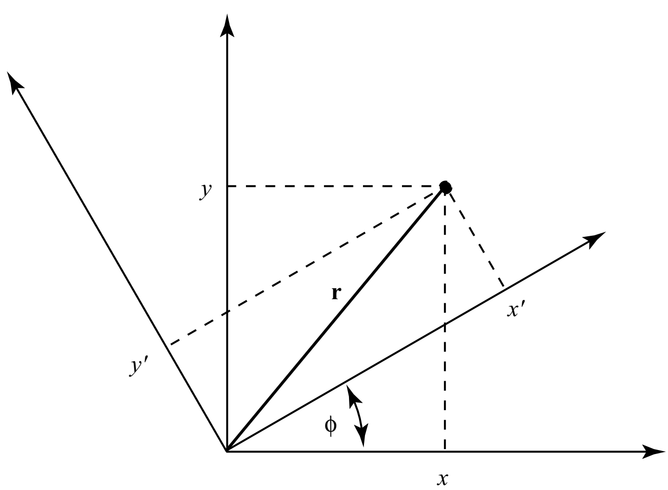


Figure 1.4 Illustration of the position vector **r** and its components in the two coordinate systems.

By straightforward trigonometry, one can readily find that the new coordinates of the vector (x', y') may be written in terms of the old coordinates as

$$x' = x\cos\phi + y\sin\phi, \tag{1.13}$$

$$y' = -x\sin\phi + y\cos\phi. \tag{1.14}$$

These equations are based on the assumption that the magnitude and direction of the vector is independent of the coordinate system, and this assumption should hold for anything we refer to as a vector. We therefore *define* a vector as a quantity whose components transform under rotations just as the position vector **r** does, i.e. a vector **A** with components A_x and A_y in the unprimed system should have components

$$A'_x = A_x\cos\phi + A_y\sin\phi, \tag{1.15}$$

$$A'_y = -A_x\sin\phi + A_y\cos\phi \tag{1.16}$$

in the primed system.

It is important to emphasize again that we are only rotating the coordinate axes, and that the vector **A** does not change: (A_x, A_y) and (A'_x, A'_y) are *representations* of the vector, different ways of describing the same physical property. Indeed, another way to define a vector is that it is a quantity with magnitude and direction that is independent of the coordinate system.

We can also interpret Eqs. (1.15) and (1.16) in an entirely different manner: if we were to physically rotate the vector **A** over an angle $-\phi$ about the origin of the coordinate system, the new direction of the vector in the *same* coordinate system would be given by (A'_x, A'_y). A rotation of the coordinate system in the direction ϕ is mathematically equivalent to a rotation of the vector by an angle $-\phi$.

To generalize the discussion to three (or more) dimensions, it helps to modify the notation somewhat. We write

$$x \to x_1, \tag{1.17}$$

$$y \to x_2, \tag{1.18}$$

and define

$$\begin{aligned} a_{11} &= \cos\phi, \\ a_{12} &= \sin\phi = \cos(\pi/2 - \phi) = \cos(\phi - \pi/2), \\ a_{21} &= -\sin\phi = -a_{12} = \cos(\phi + \pi/2), \\ a_{22} &= \cos\phi. \end{aligned} \tag{1.19}$$

With these definitions, our formulas for a coordinate transformation become

$$x'_1 = a_{11}x_1 + a_{12}x_2, \tag{1.20}$$

$$x'_2 = a_{21}x_1 + a_{22}x_2. \tag{1.21}$$

These transformations may be written in a summation format,

$$x_i' = \sum_{j=1}^{2} a_{ij} x_j, \quad i = 1, 2, \tag{1.22}$$

where x_i' is the ith component of the vector in the primed frame, x_j is the jth component of the vector in the unprimed frame, and the notation $\sum_{j=n}^{m}$ indicates summation over all terms with index j ranging from n to m. The quantity a_{ij} can be seen from Eqs. (1.19) to be the *direction cosine* with the respect to the ith primed coordinate and the jth unprimed coordinate.

What happens if we run the rotation in reverse? We can still use Eq. (1.21), but we replace ϕ by $-\phi$ and switch the primed and unprimed coordinates, i.e. the primed coordinates are now the start of the rotation and the unprimed coordinates are now the end of the rotation. With these changes, Eq. (1.21) becomes

$$x_1 = a_{11} x_1' - a_{12} x_2', \tag{1.23}$$

$$x_2 = -a_{21} x_1' + a_{22} x_2'. \tag{1.24}$$

Noting that $a_{12} = -a_{21}$, these formulas can be rewritten in the compact summation form,

$$x_j = \sum_{i=1}^{2} a_{ij} x_i', \quad j = 1, 2. \tag{1.25}$$

Generalizing to N dimensions may now be done by analogy, simply introducing higher-order direction cosines a_{ij} and an N-dimensional position vector $\mathbf{r} = (x_1, \ldots, x_N)$, which will satisfy the relations

$$x_i' = \sum_{j=1}^{N} a_{ij} x_j, \quad i = 1, \ldots, N. \tag{1.26}$$

The a_{ij}s may be written in a differential form with respect to the two coordinate systems as

$$a_{ij} = \frac{\partial x_i'}{\partial x_j}. \tag{1.27}$$

This formula can be derived by taking partial derivatives of the transformation equations for $\mathbf{r}$, namely Eq. (1.26). The quantity a_{ij} has more components than a vector and will be seen to be a *tensor*, which can be represented in matrix form; we will discuss such beasts in Chapter 5.

The reverse rotation may also be written by analogy,

$$x_j = \sum_{i=1}^{N} a_{ij} x_i', \tag{1.28}$$

and from this expression we may also write

$$a_{ij} = \frac{\partial x_j}{\partial x'_i}. \tag{1.29}$$

It is evident that the coordinates of a vector must in the end be unchanged if the axes are first rotated and then rotated back to their original positions; from this we can derive an *orthogonality condition* for the coefficients a_{ij}. We begin with the transformation of the vector **V** and its reverse,

$$V_k = \sum_{i=1}^{N} a_{ik} V'_i, \tag{1.30}$$

$$V'_i = \sum_{j=1}^{N} a_{ij} V_j. \tag{1.31}$$

On substitution of the latter equation into the former, we have

$$V_k = \sum_{i=1}^{N} a_{ik} \left[\sum_{j=1}^{N} a_{ij} V_j \right] = \sum_{j=1}^{N} \left[\sum_{i=1}^{N} a_{ik} a_{ij} \right] V_j. \tag{1.32}$$

The left-hand side of this equation is the kth component of the vector in the unprimed frame. The right-hand side of the equation is a weighted sum of all components of the vector in the unprimed frame. By the use of Eqs. (1.27) and (1.29), we may write the quantity in square brackets as

$$\sum_{i=1}^{N} a_{ik} a_{ij} = \sum_{i=1}^{N} \frac{\partial x_k}{\partial x'_i} \frac{\partial x'_i}{\partial x_j} = \frac{\partial x_k}{\partial x_j}, \tag{1.33}$$

where the final step is the result of application of the chain rule of calculus. Because the variables x_j, for $j = 1, \ldots, N$, are independent of one another, we readily find that

$$\sum_{i=1}^{N} a_{ik} a_{ij} = \delta_{jk}, \tag{1.34}$$

where δ_{jk} is the *Kronecker delta*, defined as

$$\delta_{jk} = \begin{cases} 1 & \text{for} \quad j = k, \\ 0 & \text{for} \quad j \neq k. \end{cases} \tag{1.35}$$

We will see a lot of the Kronecker delta in the future – remember it! It is another example of a *tensor*, like the rotation tensor a_{ij}.

1.3 Vector multiplication

In Section 1.1, we looked at the addition of vectors, which may be considered a generalization of the addition of ordinary numbers. One can envision that there exists a generalized form of multiplication for vectors, as well; with three components for each vector, however, there are a large number of possibilities for what we might call "vector multiplication". Just as vectors themselves are invariant under a change of coordinates, any vector multiplication should also be invariant under a change of coordinates. It turns out that for three-dimensional vectors there exist four possibilities, three of which we discuss here.[2]

1.3.1 Multiplication by a scalar

The simplest form of multiplication involving vectors is the multiplication of a vector by a scalar. The effect of such a multiplication is the "scaling" of each component of the vector equally by the scalar α, i.e.

$$\alpha\mathbf{V} = \alpha(V_x\hat{\mathbf{x}} + V_y\hat{\mathbf{y}} + V_z\hat{\mathbf{z}}) = (\alpha V_x)\hat{\mathbf{x}} + (\alpha V_y)\hat{\mathbf{y}} + (\alpha V_z)\hat{\mathbf{z}}. \tag{1.36}$$

It is clear from the above that the act of multiplying by a scalar does not change the direction of the vector, but only scales its length by the factor α; we may formally write $|\alpha\mathbf{V}| = |\alpha||\mathbf{V}|$. The result of the multiplication is also a vector, as it is clear that this product is invariant under a rotation of the coordinate axes, which does not affect the vector length.

It is to be noted that we may also consider scalar multiplication "backwards", i.e. that it represents the multiplication of a scalar by a vector, with the end result being a vector. This interpretation will be employed in Chapter 2 to help categorize the different types of vector differentiation.

1.3.2 Scalar or dot product

The *scalar product* (or "dot product") between two vectors is represented by a dot and is defined as

$$\mathbf{A}\cdot\mathbf{B} \equiv AB\cos\theta, \tag{1.37}$$

where A and B are the magnitudes of vectors $\mathbf{A}$ and $\mathbf{B}$, respectively, and θ is the angle between the two vectors.

The rotational invariance of this quantity is almost obvious from its definition, for we know that the magnitudes of vectors and the angles between any two vectors are all unchanged under rotations. We will confirm this more rigorously in a moment.

In terms of components in a particular Cartesian coordinate system, the scalar product is given by[3]

$$\mathbf{A}\cdot\mathbf{B} = \sum_i A_iB_i = \sum_i B_iA_i = \mathbf{B}\cdot\mathbf{A}. \tag{1.38}$$

[2] A discussion of the fourth, the direct product, will be deferred until Section 5.6.
[3] From now on, we no longer write the upper and lower ranges of the summations.

We can use this representation of the scalar product to rigorously prove that it is invariant under rotations. We start with the scalar product in the primed coordinate system, and substitute into it the representation of the primed vectors in terms of the unprimed coordinates, Eq. (1.31),

$$\sum_i A_i' B_i' = \sum_i \left[\sum_j a_{ij} A_j \sum_k a_{ik} B_k \right] = \sum_j \sum_k \left[\sum_i a_{ij} a_{ik} \right] A_j B_k. \tag{1.39}$$

The expression in the last brackets is simply the orthogonality relation between the direction cosines, Eq. (1.34), and may be set to δ_{jk}. We thus have

$$\sum_i A_i' B_i' = \sum_i A_i B_i \tag{1.40}$$

and we have proven that the scalar product is invariant under rotations.

The dot product may be used to demonstrate another familiar geometrical formula, the *law of cosines*. Defining

$$\mathbf{C} = \mathbf{A} + \mathbf{B}, \tag{1.41}$$

it follows from the parallelogram law of vector addition that **A**, **B**, and **C** form the sides of a triangle. If we take the dot product of **C** with itself,

$$\mathbf{C} \cdot \mathbf{C} = (\mathbf{A} + \mathbf{B}) \cdot (\mathbf{A} + \mathbf{B}) = \mathbf{A} \cdot \mathbf{A} + \mathbf{B} \cdot \mathbf{B} + 2\mathbf{A} \cdot \mathbf{B}. \tag{1.42}$$

The dot product of a vector with itself is simply the squared magnitude of the vector, and the dot product of **A** and **B** is defined by Eq. (1.37). We thus arrive at

$$C^2 = A^2 + B^2 + 2AB\cos\theta, \tag{1.43}$$

the law of cosines.

1.3.3 Vector or cross product

The third form of rotationally invariant product involving vectors is the *vector product* or "cross product". Just as the scalar product is named such because the result of the product is a scalar, the result of the vector product is another vector. It is represented by a cross between vectors,

$$\mathbf{C} = \mathbf{A} \times \mathbf{B}. \tag{1.44}$$

In evident contrast with the scalar product, the magnitude of the vector product is defined as

$$C = AB\sin\theta, \tag{1.45}$$

where θ is again the angle between the vectors **A** and **B**. From this definition, it is to be noted that the magnitude of **C** is the area of the parallelogram formed by **A** and **B**, and that