Chapter 8

Linear Functions and Matrices

Functions (also known as ‘mappings,’ or ‘maps’) between Euclidean spaces are ubiquitous in physics, engineering, and mathematics. Examples abound. Scalar functions of a single variable are the subject of differential and integral calculus, while scalar and vector functions of variables in $\mathbb{R}^3$ are studied in vector calculus, with applications in mechanics, electromagnetism, fluid mechanics, and elasticity.

Among the most important functions are **linear** functions. A precise definition will be given later in this chapter, but, for the moment, we can think of lines and planes that pass through the origin as prototypical linear functions. Though less general than their non-linear counterparts, linear functions are no less important because many non-linear functions benefit from a locally linear analysis. The derivative (Fig. 3.1) provides an example of a linear approximation near a point of any function (where the derivative exists). Higher-dimensional analogues of this construction lead to the notion of locally linear functions (planes, hyperplanes) near any point of a smooth surface or hypersurface. Changes of integration variables between multivariable Cartesian coordinates and other coordinate systems involve the Jacobian, which determines how the volume element (an infinitesimal n-dimensional hypercube) changes under the variable transformation.

In this chapter, we formalize the definition of a function between Euclidean spaces. We then specialize our discussion to linear functions and develop their key properties. Examples are used illustrate the various concepts associated with functions. The main theorem in this chapter establishes the equivalence between linear functions and matrices. Of these, square matrices are the most common because they are used for operations that map $\mathbb{R}^n$ onto itself, such as rotations and other rigid-body transformations used throughout physics, but especially in classical and quantum mechanics. We conclude with several examples of common linear transformations and their matrix realizations.

8.1 Functions between Euclidean Spaces

Recall that a function is a relation between two sets: an input set and a set of allowed outputs with the property that each element of the input set is related to one and only one element of the output set. The following definition provides the standard nomenclature for the input and output sets:

DEFINITION 8.1. The **domain** of a function f is the set over which it is defined, The **range** of f is the set of values taken by f over its domain.

The notation $f\colon \mathbb{R}^n \to \mathbb{R}^n$ indicates a function f whose domain is a (proper or improper) subset of $\mathbb{R}^n$ and whose range is a (proper or improper) subset of $\mathbb{R}^m$.

EXAMPLE 8.1. A function $f\colon \mathbb{R} \to \mathbb{R}$ is represented by a curve in the x-y plane, with the input along the x-axis and the output along the y-axis. Figure 8.1 shows three such functions that illustrate three domains and ranges that commonly arise. Figure 8.1(a) shows the function $y = x^3$, for which the domain and range are both the entire real line $\mathbb{R}$. For the function $y = \tanh x$, shown in Fig. 8.1(b), the domain is the entire real line $\mathbb{R}$, but the range is the subset $(-1, 1)$ of $\mathbb{R}$. The range is an open interval because the limit points (± 1) of the interval are attained only asymptotically as $x \to \pm\infty$. Finally, Fig. 8.1(c) shows the function $y = \sqrt{1-x^2}$, for which the domain is the subset $[-1, 1]$ and the range the subset $[0, 1]$ of $\mathbb{R}$. ▮

EXAMPLE 8.2. A function $f\colon \mathbb{R}^2 \to \mathbb{R}$ is represented as a surface in $\mathbb{R}^3$. Figure 8.2(a) shows the function $f = x^2 + y^2$, whose surface is a paraboloid (a surface generated by rotating a parabola about its axis of symmetry). The domain is all of $\mathbb{R}^2$, which is represented by the x-y plane, and the range, which is plotted along the z-axis, is the non-negative real line $[0, \infty)$. Functions $f\colon \mathbb{R} \to \mathbb{R}^3$ have the

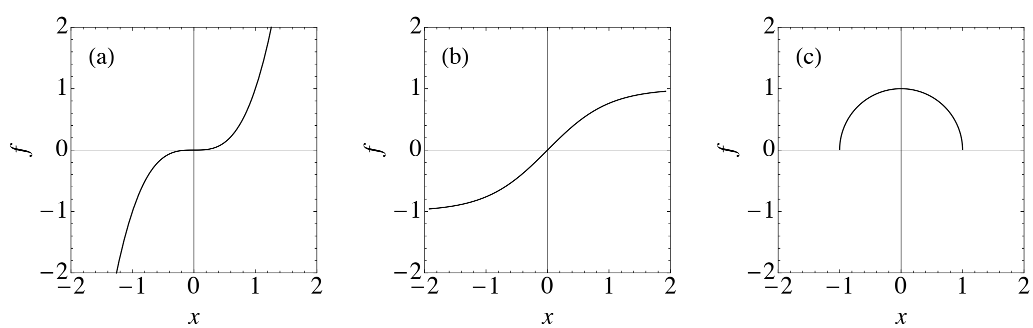

Figure 8.1: The functions (a) $y = x^3$, (b) $y = \tanh x$, and (c) $y = \sqrt{1-x^2}$. The domain and range in (a) are both $\mathbb{R}$, in (b) the domain is $\mathbb{R}$ and the range is $(-1, 1)$, and in (c), the domain is $[-1, 1]$ and the range $[0, 1]$.

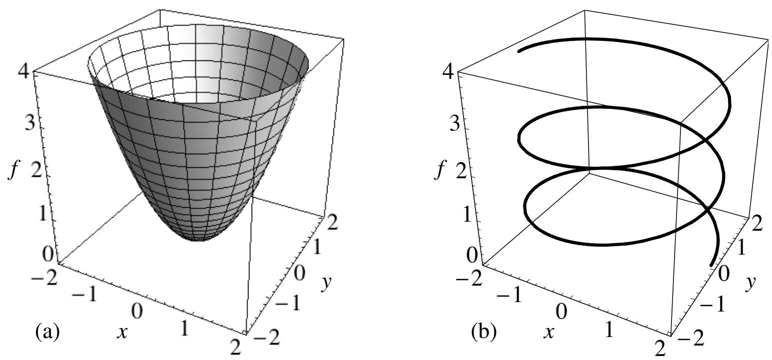

Figure 8.2: The functions (a) $f = x^2 + y^2$ and (b) $f = (2\cos 4t, 2\sin 4t, t)$. The domain of (a) is $\mathbb{R}^2$ and the range is the subset $[0,\infty)$ of $\mathbb{R}$. For (b), the domain is the subset $[0,\infty)$ of $\mathbb{R}$, and the range is the subset $[-2,2]\times[-2,2]\times[0,\infty)$ of $\mathbb{R}^3$.

form $(x(t), y(t), z(t))$, which is the parametric form of a curve in $\mathbb{R}^3$. Such functions occur in mechanics as particle trajectories. Figure 8.2(b) shows the function $f = (2\cos 4t, 2\sin 4t, t)$ for $t \geq 0$, which is the domain of the function. The range is the subset $[-2,2]\times[-2,2]\times[0,\infty)$ of $\mathbb{R}^3$. The spiral is the type of trajectory seen for a charged particle in a magnetic field (cf. Example 3.5). ∎

8.2 Linear and Affine Functions

A central theme of differential calculus is the approximation of nonlinear functions by linear functions determined by the derivative of the nonlinear function. Thus, as discussed in the introduction, linear functions have a far broader range of applications than their properties might suggest. This section formalizes the definition linear and affine functions, which provides the basis for establishing a direct connection to matrices. We begin with the definition of a linear function:

DEFINITION 8.2. A function $f\colon \mathbb{R}^n \to \mathbb{R}^n$ is **linear** if, for any vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$,

$$f(\boldsymbol{x}+\boldsymbol{y}) = f(\boldsymbol{x}) + f(\boldsymbol{y}), \tag{8.1}$$

and, for any $\boldsymbol{x} \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$,

$$f(\lambda\boldsymbol{x}) = \lambda f(\boldsymbol{x}). \tag{8.2}$$

These conditions can be subsumed by a single requirement for linearity:

$$f(\lambda\boldsymbol{x}+\mu\boldsymbol{y}) = \lambda f(\boldsymbol{x}) + \mu f(\boldsymbol{y}), \tag{8.3}$$

for any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and $\lambda, \mu \in \mathbb{R}$.

EXAMPLE 8.3. Consider the function $f = ax$, which is the equation of a straight line through the origin with slope a. To test for linearity, we write

$$f(\lambda x + \mu y) = a(\lambda x + \mu y) = \lambda(ax) + \mu(ay) = \lambda f(x) + \mu f(y), \tag{8.4}$$

which, according (8.3) fulfills the condition for linearity. Now consider the function $f = ax + b$, which is a straight line with slope a and y-intercept b. To test for linearity, we again write

$$f(\lambda x + \mu y) = a(\lambda x + \mu y) + b = \lambda(ax) + \mu(ay) + b, \tag{8.5}$$

which, because of the additive factor of b, does not equal $\lambda f(x) + \mu f(y)$, so this function is not linear. In fact, such functions are called **affine** because they are a combination of a linear transformation (in this case ax) followed by a translation (in this case, b). Linear and affine functions share several properties, as we will show below, but linear transformations have additional special properties.

Finally, consider the equation of a plane (4.20): $Ax + By + Cz = D$. Rearranging this equation into a form where z is expressed as a function of x and y:

$$z(x,y) = \frac{D}{C} - \frac{A}{C}x - \frac{B}{C}y, \tag{8.6}$$

shows that z: $\mathbb{R}^2 \to \mathbb{R}$. Thus, by defining $\boldsymbol{r} = (x,y)$ and $\boldsymbol{r}' = (x',y')$, we test for linearity by writing

$$\begin{aligned} z(\lambda\boldsymbol{r} + \mu\boldsymbol{r}') &= \frac{D}{C} - \frac{A}{C}(\lambda x + \mu x') - \frac{B}{C}(\lambda y + \mu y') \\ &= \frac{D}{C} - \lambda\left(\frac{A}{C}x + \frac{B}{C}y\right) - \mu\left(\frac{A}{C}x' + \frac{B}{C}y'\right), \end{aligned} \tag{8.7}$$

which shows that z is a linear function only if $D = 0$; otherwise, it is affine. ∎

There are several useful properties of linear functions, some of which are shared by affine functions. The following theorems prove several such properties.

THEOREM 8.1. A linear function f: $\mathbb{R}^n \to \mathbb{R}^m$ maps the origin of $\mathbb{R}^n$ to the origin of $\mathbb{R}^m$.

Proof. We invoke (8.3) with $\lambda\boldsymbol{x} + \mu\boldsymbol{y} = \boldsymbol{0}$. Then, by using the properties of linear functions, we have, with $\mu\boldsymbol{y} = -\lambda\boldsymbol{x}$,

$$\begin{aligned} f(\boldsymbol{0}) &= f(\lambda\boldsymbol{x} + \mu\boldsymbol{y}) = f(\lambda\boldsymbol{x}) + f(\mu\boldsymbol{y}) \\ &= f(\lambda\boldsymbol{x}) + f(-\lambda\boldsymbol{x}) = \lambda f(\boldsymbol{x}) - \lambda f(\boldsymbol{x}) = \boldsymbol{0}. \end{aligned} \tag{8.8}$$

Affine functions clearly do not share this property because the translational part of the function translates the origin in $\mathbb{R}^m$. □

THEOREM 8.2. A linear function $f\colon \mathbb{R}^n \to \mathbb{R}^m$ maps straight lines in $\mathbb{R}^n$ to straight lines $\mathbb{R}^m$.

Proof. From (4.1), the equation of a straight line in $\mathbb{R}^n$ is

$$\boldsymbol{x} = \boldsymbol{x}_0 + \lambda \boldsymbol{d}, \tag{8.9}$$

in which $\boldsymbol{x}$ is any point on the line, $\boldsymbol{x}_0$ is a reference point on the line, $\lambda \in \mathbb{R}$, and $\boldsymbol{d}$ is the direction of the line. Thus, using the fact that f is a linear function,

$$f(\boldsymbol{x}_0 + \lambda \boldsymbol{d}) = f(\boldsymbol{x}_0) + f(\lambda \boldsymbol{d}) = f(\boldsymbol{x}_0) + \lambda f(\boldsymbol{d}), \tag{8.10}$$

which is the equation of a line in $\mathbb{R}^m$, with reference point $f(\boldsymbol{x}_0)$ and direction vector $f(\boldsymbol{d})$. □

Building on the discussion in Example 8.3, the general form of an affine function $g\colon \mathbb{R}^n \to \mathbb{R}^m$ is $g = f + \boldsymbol{b}$, where f is a linear function and $\boldsymbol{b}$ is a vector in $\mathbb{R}^m$. Thus, with (8.9) as input and using (8.10), the output of an affine function is

$$g(\boldsymbol{x}_0 + \lambda \boldsymbol{d}) = f(\boldsymbol{x}_0 + \lambda \boldsymbol{d}) + \boldsymbol{b} = f(\boldsymbol{x}_0) + \boldsymbol{b} + \lambda f(\boldsymbol{d}), \tag{8.11}$$

which is a straight line with reference point $f(\boldsymbol{x}_0) + \boldsymbol{b}$ and direction vector $f(\boldsymbol{d})$. Using analogous steps, we can also show that linear and affine functions both map parallel lines in $\mathbb{R}^n$ to parallel lines in $\mathbb{R}^m$.

8.3 Matrices Associated with Linear Functions

Several general properties of linear functions were discussed in the preceding section. There are other properties for particular types of linear functions. A systematic development of such properties relies on the association of linear transformations with matrices. The basis of this association is the following theorem:

THEOREM 8.3. A function $f\colon \mathbb{R}^n \to \mathbb{R}^m$ is linear if and only if there exists an $m \times n$ matrix A such that $f = \mathsf{A}\boldsymbol{x}$, for all $\boldsymbol{x} \in \mathbb{R}^n$.

Proof. We first suppose that $f = \mathsf{A}\boldsymbol{x}$ and show that the linearity of f follows. For any two vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and $\lambda, \mu \in \mathbb{R}$, the property (7.28) of matrices yields

$$f(\lambda \boldsymbol{x} + \mu \boldsymbol{y}) = \mathsf{A}(\lambda \boldsymbol{x} + \mu \boldsymbol{y}) = \lambda \mathsf{A}\boldsymbol{x} + \mu \mathsf{A}\boldsymbol{y} = \lambda f(\boldsymbol{x}) + \mu f(\boldsymbol{y}), \tag{8.12}$$

Thus, according to (8.3), f is a linear function.

Now suppose that $f\colon \mathbb{R}^n \to \mathbb{R}^m$ is a linear function. We must show that $f]\mathsf{A}\boldsymbol{x}$ for all $\boldsymbol{x} \in \mathbb{R}^n$. We begin by expressing all vectors in $\mathbb{R}^n$ in terms of the standard basis. These vectors, denoted as $\boldsymbol{e}_i$, for $i = 1, 2, \ldots n$ are column vectors with a 1 in the ith row and zero in all other rows:

$$\boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad \boldsymbol{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \quad \boldsymbol{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}. \tag{8.13}$$

This basis is the natural extension to $\mathbb{R}^n$ of the familiar unit vectors $\boldsymbol{i} \equiv \boldsymbol{e}_1$, $\boldsymbol{j} \equiv \boldsymbol{e}_2$, and $\boldsymbol{k} \equiv \boldsymbol{e}_3$ in $\mathbb{R}^3$:

$$\boldsymbol{i} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \boldsymbol{j} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad \boldsymbol{k} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \tag{8.14}$$

in terms of which any vectors in $\mathbb{R}^3$ can be expressed: $\boldsymbol{x} = x_1\boldsymbol{i} + x_2\boldsymbol{j} + x_3\boldsymbol{k}$. Similarly, for any $\boldsymbol{x} \in \mathbb{R}^n$,

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + \cdots + x_n\boldsymbol{e}_n\,. \tag{8.15}$$

Hence, by using the two conditions for linearity in (8.1) and (8.2), we have,

$$\begin{aligned} f(\boldsymbol{x}) &= f(x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + \cdots + x_n\boldsymbol{e}_n) \\ &= f(x_1\boldsymbol{e}_1) + f(x_1\boldsymbol{e}_2) + \cdots + f(x_n\boldsymbol{e}_n) \\ &= x_1 f(\boldsymbol{e}_1) + x_2 f(\boldsymbol{e}_2) + \cdots + x_n f(\boldsymbol{e}_n)\,. \end{aligned} \tag{8.16}$$

As f maps vectors in $\mathbb{R}^n$ to vectors in $\mathbb{R}^m$, $f(\boldsymbol{e}_i)$ is an element of $\mathbb{R}^m$, which is an m-dimensional column vector. We denote this vector as $\boldsymbol{a}_i$ and its components as

$$\boldsymbol{a}_i = \begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{m1} \end{pmatrix}. \tag{8.17}$$

Combining this result with (8.16) yields

$$\begin{aligned} f(\boldsymbol{x}) &= x_1 f(\boldsymbol{e}_1) + x_2 f(\boldsymbol{e}_2) + \cdots + x_n f(\boldsymbol{e}_n) \\ &= x_1 \boldsymbol{a}_1 + x_2 \boldsymbol{a}_2 + \cdots + x_n \boldsymbol{a}_n \\ &= \begin{pmatrix} x_1 a_{11} + x_2 a_{12} + \cdots + x_n a_{1n} \\ x_1 a_{21} + x_2 a_{22} + \cdots + x_n a_{2n} \\ \vdots \quad\quad \vdots \quad\quad\quad \vdots \\ x_1 a_{m1} + x_2 a_{m2} + \cdots + x_n a_{mn} \end{pmatrix} \\ &= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \equiv \mathsf{A}\boldsymbol{x}, \end{aligned} \tag{8.18}$$

which established the equivalence between linear functions $f\colon \mathbb{R}^n \to \mathbb{R}^m$ and $m \times n$ matrices. □

We will explore some of the consequences of Theorem 8.3 once we examine some of the implications of this theorem for the algebraic structure of matrices.

8.4 Linear Functions and Matrices[1]

Suppose we have two linear functions $f\colon \mathbb{R}^n \to \mathbb{R}^n$ and $g\colon \mathbb{R}^n \to \mathbb{R}^n$. The sum, $f(\boldsymbol{x}) + g(\boldsymbol{x})$, multiplication by a scalar $\lambda \in \mathbb{R}$, $\lambda f(\boldsymbol{x})$, and composition $(f \circ g)(\boldsymbol{x})$ are well-defined operations. According to Theorem 8.3, there are $n \times n$ matrices A and B such that $f(\boldsymbol{x}) = \mathsf{A}\boldsymbol{x}$ and $g(\boldsymbol{x}) = \mathsf{B}\boldsymbol{x}$ for all $\boldsymbol{x} \in \mathbb{R}^n$. The transcription of the foregoing function operations into matrix forms are:

$$\begin{aligned} f(\boldsymbol{x}) + g(\boldsymbol{x}) &= \mathsf{A}\boldsymbol{x} + \mathsf{B}\boldsymbol{x} = (\mathsf{A} + \mathsf{B})\boldsymbol{x}, \\ \lambda f(\boldsymbol{x}) &= \lambda(\mathsf{A}\boldsymbol{x}) = (\lambda \mathsf{A})\boldsymbol{x}, \\ (f \circ g)(\boldsymbol{x}) &= f(g(\boldsymbol{x})) = \mathsf{A}(\mathsf{B}\boldsymbol{x}) = (\mathsf{AB})\boldsymbol{x}, \end{aligned} \tag{8.19}$$

which introduces the sum of two matrices, the multiplication of matrices by a real constant, and the multiplication of two matrices. We consider each of these operations in the following sections. Although each operation is valid for square matrices, rectangular matrices (with certain restrictions for multiplication) may also be used.

[1] For information only; not examinable.

8.4.1 Matrix Addition

Addition is defined for any matrices that have the same numbers of rows and columns. For $m \times n$ matrices $\mathsf{A} = (a_{ij})$ and $\mathsf{B} = (b_{ij})$, where $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$, the sum $\mathsf{A} + \mathsf{B}$ is defined as the addition of entries which are in the same position in the two matrices, that is, elements which have the same values of i and j:

$$\begin{aligned} \mathsf{A}+\mathsf{B} &= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix} \\ &= \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{pmatrix}. \end{aligned} \tag{8.20}$$

This sum can be written in abbreviated form as $\mathsf{A} + \mathsf{B} = (a_{ij} + b_{ij})$ once the numbers of rows and columns are specified. This definition makes apparent that A and B must have the same number of rows and columns, that is, both must be $m \times n$ matrices, for their sum to be well-defined.

EXAMPLE 8.4. Consider the sum of the two matrices

$$\mathsf{A} = \begin{pmatrix} 2 & 5 \\ -1 & 6 \end{pmatrix}, \qquad \mathsf{B} = \begin{pmatrix} 1 & -1 \\ 3 & 9 \end{pmatrix}. \tag{8.21}$$

These matrices have the same numbers of rows and columns, so, from (8.20), their sum is calculated as

$$\begin{aligned} \mathsf{A}+\mathsf{B} &= \begin{pmatrix} 2 & 5 \\ -1 & 6 \end{pmatrix} + \begin{pmatrix} 1 & -1 \\ 3 & 9 \end{pmatrix} \\ &= \begin{pmatrix} 2+1 & 5+(-1) \\ -1+3 & 6+9 \end{pmatrix} \qquad (8.22) \\ &= \begin{pmatrix} 3 & 4 \\ 3 & 15 \end{pmatrix}. \qquad (8.23) \end{aligned}$$

∎

8.4.2 Multiplication of a Matrix by a Real Scalar

Suppose $\mathsf{A} = (a_{ij})$ is an $m \times n$ matrix. Then the matrix $\lambda\mathsf{A}$, where $\lambda \in \mathbb{R}$ is the matrix $\lambda\mathsf{A} = (\lambda a_{ij})$, where each element is multiplied by λ. In matrix form,

$$\lambda\mathsf{A} = \lambda \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \cdots & \lambda a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \cdots & \lambda a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda a_{m1} & \lambda a_{m2} & \cdots & \lambda a_{mn} \end{pmatrix}. \tag{8.24}$$

In the special case that $\lambda = -1$, $(-1)\mathsf{A} = (-a_{ij})$, which enables us to define the difference $\mathsf{A} - \mathsf{B}$ between matrices A and B as $\mathsf{A} - \mathsf{B} = (a_{ij} - b_{ij})$.

EXAMPLE 8.5. Consider the matrices

$$\mathsf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ -1 & -2 & -3 \end{pmatrix}, \qquad \mathsf{B} = \begin{pmatrix} 2 & 4 & 6 \\ -1 & 3 & 5 \\ 4 & -2 & -3 \end{pmatrix}. \tag{8.25}$$

Then,

$$\begin{aligned} 3\mathsf{A} = 3 \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ -1 & -2 & -3 \end{pmatrix} &= \begin{pmatrix} 3\times 1 & 3\times 2 & 3\times 3 \\ 3\times 4 & 3\times 5 & 3\times 6 \\ 3\times(-1) & 3\times(-2) & 3\times(-3) \end{pmatrix} \\ &= \begin{pmatrix} 3 & 6 & 0 \\ 12 & 15 & 18 \\ -3 & -6 & -9 \end{pmatrix}, \end{aligned} \tag{8.26}$$

and

$$\mathsf{A} - \mathsf{B} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ -1 & -2 & -3 \end{pmatrix} - \begin{pmatrix} 2 & 4 & 6 \\ -1 & 3 & 5 \\ 4 & -2 & -3 \end{pmatrix} = \begin{pmatrix} -1 & -2 & -3 \\ 5 & 2 & 1 \\ -5 & 0 & 0 \end{pmatrix} \tag{8.27}$$

∎

8.4.3 Linear Algebraic Structure of Matrices

We define the zero matrix $\mathsf{0}$ as the matrix having all entries equal to zero:

$$\mathsf{0} = \begin{pmatrix} 0 & 0 \cdots & 0 \\ 0 & 0 \cdots & 0 \\ \vdots & \vdots \ddots & \vdots \\ 0 & 0 \cdots & 0 \end{pmatrix}. \tag{8.28}$$

Then, for any $m \times n$ matrices A, B, C, and $\mathsf{0}$ and real numbers λ and μ, the algebraic rules of the preceding two subsections imply:

1. Addition is commutative:
$$\mathsf{A}+\mathsf{B} = \mathsf{B}+\mathsf{A}.$$
2. Addition is associative:
$$(\mathsf{A}+\mathsf{B})+\mathsf{C} = \mathsf{A}+(\mathsf{B}+\mathsf{C}).$$
3. There exists a unique zero matrix 0 that obeys
$$\mathsf{A}+\mathsf{0} = \mathsf{A}.$$
4. Every matrix A has a unique inverse $-\mathsf{A}$ under addition: such that
$$\mathsf{A}+(-\mathsf{A}) = \mathsf{0}.$$
5. Scalar multiplication is distributive over matrix addition,
$$\lambda(\mathsf{A}+\mathsf{B}) = \lambda\mathsf{A}+\lambda\mathsf{B},$$
$$(\lambda+\mu)\mathsf{A} = \lambda\mathsf{A}+\mu\mathsf{A},$$
and has a unit, signified by 1, such that,
$$1\mathsf{A} = \mathsf{A},$$
and successive multiplication of vectors by scalars satisfies
$$(\lambda\mu)\mathsf{A} = \lambda(\mu\mathsf{A}).$$

These laws for the operations of addition and multiplication by a real constant are analogous to the corresponding laws for vectors: $m \times n$ matrices form a linear vector space (Definition 2.1). Indeed, we may regard matrices as a generalization of vectors, namely, vectors of vectors. This can be made more precise by defining a function that maps matrices to vectors by taking the first column of the matrix, to which we successively append each remaining column to obtain an mn-dimensional column vector. As each operation described above is carried out element by element, $m \times n$ matrices form an mn-dimensional vector space.

8.4.4 Matrix Multiplication

The multiplication of matrices is a somewhat more subtle matter than the operations we have considered in the previous subsections. We first consider the multiplications of an $m \times p$ matrix A by an p-dimensional column vector $\boldsymbol{b}$:

$$\begin{aligned} \mathsf{A}\boldsymbol{b} &= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mp} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_p \end{pmatrix} \\ &= \begin{pmatrix} a_{11}b_1 + a_{12}b_2 + \cdots + a_{1p}b_p \\ a_{21}b_1 + a_{22}b_2 + \cdots + a_{2p}b_p \\ \vdots \qquad \vdots \qquad\qquad \vdots \\ a_{m1}b_1 + a_{m2}b_2 + \cdots + a_{mp}b_p \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^{p} a_{1k}b_k \\ \sum_{k=1}^{p} a_{2k}b_k \\ \vdots \\ \sum_{k=1}^{p} a_{mk}b_k \end{pmatrix}. \end{aligned} \tag{8.29}$$

Suppose we now have a $p \times n$ matrix B, which we write as n columns of p-dimensional vectors:

$$\mathsf{B} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{p1} & b_{p2} & \cdots & b_{pn} \end{pmatrix} = (\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_n), \tag{8.30}$$

in which $\boldsymbol{b}_j$ forms the jth column of B and has entries b_{ij}, with $i = 1, 2, \ldots, p$. Thus, the entries of the product AB are

$$\begin{aligned} \mathsf{AB} &= (\mathsf{A}\boldsymbol{b}_1, \mathsf{A}\boldsymbol{b}_2, \ldots, \mathsf{A}\boldsymbol{b}_n) \\ &= \begin{pmatrix} \sum_{k=1}^{p} a_{1k}b_{k1} & \sum_{k=1}^{p} a_{1k}b_{k2} & \cdots & \sum_{k=1}^{p} a_{1k}b_{kn} \\ \sum_{k=1}^{p} a_{2k}b_{k1} & \sum_{k=1}^{p} a_{2k}b_{k2} & \cdots & \sum_{k=1}^{p} a_{2k}b_{kn} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=1}^{p} a_{mk}b_{k1} & \sum_{k=1}^{p} a_{mk}b_{k2} & \cdots & \sum_{k=1}^{p} a_{mk}b_{kn} \end{pmatrix}, \end{aligned} \tag{8.31}$$

so the product AB has matrix elements

$$(\mathsf{AB})_{ij} = \sum_{k=1}^{p} a_{ik}b_{kj}, \tag{8.32}$$

for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$, that is, the product is an $m \times n$ matrix.

If $m \neq n$, then A and B can be multiplied only in the order AB. If, however, $m = n$, the the produce is well-defined and yields a $p \times p$ matrix. The more interesting case is where $p = m = n$, so that A and B are both $n \times n$ matrices, as are their products AB and BA. The two products need not be equal, and this inequality has profound ramifications for quantum mechanics, when such matrices are associated with operations such as rotations about different axes.

EXAMPLE 8.6. Consider the matrices

$$\mathsf{A} = \begin{pmatrix} 3 & -1 \\ 5 & 2 \end{pmatrix}, \qquad \mathsf{B} = \begin{pmatrix} 0 & 2 \\ 1 & 3 \end{pmatrix}, \qquad \mathsf{C} = \begin{pmatrix} 2 & 3 & 4 \\ 1 & -1 & 2 \end{pmatrix}. \tag{8.33}$$

The products AB, BA, AC, and BC are well-defined, and we find

$$\mathsf{AB} = \begin{pmatrix} 3 & -1 \\ 5 & 2 \end{pmatrix}\begin{pmatrix} 0 & 2 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} -1 & 3 \\ 2 & 16 \end{pmatrix}, \tag{8.34}$$

$$\mathsf{BA} = \begin{pmatrix} 0 & 2 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} 3 & -1 \\ 5 & 2 \end{pmatrix} = \begin{pmatrix} 10 & 4 \\ 18 & 5 \end{pmatrix}, \tag{8.35}$$

$$\mathsf{AC} = \begin{pmatrix} 3 & -1 \\ 5 & 2 \end{pmatrix}\begin{pmatrix} 2 & 3 & 4 \\ 1 & -1 & 2 \end{pmatrix} = \begin{pmatrix} 5 & 10 & 10 \\ 12 & 13 & 24 \end{pmatrix}, \tag{8.36}$$

$$\mathsf{BC} = \begin{pmatrix} 0 & 2 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} 2 & 3 & 4 \\ 1 & -1 & 2 \end{pmatrix} = \begin{pmatrix} 2 & -2 & 4 \\ 5 & 0 & 10 \end{pmatrix}. \tag{8.37}$$

We see, in particular, that $\mathsf{AB} \neq \mathsf{BA}$. ▮

8.4.5 Properties of Matrix Multiplication

We denote the unit matrix $\mathbb{1}$ as the matrix with its diagonal entries equal to 1 and all other entries equal to zero:

$$\mathbb{1} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}. \tag{8.38}$$

The following properties are valid for matrices A, B, and C for which the indicated operations are well-defined, and for any $\lambda \in \mathbb{R}$:

1. Matrix multiplication is distributive over addition:
$$(\mathsf{A}+\mathsf{B})\mathsf{C} = \mathsf{A}\mathsf{B}+\mathsf{A}\mathsf{C},$$
$$\mathsf{A}(\mathsf{B}+\mathsf{C}) = \mathsf{A}\mathsf{B}+\mathsf{A}\mathsf{C}.$$

2. Scalar multiplication commutes with matrix multiplication:
$$\lambda(\mathsf{A}\mathsf{B}) = (\lambda\mathsf{A})\mathsf{B} = \mathsf{A}(\lambda\mathsf{B}).$$

3. Matrix multiplication is associative:
$$\mathsf{A}(\mathsf{B}\mathsf{C}) = (\mathsf{A}\mathsf{B})\mathsf{C}.$$

4. Zero matrix for matrix multiplication:
$$\mathsf{A}0 = 0\mathsf{A} = 0.$$

4. Unit matrix for matrix multiplication:
$$\mathsf{A}\mathbb{1} = \mathbb{1}\mathsf{A} = \mathsf{A}.$$

8.5 Rotations

Among the most widespread applications of Theorem 8.3 is determining the matrices of rigid-body transformations, such as rotations, reflections, and inversions. These operations appear in both classical and quantum mechanics, but the algebraic properties of rotations in particular have important consequences for the quantum mechanical description of three-dimensional rotationally-invariant systems. In this section, we will derive the matrices for rotations in $\mathbb{R}^2$ and $\mathbb{R}^3$.

8.5.1 Rotations in $\mathbb{R}^2$

We consider rotations about the origin in thew x-y plane. The positive direction of rotation is usually taken in the counter-clockwise direction. For any vector $\boldsymbol{x} \in \mathbb{R}^2$,

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 . \tag{8.39}$$

The vectors $\boldsymbol{x}$, $\boldsymbol{e}_1$, and $\boldsymbol{e}_2$, and the components x_1 and x_2 are shown in Fig. 8.3(a). The matrix associated with the counter-clockwise rotation by an angle θ is obtained by following the prescription in Theorem 8.3 and referring to Fig. 8.3(b). Basic trigonometry yields

$$f(\boldsymbol{e}_1) = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}, \qquad f(\boldsymbol{e}_2) = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}. \tag{8.40}$$

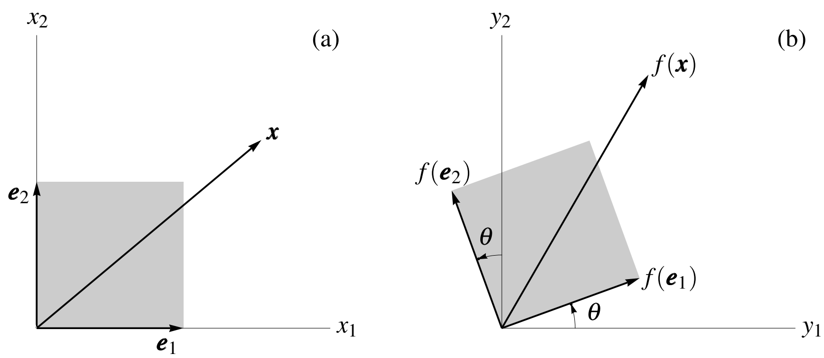

Figure 8.3: The rotation of the x-y plane about the origin in the counter-clockwise direction by an angle θ. (a) The original orientation of the basis vectors $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$, together with a vector $\boldsymbol{x}$. The shaded region is the area bounded by $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$. (b) The quantities in (a) transformed by a linear function that represents a counter-clockwise rotation by an angle θ. The shaded region is the transformed shaded area in (a). These areas are equal because the determinant of the transformation is equal to unity.

Hence,

$$\begin{aligned} f(\boldsymbol{x};\theta) &= (f(\boldsymbol{e}_1), f(\boldsymbol{e}_2))\boldsymbol{x} \\ &= \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \mathsf{R}(\theta)\boldsymbol{x}, \end{aligned} \tag{8.41}$$

which defines the matrix for a counterclockwise rotation in the x-y plane about the origin by an angle θ:

$$\boxed{\mathsf{R}(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.} \tag{8.42}$$

This matrix determines all of the properties of rotations in $\mathbb{R}^2$.

Consider first the determinant of $\mathsf{R}(\theta)$:

$$\det\big[\mathsf{R}(\theta)\big] = \begin{vmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{vmatrix} = \cos^2\theta + \sin^2\theta = 1\,. \tag{8.43}$$

the interpretation of this result is seen by observing that the two shaded regions in Fig. 8.3(a,b) have equal areas, that is, the transformation does not alter the area spanned by the basis vectors. This is what is meant by a ‘rigid-body’ transformation: the body is transformed rigidly, in this case by a rotation, with no change in volume.

We now consider a rotation by θ followed by a rotation by ϕ:

$$\begin{aligned}\mathsf{R}(\phi)\mathsf{R}(\theta) &= \begin{pmatrix}\cos\phi & -\sin\phi \\ \sin\phi & \cos\phi\end{pmatrix}\begin{pmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{pmatrix} \\ &= \begin{pmatrix}\cos\phi\cos\theta-\sin\phi\sin\theta & -\cos\phi\sin\theta-\sin\phi\cos\theta \\ \sin\phi\cos\theta+\cos\phi\sin\theta & -\sin\phi\sin\theta+\cos\phi\cos\theta\end{pmatrix} \\ &= \begin{pmatrix}\cos(\phi+\theta) & -\sin(\phi+\theta) \\ \sin(\phi+\theta) & \cos(\phi+\theta)\end{pmatrix} = \mathsf{R}(\phi+\theta)\,, \end{aligned} \tag{8.44}$$

where in writing the last matrix equality we have invoked the trigonometric identities

$$\begin{aligned}\cos(\phi+\theta) &= \cos\phi\cos\theta-\sin\phi\sin\theta\,, \\ \sin(\phi+\theta) &= \sin\phi\cos\theta+\cos\phi\sin\theta\,.\end{aligned} \tag{8.45}$$

Notice that, because ϕ and θ enter the expression for the product transformation symmetrically, there roles can be interchanged without affecting the result, so

$$\mathsf{R}(\phi)\mathsf{R}(\theta) = \mathsf{R}(\theta)\mathsf{R}(\phi)\,. \tag{8.46}$$

In other words, these rotations commute. We will see in the next section that rotations about *different* axes do *not* commute.

The rotation angle $\theta = 0$ yields the 2×2 unit matrix:

$$\mathsf{R}(0) = \begin{pmatrix}1 & 0 \\ 0 & 1\end{pmatrix}, \tag{8.47}$$

which is to be expected because $\theta = 0$ corresponds to performing no transformation at all. With this result and that in (8.44) we have that

$$\mathsf{R}(0) = \mathsf{R}(\theta-\theta) = \mathsf{R}(\theta)\mathsf{R}(-\theta) = \mathsf{R}(-\theta)\mathsf{R}(\theta)\,. \tag{8.48}$$

Hence, $[\mathsf{R}(\theta)]^{-1} = \mathsf{R}(-\theta)$, which simply says that reversing a transformation yields the original vectors.

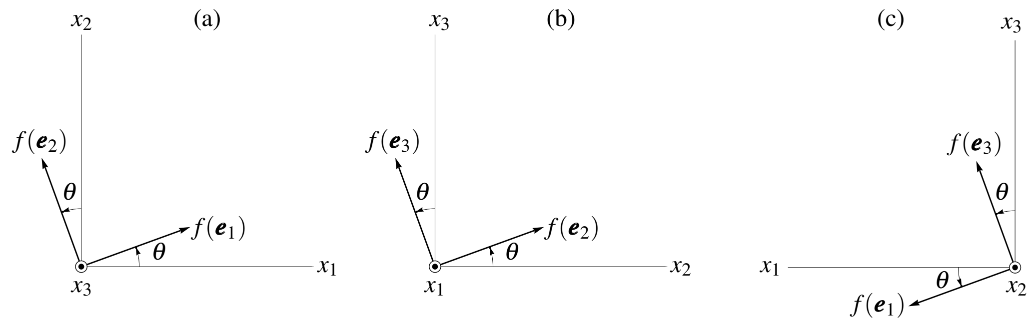

Figure 8.4: The rotations in the counter-clockwise direction by an angle θ about each of the coordinate axes of a right-handed coordinate system in $\mathbb{R}^3$: rotations about (a) the x_3 axis, (b) the x_1 axis, and (c) the x_2-axis. Each of the rotation axes point out of the plane of the page.

8.5.2 Rotations in $\mathbb{R}^3$

We will not develop the complete theory of rotations in $\mathbb{R}^3$, but consider only rotations about the three coordinate axes. This will provide an indication of the fundamental difference between these rotations and those in $\mathbb{R}^2$.

The rotations we are considering are shown in Fig. 8.4. For any vector $\boldsymbol{x} \in \mathbb{R}^3$,

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = x_1 \boldsymbol{e}_1 + x_2 \boldsymbol{e}_2 + x_3 \boldsymbol{e}_3 . \tag{8.49}$$

We first consider rotations about the x_3 axis, which is depicted in Fig. 8.4(a). The x_3 axis is unaffected by the transformation, and the x_1- and x_2-axes are transformed in the same manner as for two-dimensional rotations. Hence, we have

$$f_3(\boldsymbol{e}_1) = \begin{pmatrix} \cos\theta \\ \sin\theta \\ 0 \end{pmatrix}, \qquad f_3(\boldsymbol{e}_2) = \begin{pmatrix} -\sin\theta \\ \cos\theta \\ 0 \end{pmatrix}, \qquad f_3(\boldsymbol{e}_3) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{8.50}$$

The matrix corresponding to this linear transformation is

$$\begin{aligned} f_3(\boldsymbol{x};\theta) &= (f_3(\boldsymbol{e}_1), f_3(\boldsymbol{e}_2), f_3)\boldsymbol{e}_3))\boldsymbol{x} \\ &= \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \mathsf{R}_3(\theta)\boldsymbol{x}, \end{aligned} \tag{8.51}$$

which defines

$$\mathsf{R}_3(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{8.52}$$

As Fig. 8.4(b) shows, rotations about the x_1-axis have a similar geometry to rotations about the x_3-axis. A procedure analogous ton that used to determine $\mathsf{R}_3(\theta)$ yields

$$\mathsf{R}_1(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}. \tag{8.53}$$

Finally, we consider rotations about the x_2-axis [Fig., 8.4(c)]. We obtain

$$f_2(\boldsymbol{e}_1) = \begin{pmatrix} \cos\theta \\ 0 \\ -\sin\theta \end{pmatrix}, \qquad f_2(\boldsymbol{e}_2) = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad f_2(\boldsymbol{e}_3) = \begin{pmatrix} \sin\theta \\ 0 \\ \cos\theta \end{pmatrix}. \tag{8.54}$$

The matrix form of this linear transformation is

$$\begin{aligned} f_2(\boldsymbol{x};\theta) &= (f_2(\boldsymbol{e}_1), f_2(\boldsymbol{e}_2), f_2(\boldsymbol{e}_3))\boldsymbol{x} \\ &= \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \mathsf{R}_2(\theta)\boldsymbol{x}, \end{aligned} \tag{8.55}$$

which defines

$$\mathsf{R}_2(\theta) = \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}. \tag{8.56}$$

Each of the rotations about the coordinate aces separately have the same properties as two-dimensional rotations. However, taken together, some different properties emerge. Most notably, we calculate, for example, $\mathsf{R}_1(\theta)\mathsf{R}_2(\theta)$:

$$\begin{aligned}\mathsf{R}_1(\theta)\mathsf{R}_2(\theta) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix} \\ &= \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ -\sin^2\theta & \cos\theta & -\sin\theta\cos\theta \\ -\sin\theta\cos\theta & \sin\theta & \cos^2\theta \end{pmatrix}, \end{aligned} \tag{8.57}$$

which we compare with

$$\begin{aligned}\mathsf{R}_2(\theta)\mathsf{R}_1(\theta) &= \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \\ &= \begin{pmatrix} \cos\theta & \sin^2\theta & \sin\theta\cos\theta \\ 0 & \cos\theta & -\sin\theta \\ -\sin\theta & \sin\theta\cos\theta & \cos^2\theta \end{pmatrix}, \end{aligned} \tag{8.58}$$

so $\mathsf{R}_1(\theta)\mathsf{R}_2(\theta) \neq \mathsf{R}_2(\theta)\mathsf{R}_1(\theta)$. This is a general property of rotations in three dimensions that has important implications for quantum mechanical systems,.

8.6 Summary

The main theme of this chapter is the relationship between linear functions and matrices. A linear function $f\colon \mathbb{R}^n \to \mathbb{R}^m$ can be succinctly defines as having the property that

$$f(\lambda\boldsymbol{x}+\mu\boldsymbol{y}) = \lambda f(\boldsymbol{x}) + \mu f(\boldsymbol{y}),$$

for any $\boldsymbol{x},\boldsymbol{y} \in \mathbb{R}^n$ and $\lambda,\mu \in \mathbb{R}$.

The equivalence of linear functions and matrices was established by Theorem 8.3 and led to the transcription of operations between functions to those between matrices:

$$\begin{aligned} f(\boldsymbol{x}) + g(\boldsymbol{x}) &= \mathsf{A}\boldsymbol{x} + \mathsf{B}\boldsymbol{x} = (\mathsf{A}+\mathsf{B})\boldsymbol{x}, \\ \lambda f(\boldsymbol{x}) &= \lambda(\mathsf{A}\boldsymbol{x}) = (\lambda\mathsf{A})\boldsymbol{x}, \\ (f\circ g)(\boldsymbol{x}) &= f(g(\boldsymbol{x})) = \mathsf{A}(\mathsf{B}\boldsymbol{x}) = (\mathsf{AB})\boldsymbol{x}, \end{aligned}$$

Addition and multiplication can be defined for any $m \times n$ matrices, but the product AB can be defined only for an $m \times p$ matrix A and a $p \times n$ matrix B, in which case the product is an $m \times n$ matrix.

Rotations are among the most important linear transformations and are expressed as $n \times n$ matrices. Rotations $\mathsf{R}(\theta)$ in two dimensions about the origin in the counter-clockwise direction by an angle θ were found to have the following properties:

1. $\det\big[\mathsf{R}(\theta)\big] = 1$, which means that areas are preserved under rotations.
2. The composition of rotations is $\mathsf{R}(\theta)\mathsf{R}(\phi) = \mathsf{R}(\theta + \phi)$.
3. The composition of rotation is commutative: $\mathsf{R}(\theta)\mathsf{R}(\phi) = \mathsf{R}(\phi)\mathsf{R}(\theta)$.
4. $\mathsf{R}(0) = \mathbb{1}$, where $\mathbb{1}$ is the 2×2 unit, or identity, matrix.
5. The inverse $\big[\mathsf{R}(\theta)\big]^{-1}$ of a rotation $\mathsf{R}(\theta)$ is $\mathsf{R}(-\theta)$.

Rotations in $\mathbb{R}^3$ share all of the properties except 3: rotations in three dimensions do not generally commute. The ramifications of this non-commutativity will be explored in quantum mechanics.