

# 3

## *Linear Algebra*

*"Physics is much too hard for physicists." David Hilbert (1862-1943)*

AS THE READER IS AWARE BY NOW, calculus has its roots in physics and has become a very useful tool for modeling the physical world. Another very important area of mathematics is linear algebra. Physics students who have taken a course in linear algebra in a mathematics department might not come away with this perception. It is not until students take more advanced classes in physics that they begin to realize that a good grounding in linear algebra can lead to a better understanding of the behavior of physical systems.

In this chapter we will introduce some of the basics of linear algebra for finite dimensional vector spaces and we will reinforce these concepts through generalizations in later chapters to infinite dimensional vector spaces. In keeping with the theme of our text, we will apply some of these ideas to the coupled systems introduced in the last chapter. Such systems lead to linear and nonlinear oscillations in dynamical systems.

### 3.1 *Vector Spaces*

MUCH OF THE DISCUSSION and terminology that we will use comes from the theory of vector spaces . Up until now you may only have dealt with finite dimensional vector spaces. Even then, you might only be comfortable with two and three dimensions. We will review a little of what we know about finite dimensional spaces so that we can introduce more general function spaces.

The notion of a vector space is a generalization of three dimensional vectors and operations on them. In three dimensions, we have things called vectors<sup>1</sup> , which are arrows of a specific length and pointing in a given direction. To each vector, we can associate a point in a three dimensional Cartesian system. We just attach the tail of the vector  $\mathbf{v}$  to

Linear algebra is the backbone of most of applied mathematics and underlies many areas of physics, such as quantum mechanics.

<sup>1</sup> In introductory physics one defines a vector as any quantity that has both magnitude and direction.

the origin and the head lands at  $(x, y, z)$ .<sup>2</sup> We then use unit vectors  $\mathbf{i}$ ,  $\mathbf{j}$  and  $\mathbf{k}$  along the coordinate axes to write

$$\mathbf{v} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}.$$

Having defined vectors, we then learned how to add vectors and multiply vectors by numbers, or scalars. Under these operations, we expected to get back new vectors. Then we learned that there were two types of multiplication of vectors. We could multiply them to get a scalar or a vector. This leads to dot products and cross products, respectively. The dot product is useful for determining the length of a vector, the angle between two vectors, or if the vectors were orthogonal. The cross product is used to produce orthogonal vectors, areas of parallelograms, and volumes of parallelepipeds.

In physics you first learned about vector products when you defined work,  $W = \mathbf{F} \cdot \mathbf{r}$ . Cross products were useful in describing things like torque,  $\boldsymbol{\tau} = \mathbf{r} \times \mathbf{F}$ , or the force on a moving charge in a magnetic field,  $\mathbf{F} = q\mathbf{v} \times \mathbf{B}$ . We will return to these more complicated vector operations later when reviewing Maxwell's equations of electrodynamics.

These notions are then generalized to spaces of more than three dimensions in linear algebra courses. The properties outlined roughly above need to be preserved. So, we have to start with a space of vectors and the operations between them. We also need a set of scalars, which generally come from some *field*. However, in our applications the field will either be the set of real numbers or the set of complex numbers.<sup>3</sup>

A *vector space*  $V$  over a field  $F$  is a set that is closed under addition and scalar multiplication and satisfies the following conditions: For any  $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$  and  $a, b \in F$

1.  $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ .
2.  $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ .
3. There exists a  $\mathbf{0}$  such that  $\mathbf{0} + \mathbf{v} = \mathbf{v}$ .
4. There exists an additive inverse,  $-\mathbf{v}$ , such that  $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$ .

There are several distributive properties:

5.  $a(b\mathbf{v}) = (ab)\mathbf{v}$ .
6.  $(a + b)\mathbf{v} = a\mathbf{v} + b\mathbf{v}$ .
7.  $a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}$ .
8. There exists a multiplicative identity,  $1$ , such that  $1(\mathbf{v}) = \mathbf{v}$ .

For now, we will restrict our examples to two and three dimensions and the field will consist of the real numbers.

In three dimensions the unit vectors  $\mathbf{i}$ ,  $\mathbf{j}$ , and  $\mathbf{k}$  play an important role. Any vector in the three dimensional space can be written as a

<sup>2</sup> In multivariate calculus one concentrates on the component form of vectors. These representations are easily generalized as we will see.

<sup>3</sup> A field is a set together with two operations, usually addition and multiplication, such that we have

- Closure under addition and multiplication
- Associativity of addition and multiplication
- Commutativity of addition and multiplication
- Additive and multiplicative identity
- Additive and multiplicative inverses
- Distributivity of multiplication over addition

linear combination of these vectors,

$$\mathbf{v} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}.$$

In fact, given any three non-coplanar vectors,  $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ , all vectors can be written as a linear combination of those vectors,

$$\mathbf{v} = c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + c_3\mathbf{a}_3.$$

Such vectors are said to span the space and are called a basis for the space.

We can generalize these ideas. In an  $n$ -dimensional vector space any vector in the space can be represented as the sum over  $n$  linearly independent vectors (the equivalent of non-coplanar vectors). Such a *linearly independent* set of vectors  $\{\mathbf{v}_j\}_{j=1}^n$  satisfies the condition

$$\sum_{j=1}^n c_j \mathbf{v}_j = \mathbf{0} \quad \Leftrightarrow \quad c_j = 0.$$

Note that we will often use summation notation instead of writing out all of the terms in the sum.

This leads to the idea of a basis set. The *standard basis* in an  $n$ -dimensional vector space is a generalization of the standard basis in three dimensions ( $\mathbf{i}$ ,  $\mathbf{j}$  and  $\mathbf{k}$ ). We define

$$\mathbf{e}_k = (0, \dots, 0, \underbrace{1}_{k\text{th space}}, 0, \dots, 0), \quad k = 1, \dots, n. \quad (3.1)$$

Then, we can expand any  $\mathbf{v} \in V$  as

$$\mathbf{v} = \sum_{k=1}^n v_k \mathbf{e}_k, \quad (3.2)$$

where the  $v_k$ 's are called the components of the vector in this basis. Sometimes we will write  $\mathbf{v}$  as an  $n$ -tuple  $(v_1, v_2, \dots, v_n)$ . This is similar to the ambiguous use of  $(x, y, z)$  to denote both vectors and points in the three dimensional space.

The only other thing we will need at this point is to generalize the dot product. Recall that there are two forms for the dot product in three dimensions. First, one has that

$$\mathbf{u} \cdot \mathbf{v} = uv \cos \theta, \quad (3.3)$$

where  $u$  and  $v$  denote the length of the vectors. The other form is the component form:

$$\mathbf{u} \cdot \mathbf{v} = u_1v_1 + u_2v_2 + u_3v_3 = \sum_{k=1}^3 u_kv_k. \quad (3.4)$$

The standard basis vectors,  $\mathbf{e}_k$  are a natural generalization of  $\mathbf{i}$ ,  $\mathbf{j}$  and  $\mathbf{k}$ .

For more general vector spaces the term inner product is used to generalize the notions of dot and scalar products as we will see below.

Of course, this form is easier to generalize. So, we define the *scalar product* between two  $n$ -dimensional vectors as

$$\langle \mathbf{u}, \mathbf{v} \rangle = \sum_{k=1}^n u_k v_k. \quad (3.5)$$

Actually, there are a number of notations that are used in other texts. One can write the scalar product as  $(\mathbf{u}, \mathbf{v})$  or even in the Dirac bra-ket notation<sup>4</sup>  $\langle \mathbf{u} | \mathbf{v} \rangle$ .

We note that the (real) scalar product satisfies some simple properties. For vectors  $\mathbf{v}$ ,  $\mathbf{w}$  and real scalar  $\alpha$  we have

1.  $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$  and  $\langle \mathbf{v}, \mathbf{v} \rangle = 0$  if and only if  $\mathbf{v} = \mathbf{0}$ .
2.  $\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle$ .
3.  $\langle \alpha \mathbf{v}, \mathbf{w} \rangle = \alpha \langle \mathbf{v}, \mathbf{w} \rangle$ .

While it does not always make sense to talk about angles between general vectors in higher dimensional vector spaces, there is one concept that is useful. It is that of orthogonality, which in three dimensions is another way of saying the vectors are perpendicular to each other. So, we also say that vectors  $\mathbf{u}$  and  $\mathbf{v}$  are *orthogonal* if and only if  $\langle \mathbf{u}, \mathbf{v} \rangle = 0$ . If  $\{\mathbf{a}_k\}_{k=1}^n$  is a set of basis vectors such that

$$\langle \mathbf{a}_j, \mathbf{a}_k \rangle = 0, \quad k \neq j,$$

then it is called an *orthogonal basis*.

If in addition each basis vector is a unit vector, then one has an *orthonormal basis*. This generalization of the unit basis can be expressed more compactly. We will denote such a basis of unit vectors by  $\mathbf{e}_j$  for  $j = 1 \dots n$ . Then,

$$\langle \mathbf{e}_j, \mathbf{e}_k \rangle = \delta_{jk}, \quad (3.6)$$

where we have introduced the Kronecker delta (named after Leopold Kronecker (1823-1891))

$$\delta_{jk} \equiv \begin{cases} 0, & j \neq k \\ 1, & j = k \end{cases} \quad (3.7)$$

The process of making vectors have unit length is called *normalization*. This is simply done by dividing by the *length* of the vector. Recall that the length of a vector,  $\mathbf{v}$ , is obtained as  $v = \sqrt{\mathbf{v} \cdot \mathbf{v}}$ . So, if we want to find a unit vector in the direction of  $\mathbf{v}$ , then we simply normalize it as

$$\hat{\mathbf{v}} = \frac{\mathbf{v}}{v}.$$

Notice that we used a hat to indicate that we have a unit vector. Furthermore, if  $\{\mathbf{a}_j\}_{j=1}^n$  is a set of orthogonal basis vectors, then

$$\hat{\mathbf{a}}_j = \frac{\mathbf{a}_j}{\sqrt{\langle \mathbf{a}_j, \mathbf{a}_j \rangle}}, \quad j = 1 \dots n.$$

<sup>4</sup> The bra-ket notation was introduced by Paul Adrien Maurice Dirac (1902-1984) in order to facilitate computations of inner products in quantum mechanics. In the notation  $\langle \mathbf{u} | \mathbf{v} \rangle$ ,  $\langle \mathbf{u} |$  is the bra and  $|\mathbf{v} \rangle$  is the ket. The kets live in a vector space and represented by column vectors with respect to a given basis. The bras live in the dual vector space and are represented by row vectors. The correspondence between bra and kets is  $|\mathbf{v} \rangle = \overline{|\mathbf{v} \rangle^T}$ . One can operate on kets,  $A|\mathbf{v} \rangle$ , and make sense out of operations like  $\langle \mathbf{u} | A | \mathbf{v} \rangle$ , which are used to obtain expectation values associated with the operator. Finally, the outer product,  $|\mathbf{v} \rangle \langle \mathbf{v} |$  is used to perform vector space projections.

Orthogonal basis vectors.

Normalization of vectors.

**Example 3.1.** Find the angle between the vectors  $\mathbf{u} = (-2, 1, 3)$  and  $\mathbf{v} = (1, 0, 2)$ . we need the lengths of each vector,

$$u = \sqrt{(-2)^2 + 1^2 + 3^2} = \sqrt{14},$$

$$v = \sqrt{1^2 + 0^2 + 2^2} = \sqrt{5}.$$

We also need the scalar product of these vectors,

$$\mathbf{u} \cdot \mathbf{v} = -2 + 6 = 4.$$

This gives

$$\cos \theta = \frac{\mathbf{u} \cdot \mathbf{v}}{uv} = \frac{4}{\sqrt{5}\sqrt{14}}.$$

So,  $\theta = 61.4^\circ$ .

**Example 3.2.** Normalize the vector  $\mathbf{v} = 2\mathbf{i} + \mathbf{j} - 2\mathbf{k}$ .

The length of the vector is  $v = \sqrt{2^2 + 1^2 + (-2)^2} = \sqrt{9} = 3$ . So, the unit vector in the direction of  $\mathbf{v}$  is  $\hat{\mathbf{v}} = \frac{2}{3}\mathbf{i} + \frac{1}{3}\mathbf{j} - \frac{2}{3}\mathbf{k}$ .

Let  $\{\mathbf{a}_k\}_{k=1}^n$  be a set of orthogonal basis vectors for vector space  $V$ . We know that any vector  $\mathbf{v}$  can be represented in terms of this basis,  $\mathbf{v} = \sum_{k=1}^n v_k \mathbf{a}_k$ . If we know the basis and vector, can we find the components,  $v_k$ ? The answer is yes. We can use the scalar product of  $\mathbf{v}$  with each basis element  $\mathbf{a}_j$ . Using the properties of the scalar product, we have for  $j = 1, \dots, n$

$$\begin{aligned} \langle \mathbf{a}_j, \mathbf{v} \rangle &= \langle \mathbf{a}_j, \sum_{k=1}^n v_k \mathbf{a}_k \rangle \\ &= \sum_{k=1}^n v_k \langle \mathbf{a}_j, \mathbf{a}_k \rangle. \end{aligned} \quad (3.8)$$

Since we know the basis elements, we can easily compute the numbers

$$A_{jk} \equiv \langle \mathbf{a}_j, \mathbf{a}_k \rangle$$

and

$$b_j \equiv \langle \mathbf{a}_j, \mathbf{v} \rangle.$$

Therefore, the system (3.8) for the  $v_k$ 's is a linear algebraic system, which takes the form

$$b_j = \sum_{k=1}^n A_{jk} v_k. \quad (3.9)$$

We can write this set of equations in a more compact form. The set of numbers  $A_{jk}$ ,  $j, k = 1, \dots, n$  are the elements of an  $n \times n$  matrix  $A$  with  $A_{jk}$  being an element in the  $j$ th row and  $k$ th column. Also,  $v_j$  and  $b_j$  can be written as column vectors,  $\mathbf{v}$  and  $\mathbf{b}$ , respectively. Thus, system (3.8) can be written in matrix form as

$$A\mathbf{v} = \mathbf{b}.$$

However, if the basis is orthogonal, then the matrix  $A_{jk} \equiv \langle \mathbf{a}_j, \mathbf{a}_k \rangle$  is diagonal and the system is easily solvable. Recall that two vectors are orthogonal if and only if

$$\langle \mathbf{a}_i, \mathbf{a}_j \rangle = 0, \quad i \neq j. \quad (3.10)$$

Thus, in this case we have that

$$\langle \mathbf{a}_j, \mathbf{v} \rangle = v_j \langle \mathbf{a}_j, \mathbf{a}_j \rangle, \quad j = 1, \dots, n. \quad (3.11)$$

or

$$v_j = \frac{\langle \mathbf{a}_j, \mathbf{v} \rangle}{\langle \mathbf{a}_j, \mathbf{a}_j \rangle}. \quad (3.12)$$

In fact, if the basis is orthonormal, i.e., the basis consists of an orthogonal set of unit vectors, then  $A$  is the identity matrix and the solution takes on a simpler form:

$$v_j = \langle \mathbf{a}_j, \mathbf{v} \rangle. \quad (3.13)$$

**Example 3.3.** Consider the set of vectors  $\mathbf{a}_1 = \mathbf{i} + \mathbf{j}$  and  $\mathbf{a}_2 = \mathbf{i} - 2\mathbf{j}$ .

1. Determine the matrix elements  $A_{jk} = \langle \mathbf{a}_j, \mathbf{a}_k \rangle$ .
2. Is this an orthogonal basis?
3. Expand the vector  $\mathbf{v} = 2\mathbf{i} + 3\mathbf{j}$  in the basis  $\{\mathbf{a}_1, \mathbf{a}_2\}$ .

First, we compute the matrix elements of  $A$ :

$$\begin{aligned} A_{11} &= \langle \mathbf{a}_1, \mathbf{a}_1 \rangle = 2 \\ A_{12} &= \langle \mathbf{a}_1, \mathbf{a}_2 \rangle = -1 \\ A_{21} &= \langle \mathbf{a}_2, \mathbf{a}_1 \rangle = -1 \\ A_{22} &= \langle \mathbf{a}_2, \mathbf{a}_2 \rangle = 5 \end{aligned} \quad (3.14)$$

So,

$$A = \begin{pmatrix} 2 & -1 \\ -1 & 5 \end{pmatrix}.$$

Since  $A_{12} = A_{21} \neq 0$ , the vectors are not orthogonal. However, they are linearly independent. Obviously, if  $c_1 = c_2 = 0$ , then the linear combination  $c_1\mathbf{a}_1 + c_2\mathbf{a}_2 = \mathbf{0}$ . Conversely, we assume that  $c_1\mathbf{a}_1 + c_2\mathbf{a}_2 = \mathbf{0}$  and solve for the coefficients. Inserting the given vectors, we have

$$\begin{aligned} \mathbf{0} &= c_1(\mathbf{i} + \mathbf{j}) + c_2(\mathbf{i} - 2\mathbf{j}) \\ &= (c_1 + c_2)\mathbf{i} + (c_1 - 2c_2)\mathbf{j}. \end{aligned} \quad (3.15)$$

This implies that

$$\begin{aligned} c_1 + c_2 &= 0 \\ c_1 - 2c_2 &= 0. \end{aligned} \quad (3.16)$$

Solving this system, one has  $c_1 = 0, c_2 = 0$ . Therefore, the two vectors are linearly independent.

In order to determine the components of  $\mathbf{v}$  with respect to the new basis, we need to set up the system (3.8) and solve for the  $v_k$ 's. We have first,

$$\begin{aligned} \mathbf{b} &= \begin{pmatrix} \langle \mathbf{a}_1, \mathbf{v} \rangle \\ \langle \mathbf{a}_2, \mathbf{v} \rangle \end{pmatrix} \\ &= \begin{pmatrix} \langle \mathbf{i} + \mathbf{j}, 2\mathbf{i} + 3\mathbf{j} \rangle \\ \langle \mathbf{i} - 2\mathbf{j}, 2\mathbf{i} + 3\mathbf{j} \rangle \end{pmatrix} \\ &= \begin{pmatrix} 5 \\ -4 \end{pmatrix}. \end{aligned} \tag{3.17}$$

So, now we have to solve the system  $A\mathbf{v} = \mathbf{b}$  for  $\mathbf{v}$  :

$$\begin{pmatrix} 2 & -1 \\ -1 & 5 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 5 \\ -4 \end{pmatrix}. \tag{3.18}$$

We can solve this with matrix methods,  $\mathbf{v} = A^{-1}\mathbf{b}$ , or rewrite it as a system of two equations and two unknowns. The result is  $v_1 = \frac{7}{3}, v_2 = -\frac{1}{3}$ . Thus,  $\mathbf{v} = \frac{7}{3}\mathbf{a}_1 - \frac{1}{3}\mathbf{a}_2$ .

### 3.2 Linear Transformations

A MAIN THEME in linear algebra is to study linear transformations between vector spaces. These come in many forms and there are an abundance of applications in physics. For example, the transformation between the spacetime coordinates of observers moving in inertial frames in the theory of special relativity constitute such a transformation.

A simple example often encountered in physics courses is the rotation by a fixed angle. This is the description of points in space using two different coordinate bases, one just a rotation of the other by some angle. We begin with a vector  $\mathbf{v}$  as described by a set of axes in the standard orientation, as shown in Figure 3.1. Also displayed in this figure are the unit vectors. To find the coordinates  $(x, y)$ , one needs only draw perpendiculars to the axes and read the coordinates off the axes.

In order to derive the needed transformation we will make use of polar coordinates. In Figure 3.1 we see that the vector makes an angle of  $\phi$  with respect to the positive  $x$ -axis. The components  $(x, y)$  of the vector can be determined from this angle and the magnitude of  $\mathbf{v}$  as

$$\begin{aligned} x &= v \cos \phi \\ y &= v \sin \phi. \end{aligned} \tag{3.19}$$

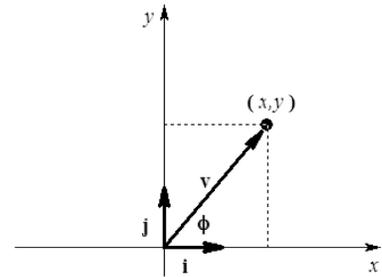


Figure 3.1: Vector  $\mathbf{v}$  in a standard coordinate system.

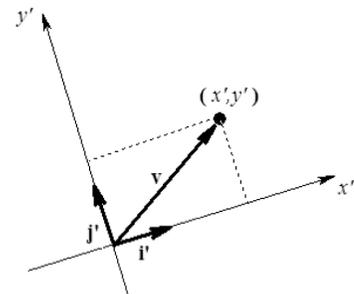


Figure 3.2: Vector  $\mathbf{v}$  in a rotated coordinate system.

We now consider another set of axes at an angle of  $\theta$  to the old. Such a system is shown in Figure 3.2. We will designate these axes as  $x'$  and  $y'$ . Note that the basis vectors are different in this system. Projections to the axes are shown. Comparing the coordinates in both systems shown in Figures 3.1-3.2, we see that the primed coordinates are not the same as the unprimed ones.

In Figure 3.3 the two systems are superimposed on each other. The polar form for the primed system is given by

$$\begin{aligned}x' &= v \cos(\phi - \theta) \\y' &= v \sin(\phi - \theta).\end{aligned}\quad (3.20)$$

We can use this form to find a relationship between the two systems. Namely, we use the addition formula for trigonometric functions to obtain

$$\begin{aligned}x' &= v \cos \phi \cos \theta + v \sin \phi \sin \theta \\y' &= v \sin \phi \cos \theta - v \cos \phi \sin \theta.\end{aligned}\quad (3.21)$$

Noting that these expressions involve products of  $v$  with  $\cos \phi$  and  $\sin \phi$ , we can use the polar form for  $x$  and  $y$  to find the desired form:

$$\begin{aligned}x' &= x \cos \theta + y \sin \theta \\y' &= -x \sin \theta + y \cos \theta.\end{aligned}\quad (3.22)$$

This is an example of a transformation between two coordinate systems. It is called a rotation by  $\theta$ . We can designate it generally by

$$(x', y') = \hat{R}_\theta(x, y).$$

It is referred to as a passive transformation, because it does not affect the vector. [Note: We will use the hat for the passive rotation.]

An active rotation is one in which one rotates the vector, such as shown in Figure 3.4. One can derive a similar transformation for how the coordinate of the vector change under such a transformation. Denoting the new vector as  $\mathbf{v}'$  with new coordinates  $(x'', y'')$ , we have

$$\begin{aligned}x'' &= x \cos \theta - y \sin \theta \\y'' &= x \sin \theta + y \cos \theta.\end{aligned}\quad (3.23)$$

We can designate this transformation by

$$(x'', y'') = R_\theta(x, y)$$

and see that the active and passive rotations are related,

$$R_\theta(x, y) = \hat{R}_{-\theta}(x, y).$$

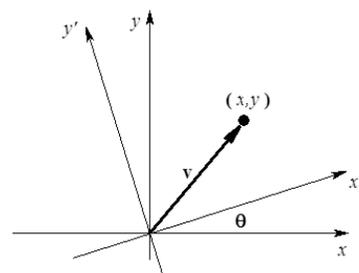


Figure 3.3: Comparison of the coordinate systems.

Passive rotation.

Active rotation.

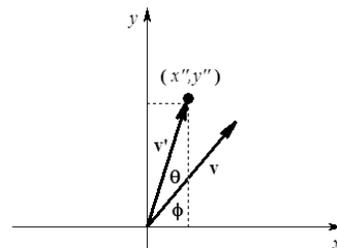


Figure 3.4: Rotation of vector  $\mathbf{v}$

### 3.3 Matrices

LINEAR TRANSFORMATIONS such as the rotation in the last section can be represented by matrices. Such matrix representations often become the core of a linear algebra class to the extent that one loses sight of their meaning. We will review matrix representations and show how they are useful in solving coupled systems of differential equations later in the chapter.

We begin with the rotation transformation as applied to the axes in Equation (3.22). We write vectors like  $\mathbf{v}$  as a column matrix

$$\mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix}.$$

We can also write the trigonometric functions in a  $2 \times 2$  matrix form as

$$\hat{R}_\theta = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}.$$

Then, the transformation takes the form

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \quad (3.24)$$

This can be written in the more compact form

$$\mathbf{v}' = \hat{R}_\theta \mathbf{v}.$$

In using the matrix form of the transformation, we have employed the definition of matrix multiplication. Namely, we have multiplied a  $2 \times 2$  matrix times a  $2 \times 1$  matrix. (Note that an  $n \times m$  matrix has  $n$  rows and  $m$  columns.) The multiplication proceeds by selecting the  $i$ th row of the first matrix and the  $j$ th column of the second matrix. Multiply corresponding elements of each and add them. Then, place the result into the  $ij$ th entry of the product matrix. This operation can only be performed if the number of columns of the first matrix is the same as the number of columns of the second matrix.

**Example 3.4.** *As an example, we multiply a  $3 \times 2$  matrix times a  $2 \times 2$  matrix to obtain a  $3 \times 2$  matrix:*

$$\begin{aligned} \begin{pmatrix} 1 & 2 \\ 5 & -1 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix} &= \begin{pmatrix} 1(3) + 2(1) & 1(2) + 2(4) \\ 5(3) + (-1)(1) & 5(2) + (-1)(4) \\ 3(3) + 2(1) & 3(2) + 2(4) \end{pmatrix} \\ &= \begin{pmatrix} 5 & 10 \\ 14 & 6 \\ 11 & 14 \end{pmatrix} \end{aligned} \quad (3.25)$$

In Equation (3.24), we have the row  $(\cos \theta, \sin \theta)$  and column  $(x, y)^T$ . Combining these we obtain  $x \cos \theta + y \sin \theta$ . This is  $x'$ . We perform the same operation for the second row:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \cos \theta + y \sin \theta \\ -x \sin \theta + y \cos \theta \end{pmatrix}. \quad (3.26)$$

In the last section we also introduced active rotations. These were rotations of vectors keeping the coordinate system fixed. Thus, we start with a vector  $\mathbf{v}$  and rotate it by  $\theta$  to get a new vector  $\mathbf{u}$ . That transformation can be written as

$$\mathbf{u} = R_\theta \mathbf{v}, \quad (3.27)$$

where

$$R_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

Now consider a rotation by  $-\theta$ . Due to the symmetry properties of the sines and cosines, we have

$$R_{-\theta} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}.$$

We see that if the 12 and 21 elements of this matrix are interchanged we recover  $R_\theta$ . This is an example of what is called the *transpose* of  $R_\theta$ . Given a matrix,  $A$ , its transpose  $A^T$  is the matrix obtained by interchanging the rows and columns of  $A$ . Formally, let  $A_{ij}$  be the elements of  $A$ . Then

$$A_{ij}^T = A_{ji}.$$

Matrix transpose.

It is also the case that these matrices are inverses of each other. We can understand this in terms of the nature of rotations. We first rotate the vector by  $\theta$  as  $\mathbf{u} = R_\theta \mathbf{v}$  and then rotate  $\mathbf{u}$  by  $-\theta$  obtaining  $\mathbf{w} = R_{-\theta} \mathbf{u}$ . Thus, the "composition" of these two transformations leads to

$$\mathbf{w} = R_{-\theta} \mathbf{u} = R_{-\theta} (R_\theta \mathbf{v}). \quad (3.28)$$

We can view this as a net transformation from  $\mathbf{v}$  to  $\mathbf{w}$  given by

$$\mathbf{w} = (R_{-\theta} R_\theta) \mathbf{v},$$

where the transformation matrix for the composition is given by  $R_{-\theta} R_\theta$ . Actually, if you think about it, we should end up with the original vector. We can compute the resulting matrix by carrying out the multiplication. We obtain

$$R_{-\theta} R_\theta = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (3.29)$$

This is the  $2 \times 2$  identity matrix. We note that the product of these two matrices yields the identity. This is like the multiplication of numbers. If  $ab = 1$ , then  $a$  and  $b$  are multiplicative inverses of each other. So, we see here that  $R_\theta$  and  $R_{-\theta}$  are inverses of each other as well. In fact, we have determined that

$$R_{-\theta} = R_\theta^{-1} = R_\theta^T, \tag{3.30}$$

where the  $T$  designates the transpose. We note that matrices satisfying the relation  $A^T = A^{-1}$  are called *orthogonal matrices*.

Orthogonal matrices.

We can easily extend this discussion to three dimensions. Such rotations in the  $xy$ -plane can be viewed as rotations about the  $z$ -axis. Rotating a vector about the  $z$ -axis by angle  $\alpha$  will leave the  $z$ -component fixed. This can be represented by the rotation matrix

$$R_z(\alpha) = \begin{pmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{3.31}$$

We can also rotate vectors about the other axes, so that we would have two other rotation matrices:

$$R_y(\beta) = \begin{pmatrix} \cos \beta & 0 & -\sin \beta \\ 0 & 1 & 0 \\ \sin \beta & 0 & \cos \beta \end{pmatrix}. \tag{3.32}$$

$$R_x(\gamma) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \gamma & \sin \gamma \\ 0 & -\sin \gamma & \cos \gamma \end{pmatrix}. \tag{3.33}$$

As before, passive rotations of the coordinate axes are obtained by replacing the angles above by their negatives; e.g.,  $\hat{R}_x(\gamma) = R_x(-\gamma)$ .<sup>5</sup>

<sup>5</sup> In classical dynamics one describes a general rotation in terms of the so-called Euler angles. These are the angles  $(\phi, \theta, \psi)$  such that the combined rotation  $\hat{R}_z(\psi)\hat{R}_x(\theta)\hat{R}_z(\phi)$  rotates the initial coordinate system into a new one.

We can generalize what we have seen with the simple example of rotation to other linear transformations. We begin with a vector  $\mathbf{v}$  in an  $n$ -dimensional vector space. We can consider a transformation  $L$  that takes  $\mathbf{v}$  into a new vector  $\mathbf{u}$  as

$$\mathbf{u} = L(\mathbf{v}).$$

We will restrict ourselves to linear transformations between two  $n$ -dimensional vector spaces. A *linear transformation* satisfies the following condition:

$$L(\alpha\mathbf{a} + \beta\mathbf{b}) = \alpha L(\mathbf{a}) + \beta L(\mathbf{b}) \tag{3.34}$$

for any vectors  $\mathbf{a}$  and  $\mathbf{b}$  and scalars  $\alpha$  and  $\beta$ .<sup>6</sup>

Such linear transformations can be represented by matrices. Take any vector  $\mathbf{v}$ . It can be represented in terms of a basis. Let's use the standard basis  $\{\mathbf{e}_i\}$ ,  $i = 1, \dots, n$ . Then we have

$$\mathbf{v} = \sum_{i=1}^n v_i \mathbf{e}_i.$$

<sup>6</sup> In section we define a linear operator using two conditions,  $L(\mathbf{a} + \mathbf{b}) = L(\mathbf{a}) + L(\mathbf{b})$  and  $L(\alpha\mathbf{a}) = \alpha L(\mathbf{a})$ . The reader can show that this is equivalent to the condition presented here. Furthermore, all linear transformations take the origin to the origin,  $L(\mathbf{0}) = \mathbf{0}$ .

Now consider the effect of the transformation  $L$  on  $\mathbf{v}$ , using the linearity property:

$$L(\mathbf{v}) = L\left(\sum_{i=1}^n v_i \mathbf{e}_i\right) = \sum_{i=1}^n v_i L(\mathbf{e}_i). \quad (3.35)$$

Thus, we see that determining how  $L$  acts on  $\mathbf{v}$  requires that we know how  $L$  acts on the basis vectors. Namely, we need  $L(\mathbf{e}_i)$ . Since  $\mathbf{e}_i$  is a vector, this produces another vector in the space. But the resulting vector can be expanded in the basis. Let's assume that the resulting vector takes the form

$$L(\mathbf{e}_i) = \sum_{j=1}^n L_{ji} \mathbf{e}_j, \quad (3.36)$$

where  $L_{ji}$  is the  $j$ th component of  $L(\mathbf{e}_i)$  for each  $i = 1, \dots, n$ . The matrix of  $L_{ji}$ 's is called the matrix representation of the operator  $L$ .

Typically, in a linear algebra class you start with matrices and do not see this connection to linear operators. However, there will be times that you will need this connection to understand why matrices are involved. Furthermore, the matrix representation depends on the basis used. We used the standard basis above. However, you could have started with a different basis, such as dictated by another coordinate system. We will not go further into this point at this time and just stick with the standard basis.

**Example 3.5.** Consider the linear transformation of  $\mathbf{u} = (u, v)$  into  $\mathbf{x} = (x, y)$  by

$$L(u, v) = (3u - v, v + u) = (x, y).$$

The matrix representation for this transformation is found by considering how  $L$  acts on the basis vectors. We have  $L(1, 0) = (3, 1)$  and  $L(0, 1) = (-1, 1)$ . Thus, the representation is given as

$$L = \begin{pmatrix} 3 & -1 \\ 1 & 1 \end{pmatrix}.$$

Now that we know how  $L$  acts on basis vectors, what does this have to say about how  $L$  acts on any other vector in the space? We insert expression (3.36) into Equation (3.35). Then we find

$$\begin{aligned} L(\mathbf{v}) &= \sum_{i=1}^n v_i L(\mathbf{e}_i) \\ &= \sum_{i=1}^n v_i \left( \sum_{j=1}^n L_{ji} \mathbf{e}_j \right) \\ &= \sum_{j=1}^n \left( \sum_{i=1}^n v_i L_{ji} \right) \mathbf{e}_j. \end{aligned} \quad (3.37)$$

Since  $L(\mathbf{v}) = \mathbf{u}$ , we see that the  $j$ th component of  $\mathbf{u}$  can be written as

$$u_j = \sum_{i=1}^n L_{ji}v_i, \quad j = 1 \dots n. \quad (3.38)$$

This equation can be written in matrix form as

$$\mathbf{u} = L\mathbf{v},$$

where  $L$  now takes the role of a matrix. It is similar to the multiplication of the rotation matrix times a vector as seen in the last section. We will just work with matrix representations from here on.

**Example 3.6.** For the transformation  $L(u, v) = (3u - v, v + u) = (x, y)$  in the last example, what does  $\mathbf{v} = 5\mathbf{i} + 3\mathbf{j}$  get mapped into? We know the matrix representation from the previous example, so we have

$$\mathbf{u} = \begin{pmatrix} 3 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 5 \\ 3 \end{pmatrix} = \begin{pmatrix} 12 \\ 2 \end{pmatrix}.$$

Next, we can compose transformations like we had done with the two rotation matrices. Let  $\mathbf{u} = A(\mathbf{v})$  and  $\mathbf{w} = B(\mathbf{u})$  for two transformations  $A$  and  $B$ . (Thus,  $\mathbf{v} \rightarrow \mathbf{u} \rightarrow \mathbf{w}$ .) Then a composition of these transformations is given by

$$\mathbf{w} = B(\mathbf{u}) = B(A\mathbf{v}).$$

This can be viewed as a transformation from  $\mathbf{v}$  to  $\mathbf{w}$  as

$$\mathbf{w} = BA(\mathbf{v}),$$

where the matrix representation of  $BA$  is given by the product of the matrix representations of  $A$  and  $B$ .

To see this, we look at the  $ij$ th element of the matrix representation of  $BA$ . We first note that the transformation from  $\mathbf{v}$  to  $\mathbf{w}$  is given by

$$w_i = \sum_{j=1}^n (BA)_{ij}v_j. \quad (3.39)$$

However, if we use the successive transformations, we have

$$\begin{aligned} w_i &= \sum_{k=1}^n B_{ik}u_k \\ &= \sum_{k=1}^n B_{ik} \left( \sum_{j=1}^n A_{kj}v_j \right) \\ &= \sum_{j=1}^n \left( \sum_{k=1}^n B_{ik}A_{kj} \right) v_j. \end{aligned} \quad (3.40)$$

We have two expressions for  $w_i$  as sums over  $v_j$ . So, the coefficients must be equal. This leads to our result:

$$(BA)_{ij} = \sum_{k=1}^n B_{ik}A_{kj}. \quad (3.41)$$

Thus, we have found the component form of matrix multiplication, which resulted from the composition of two linear transformations. This agrees with our earlier example of matrix multiplication: The  $ij$ -th component of the product is obtained by multiplying elements in the  $i$ th row of  $B$  and the  $j$ th column of  $A$  and summing.

**Example 3.7.** Consider the rotation in two dimensions of the axes by an angle  $\theta$ . Now apply the scaling transformation<sup>7</sup>

$$L_s = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}.$$

<sup>7</sup>This scaling transformation will rescale  $x$ -components by  $a$  and  $y$ -components by  $b$ . If either is negative, it will also provide an additional reflection.

what is the matrix representation of this combination of transformations? The result is a simple product of the matrix representations (in reverse order of application):

$$L_s \hat{R} = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} = \begin{pmatrix} a \cos \theta & a \sin \theta \\ -b \sin \theta & b \cos \theta \end{pmatrix}.$$

There are many other properties of matrices and types of matrices that one may encounter. We will list a few.

First of all, there is the  $n \times n$  identity matrix,  $I$ . The identity is defined as that matrix satisfying

$$IA = AI = A \quad (3.42)$$

Identity matrix.

for any  $n \times n$  matrix  $A$ . The  $n \times n$  identity matrix takes the form

$$I = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \vdots & 1 \end{pmatrix} \quad (3.43)$$

A component form is given by the Kronecker delta. Namely, we have that

Kronecker delta,  $\delta_{ij}$ .

$$I_{ij} = \delta_{ij} \equiv \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases} \quad (3.44)$$

The inverse of matrix  $A$  is that matrix  $A^{-1}$  such that

$$AA^{-1} = A^{-1}A = I. \quad (3.45)$$

There is a systematic method for determining the inverse in terms of cofactors, which we describe a little later. However, the inverse of a  $2 \times 2$  matrix is easily obtained without learning about cofactors. Let

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Now consider the matrix

$$B = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Multiplying these matrices, we find that

$$AB = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} ad - bc & 0 \\ 0 & ad - bc \end{pmatrix}.$$

This is not quite the identity, but it is a multiple of the identity. We just need to divide by  $ad - bc$ . So, we have found the inverse matrix:

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Inverse of a  $2 \times 2$  matrix.

We leave it to the reader to show that  $A^{-1}A = I$ .

The factor  $ad - bc$  is the difference in the products of the diagonal and off-diagonal elements of matrix  $A$ . This factor is called the *determinant* of  $A$ . It is denoted as  $\det(A)$ ,  $\det A$  or  $|A|$ . Thus, we define

$$\det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc. \quad (3.46)$$

For higher dimensional matrices one can write the definition of the determinant. We will for now just indicate the process for  $3 \times 3$  matrices. We write matrix  $A$  as

Determinant of a  $3 \times 3$  matrix.

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}. \quad (3.47)$$

The determinant of  $A$  can be computed in terms of simpler  $2 \times 2$  determinants. We define

$$\begin{aligned} \det A &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\ &= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}. \end{aligned} \quad (3.48)$$

There are many other properties of determinants. For example, if two rows, or columns, of a matrix are multiples of each other, then  $\det A = 0$ . If one multiplies one row, or column, of a matrix by a constant,  $k$ , then the determinant of the matrix is multiplied by  $k$ .

If  $\det A = 0$ ,  $A$  is called a *singular* matrix. Otherwise, it is called *nonsingular*. If a matrix is nonsingular, then the inverse exists. From our example for a general  $2 \times 2$  system, the inverse exists if  $ad - bc \neq 0$ .

Computing the inverse of a larger matrix is a little more complicated. One first constructs the matrix of cofactors. The  $ij$ -th cofactor is obtained by computing the determinant of the matrix resulting from eliminating the  $i$ th row and  $j$ th column of  $A$  and multiplying by either  $+1$  or  $-1$ . Thus,

$$C_{ij} = (-1)^{i+j} \det (a_{ij}).$$

The matrix of cofactors.

Then, the inverse matrix is obtained by dividing the transpose of the matrix of cofactors by the determinant of  $A$ . Thus,

$$(A^{-1})_{ij} = \frac{C_{ji}}{\det A}.$$

This is best shown by example.

**Example 3.8.** Find the inverse of the matrix

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 3 & 2 \\ 1 & -2 & 1 \end{pmatrix}.$$

The determinant of this matrix is easily found as

$$\det A = \begin{vmatrix} 1 & 2 & -1 \\ 0 & 3 & 2 \\ 1 & -2 & 1 \end{vmatrix} = 1 \begin{vmatrix} 3 & 2 \\ -2 & 1 \end{vmatrix} + 1 \begin{vmatrix} 2 & -1 \\ 3 & 2 \end{vmatrix} = 14.$$

Next, we construct the matrix of cofactors:

$$C_{ij} = \begin{pmatrix} + \begin{vmatrix} 3 & 2 \\ -2 & 1 \end{vmatrix} & - \begin{vmatrix} 0 & 2 \\ 1 & 1 \end{vmatrix} & + \begin{vmatrix} 0 & 3 \\ 1 & -2 \end{vmatrix} \\ - \begin{vmatrix} 2 & -1 \\ -2 & 1 \end{vmatrix} & + \begin{vmatrix} 1 & -1 \\ 1 & 1 \end{vmatrix} & - \begin{vmatrix} 1 & 2 \\ 1 & -2 \end{vmatrix} \\ + \begin{vmatrix} 2 & -1 \\ 3 & 2 \end{vmatrix} & - \begin{vmatrix} 1 & -1 \\ 0 & 2 \end{vmatrix} & + \begin{vmatrix} 1 & 2 \\ 0 & 3 \end{vmatrix} \end{pmatrix}.$$

Computing the  $2 \times 2$  determinants, we obtain

$$C_{ij} = \begin{pmatrix} 7 & -2 & -3 \\ 0 & 2 & 4 \\ 7 & -2 & 3 \end{pmatrix}.$$

Finally, we compute the inverse as

$$\begin{aligned}
 A^{-1} &= \frac{1}{14} \begin{pmatrix} 7 & -2 & -3 \\ 0 & 2 & 4 \\ 7 & -2 & 3 \end{pmatrix}^T \\
 &= \frac{1}{14} \begin{pmatrix} 7 & 0 & 7 \\ -2 & 2 & -2 \\ -3 & 4 & 3 \end{pmatrix} \\
 &= \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ -\frac{1}{7} & \frac{1}{7} & -\frac{1}{7} \\ -\frac{3}{14} & \frac{2}{7} & \frac{3}{14} \end{pmatrix}. \tag{3.49}
 \end{aligned}$$

Another operation that we have seen earlier is the *transpose* of a matrix. The transpose of a matrix is a new matrix in which the rows and columns are interchanged. If write an  $n \times m$  matrix  $A$  in standard form as

Matrix transpose.

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix}, \tag{3.50}$$

then the transpose is defined as

$$A^T = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1m} & a_{2m} & \dots & a_{nm} \end{pmatrix}. \tag{3.51}$$

In index form, we have

$$(A^T)_{ij} = A_{ji}, \quad i, j = 1, \dots, n.$$

As we had seen in the last section, a matrix satisfying

$$A^T = A^{-1}, \quad \text{or} \quad AA^T = A^T A = I,$$

is called an orthogonal matrix. One also can show that

$$(AB)^T = B^T A^T.$$

Finally, the *trace* of a square matrix is the sum of its diagonal elements:

Trace of a matrix.

$$\text{Tr}(A) = a_{11} + a_{22} + \dots + a_{nn} = \sum_{i=1}^n a_{ii}.$$

We can show that for two square matrices

$$\text{Tr}(AB) = \text{Tr}(BA).$$

A standard application of determinants is the solution of a system of linear algebraic equations using Cramer's Rule. As an example, we consider a simple system of two equations and two unknowns. Let's consider this system of two equations and two unknowns,  $x$  and  $y$ , in the form

$$\begin{aligned} ax + by &= e, \\ cx + dy &= f. \end{aligned} \quad (3.52)$$

The standard way to solve this is to eliminate one of the variables. (Just imagine dealing with a bigger system!). So, we can eliminate the  $x$ 's. Multiply the first equation by  $c$  and the second equation by  $a$  and subtract. We then get

$$(bc - ad)y = (ec - fa).$$

If  $bc - ad \neq 0$ , then we can solve to  $y$ , getting

$$y = \frac{ec - fa}{bc - ad}$$

. Similarly, we find

$$x = \frac{ed - bf}{ad - bc}.$$

We note the the denominators can be replaced with the determinant of the matrix of coefficients,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

In fact, we can also replace each numerator with a determinant. Thus, our solutions may be written as

$$\begin{aligned} x &= \frac{\begin{vmatrix} e & b \\ f & d \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}} \\ y &= \frac{\begin{vmatrix} a & e \\ c & f \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}}. \end{aligned} \quad (3.53)$$

This is Cramer's Rule for writing out solutions of systems of equations. Note that each variable is determined by placing a determinant with  $e$  and  $f$  placed in the column of the coefficient matrix corresponding to the order of the variable in the equation. The denominator is

Cramer's Rule for solving algebraic systems of equations.

the determinant of the coefficient matrix. This construction is easily extended to larger systems of equations.

Cramer's Rule can be extended to higher dimensional systems. As an example, we now solve a system of three equations and three unknowns.

**Example 3.9.** *Solve the system of equations*

$$\begin{aligned}x + 2y - z &= 1, \\3y + 2z &= 2 \\x - 2y + z &= 0.\end{aligned}\tag{3.54}$$

First, one writes the system in the form  $L\mathbf{x} = \mathbf{b}$ , where  $L$  is the coefficient matrix

$$L = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 3 & 2 \\ 1 & -2 & 1 \end{pmatrix}$$

and

$$\mathbf{b} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}.$$

The solution is generally,  $\mathbf{x} = L^{-1}\mathbf{b}$  if  $L^{-1}$  exists. So, we check that  $\det L = 14 \neq 0$ . Thus,  $L$  is nonsingular and its inverse exists.

So, the solution of this system of three equations and three unknowns can now be found using Cramer's rule. Thus, we have

$$\begin{aligned}x &= \frac{\begin{vmatrix} 1 & 2 & -1 \\ 2 & 3 & 2 \\ 0 & -2 & 1 \end{vmatrix}}{\det L} = \frac{7}{14} = \frac{1}{2}, \\y &= \frac{\begin{vmatrix} 1 & 1 & -1 \\ 0 & 2 & 2 \\ 1 & 0 & 1 \end{vmatrix}}{\det L} = \frac{6}{14} = \frac{3}{7}, \\z &= \frac{\begin{vmatrix} 1 & 2 & 1 \\ 0 & 3 & 2 \\ 1 & -2 & 0 \end{vmatrix}}{\det L} = \frac{5}{14}.\end{aligned}\tag{3.55}$$

We end this section by summarizing the rule for the existence of solutions of systems of algebraic equations,  $L\mathbf{x} = \mathbf{b}$ .

1. If  $\det L \neq 0$ , then there exists a unique solution,  $\mathbf{x} = L^{-1}\mathbf{b}$ . In particular, if  $\mathbf{b} = \mathbf{0}$ , the system is homogeneous and only has the trivial solution,  $\mathbf{x} = \mathbf{0}$ .

2. If  $\det L = 0$ , then the system does not have a unique solution. Either there is no solution, or an infinite number of solutions. For example, the system

$$\begin{aligned} 2x + y &= 5, \\ 4x + 2y &= 2, \end{aligned} \quad (3.56)$$

has no solutions, while

$$\begin{aligned} 2x + y &= 0, \\ 4x + 2y &= 0, \end{aligned} \quad (3.57)$$

has an infinite number of solutions ( $y = -2x$ ).

### 3.4 Eigenvalue Problems

#### 3.4.1 An Introduction to Coupled Systems

RECALL THAT one of the reasons we have seemingly digressed into topics in linear algebra and matrices is to solve a coupled system of differential equations. The simplest example is a system of linear differential equations of the form

$$\begin{aligned} \frac{dx}{dt} &= ax + by \\ \frac{dy}{dt} &= cx + dy. \end{aligned} \quad (3.58)$$

We note that this system is coupled. We cannot solve either equation without knowing either  $x(t)$  or  $y(t)$ . A much easier problem would be to solve an uncoupled system like

Uncoupled system.

$$\begin{aligned} \frac{dx}{dt} &= \lambda_1 x \\ \frac{dy}{dt} &= \lambda_2 y. \end{aligned} \quad (3.59)$$

The solutions are quickly found to be

$$\begin{aligned} x(t) &= c_1 e^{\lambda_1 t}, \\ y(t) &= c_2 e^{\lambda_2 t}. \end{aligned} \quad (3.60)$$

Here  $c_1$  and  $c_2$  are two arbitrary constants.

We can determine particular solutions of the system by specifying  $x(t_0) = x_0$  and  $y(t_0) = y_0$  at some time  $t_0$ . Thus,

$$\begin{aligned} x(t) &= x_0 e^{\lambda_1 t}, \\ y(t) &= y_0 e^{\lambda_2 t}. \end{aligned} \quad (3.61)$$

Wouldn't it be nice if we could transform the more general system into one that is not coupled? Let's write these systems in more general form. We write the coupled system as

$$\frac{d}{dt}\mathbf{x} = A\mathbf{x}$$

and the uncoupled system as

$$\frac{d}{dt}\mathbf{y} = \Lambda\mathbf{y},$$

where

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

We note that  $\Lambda$  is a diagonal matrix.

Now, we seek a transformation between  $\mathbf{x}$  and  $\mathbf{y}$  that will transform the coupled system into the uncoupled system. Thus, we define the transformation

$$\mathbf{x} = S\mathbf{y}. \quad (3.62)$$

Inserting this transformation into the coupled system we have

$$\begin{aligned} \frac{d}{dt}\mathbf{x} &= A\mathbf{x} \Rightarrow \\ \frac{d}{dt}S\mathbf{y} &= AS\mathbf{y} \Rightarrow \\ S\frac{d}{dt}\mathbf{y} &= AS\mathbf{y}. \end{aligned} \quad (3.63)$$

Multiply both sides by  $S^{-1}$ . [We can do this if we are dealing with an invertible transformation; i.e., a transformation in which we can get  $\mathbf{y}$  from  $\mathbf{x}$  as  $\mathbf{y} = S^{-1}\mathbf{x}$ .] We obtain

$$\frac{d}{dt}\mathbf{y} = S^{-1}AS\mathbf{y}.$$

Noting that

$$\frac{d}{dt}\mathbf{y} = \Lambda\mathbf{y},$$

we have

$$\Lambda = S^{-1}AS. \quad (3.64)$$

The expression  $S^{-1}AS$  is called a *similarity transformation* of matrix  $A$ . So, in order to uncouple the system, we seek a similarity transformation that results in a diagonal matrix. This process is called the *diagonalization* of matrix  $A$ . We do not know  $S$ , nor do we know  $\Lambda$ . We can rewrite this equation as

$$AS = S\Lambda.$$

We can solve this equation if  $S$  is *real symmetric*, i.e.  $S^T = S$ . [In the case of complex matrices, we need the matrix to be Hermitian,  $\bar{S}^T = S$  where the bar denotes complex conjugation. Further discussion of diagonalization is left for the end of the chapter.]

We first show that  $S\Lambda = \Lambda S$ . We look at the  $ij$ th component of  $S\Lambda$  and rearrange the terms in the matrix product.

$$\begin{aligned}
 (S\Lambda)_{ij} &= \sum_{k=1}^n S_{ik}\Lambda_{kj} \\
 &= \sum_{k=1}^n S_{ik}\lambda_j I_{kj} \\
 &= \sum_{k=1}^n \lambda_j I_{jk} S_{ki}^T \\
 &= \sum_{k=1}^n \Lambda_{jk} S_{ki} \\
 &= (\Lambda S)_{ij}
 \end{aligned} \tag{3.65}$$

This result leads us to the fact that  $S$  satisfies the equation

$$AS = \Lambda S.$$

Therefore, one has that the columns of  $S$  (denoted  $\mathbf{v}$ ) satisfy an equation of the form

$$A\mathbf{v} = \lambda\mathbf{v}. \tag{3.66}$$

This is an equation for vectors  $\mathbf{v}$  and numbers  $\lambda$  given matrix  $A$ . It is called an *eigenvalue problem*. The vectors are called *eigenvectors* and the numbers,  $\lambda$ , are called *eigenvalues*. In principle, we can solve the eigenvalue problem and this will lead us to solutions of the uncoupled system of differential equations.

### 3.4.2 Example of an Eigenvalue Problem

WE WILL DETERMINE the eigenvalues and eigenvectors for

$$A = \begin{pmatrix} 1 & -2 \\ -3 & 2 \end{pmatrix}$$

In order to find the eigenvalues and eigenvectors of this equation, we need to solve

$$A\mathbf{v} = \lambda\mathbf{v}. \tag{3.67}$$

Let  $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ . Then the eigenvalue problem can be written out. We have that

$$A\mathbf{v} = \lambda\mathbf{v}$$

$$\begin{aligned} \begin{pmatrix} 1 & -2 \\ -3 & 2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} &= \lambda \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \\ \begin{pmatrix} v_1 - 2v_2 \\ -3v_1 + 2v_2 \end{pmatrix} &= \begin{pmatrix} \lambda v_1 \\ \lambda v_2 \end{pmatrix}. \end{aligned} \quad (3.68)$$

So, we see that the system becomes

$$\begin{aligned} v_1 - 2v_2 &= \lambda v_1, \\ -3v_1 + 2v_2 &= \lambda v_2. \end{aligned} \quad (3.69)$$

This can be rewritten as

$$\begin{aligned} (1 - \lambda)v_1 - 2v_2 &= 0, \\ -3v_1 + (2 - \lambda)v_2 &= 0. \end{aligned} \quad (3.70)$$

This is a homogeneous system. We can try to solve it using elimination, as we had done earlier when deriving Cramer's Rule. We find that multiplying the first equation by  $2 - \lambda$ , the second by 2 and adding, we get

$$[(1 - \lambda)(2 - \lambda) - 6]v_1 = 0.$$

If the factor in the brackets is not zero, we obtain  $v_1 = 0$ . Inserting this into the system gives  $v_2 = 0$  as well. Thus, we find  $\mathbf{v}$  is the zero vector. However, this does not get us anywhere. We could have guessed this solution. This simple solution is the solution of all eigenvalue problems and is called the trivial solution. When solving eigenvalue problems, we only look for nontrivial solutions!

So, we have to stipulate that the factor in the brackets is zero. This means that  $v_1$  is still unknown. This situation will always occur for eigenvalue problems. The general eigenvalue problem can be written as

$$A\mathbf{v} - \lambda\mathbf{v} = 0,$$

or by inserting the identity matrix,

$$A\mathbf{v} - \lambda I\mathbf{v} = 0.$$

Finally, we see that we always get a homogeneous system,

$$(A - \lambda I)\mathbf{v} = 0.$$

The factor that has to be zero can be seen now as the determinant of this system. Thus, we require

$$\det(A - \lambda I) = 0. \quad (3.71)$$

We write out this condition for the example at hand. We have that

$$\begin{vmatrix} 1 - \lambda & -2 \\ -3 & 2 - \lambda \end{vmatrix} = 0.$$

This will always be the starting point in solving eigenvalue problems. Note that the matrix is  $A$  with  $\lambda$ 's subtracted from the diagonal elements.

Computing the determinant, we have

$$(1 - \lambda)(2 - \lambda) - 6 = 0,$$

or

$$\lambda^2 - 3\lambda - 4 = 0.$$

We therefore have obtained a condition on the eigenvalues! It is a quadratic and we can factor it:

$$(\lambda - 4)(\lambda + 1) = 0.$$

So, our eigenvalues are  $\lambda = 4, -1$ .

The second step is to find the eigenvectors. We have to do this for each eigenvalue. We first insert  $\lambda = 4$  into our system:

$$\begin{aligned} -3v_1 - 2v_2 &= 0, \\ -3v_1 - 2v_2 &= 0. \end{aligned} \tag{3.72}$$

Note that these equations are the same. So, we have one equation in two unknowns. We will not get a unique solution. This is typical of eigenvalue problems. We can pick anything we want for  $v_2$  and then determine  $v_1$ . For example,  $v_2 = 1$  gives  $v_1 = -2/3$ . A nicer solution would be  $v_2 = 3$  and  $v_1 = -2$ . These vectors are different, but they point in the same direction in the  $v_1v_2$  plane.

For  $\lambda = -1$ , the system becomes

$$\begin{aligned} 2v_1 - 2v_2 &= 0, \\ -3v_1 + 3v_2 &= 0. \end{aligned} \tag{3.73}$$

While these equations do not at first look the same, we can divide out the constants and see that once again we get the same equation,

$$v_1 = v_2.$$

Picking  $v_2 = 1$ , we get  $v_1 = 1$ .

In summary, the solution to our eigenvalue problem is

$$\begin{aligned} \lambda = 4, \quad \mathbf{v} &= \begin{pmatrix} -2 \\ 3 \end{pmatrix} \\ \lambda = -1, \quad \mathbf{v} &= \begin{pmatrix} 1 \\ 1 \end{pmatrix} \end{aligned}$$

### 3.4.3 Eigenvalue Problems - A Summary

IN THE LAST SUBSECTION we were introduced to eigenvalue problems as a way to obtain a solution to a coupled system of linear differential equations. Eigenvalue problems appear in many contexts in physical applications. In this section we will summarize the method of solution of eigenvalue problems based upon our discussion in the last section. In the next subsection we will look at another problem that is a bit more geometric and will give us more insight into the process of diagonalization. We will return to our coupled system in a later section and provide more examples of solving eigenvalue problems.

We seek *nontrivial solutions* to the eigenvalue problem

$$A\mathbf{v} = \lambda\mathbf{v}. \quad (3.74)$$

We note that  $\mathbf{v} = \mathbf{0}$  is an obvious solution. Furthermore, it does not lead to anything useful. So, it is a trivial solution. Typically, we are given the matrix  $A$  and have to determine the eigenvalues,  $\lambda$ , and the associated eigenvectors,  $\mathbf{v}$ , satisfying the above eigenvalue problem. Later in the course we will explore other types of eigenvalue problems.

For now we begin to solve the eigenvalue problem for  $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ . Inserting this into Equation (3.74), we obtain the homogeneous algebraic system

$$\begin{aligned} (a - \lambda)v_1 + bv_2 &= 0, \\ cv_1 + (d - \lambda)v_2 &= 0. \end{aligned} \quad (3.75)$$

The solution of such a system would be unique if the determinant of the system is not zero. However, this would give the trivial solution  $v_1 = 0, v_2 = 0$ . To get a nontrivial solution, we need to force the determinant to be zero. This yields the eigenvalue equation

$$0 = \begin{vmatrix} a - \lambda & b \\ c & d - \lambda \end{vmatrix} = (a - \lambda)(d - \lambda) - bc.$$

This is a quadratic equation for the eigenvalues that would lead to nontrivial solutions. If we expand the right side of the equation, we find that

$$\lambda^2 - (a + d)\lambda + ad - bc = 0.$$

This is the same equation as the characteristic equation for the general constant coefficient differential equation considered in the last chapter as we will later show in Equation (2.113). Thus, the eigenvalues correspond to the solutions of the characteristic polynomial for the system.

Once we find the eigenvalues, then there are possibly an infinite number solutions to the algebraic system. We will see this in the examples.

The method for solving eigenvalue problems, as you have seen, consists of just a few simple steps. We list these steps as follows:

<b>Solving Eigenvalue Problems</b>
a) Write the coefficient matrix; b) Find the eigenvalues from the equation $\det(A - \lambda I) = 0$ ; and, c) Solve the linear system $(A - \lambda I)\mathbf{v} = 0$ for each $\lambda$ .

### 3.5 Matrix Formulation of Planar Systems

WE HAVE INVESTIGATED several linear systems in the plane in the last chapter. However, we need a deeper insight into the solutions of planar systems. So, in this section we will recast the first order linear systems into matrix form. This will lead to a better understanding of first order systems and allow for extensions to higher dimensions and the solution of nonhomogeneous equations. In particular, we can see that the solutions obtained for planar systems in the last chapters are intimately connected to the underlying eigenvalue problems.

We start with the usual homogeneous system in Equation (2.110). Let the unknowns be represented by the vector

$$\mathbf{x}(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}.$$

Then we have that

$$\mathbf{x}' = \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \equiv A\mathbf{x}.$$

Here we have introduced the *coefficient matrix*  $A$ . This is a first order vector differential equation,

$$\mathbf{x}' = A\mathbf{x}.$$

Formerly, we can write the solution as<sup>8</sup>

$$\mathbf{x} = \mathbf{x}_0 e^{At}.$$

We would like to investigate the solution of our system. Our investigations will lead to new techniques for solving linear systems using matrix methods.

<sup>8</sup> The exponential of a matrix is defined using the Maclaurin series expansion

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

So, we define

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!} = I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \dots \quad (3.76)$$

In general, it is difficult computing  $e^A$  unless  $A$  is diagonal.

We begin by recalling the solution to the specific problem (2.117). We obtained the solution to this system as

$$\begin{aligned}x(t) &= c_1 e^t + c_2 e^{-4t}, \\y(t) &= \frac{1}{3}c_1 e^t - \frac{1}{2}c_2 e^{-4t}.\end{aligned}\tag{3.77}$$

This can be rewritten using matrix operations. Namely, we first write the solution in vector form.

$$\begin{aligned}\mathbf{x} &= \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} \\&= \begin{pmatrix} c_1 e^t + c_2 e^{-4t} \\ \frac{1}{3}c_1 e^t - \frac{1}{2}c_2 e^{-4t} \end{pmatrix} \\&= \begin{pmatrix} c_1 e^t \\ \frac{1}{3}c_1 e^t \end{pmatrix} + \begin{pmatrix} c_2 e^{-4t} \\ -\frac{1}{2}c_2 e^{-4t} \end{pmatrix} \\&= c_1 \begin{pmatrix} 1 \\ \frac{1}{3} \end{pmatrix} e^t + c_2 \begin{pmatrix} 1 \\ -\frac{1}{2} \end{pmatrix} e^{-4t}.\end{aligned}\tag{3.78}$$

We see that our solution is in the form of a linear combination of vectors of the form

$$\mathbf{x} = \mathbf{v}e^{\lambda t}$$

with  $\mathbf{v}$  a constant vector and  $\lambda$  a constant number. This is similar to how we began to find solutions to second order constant coefficient equations. So, for the general problem (3.5) we insert this guess. Thus,

$$\begin{aligned}\mathbf{x}' &= A\mathbf{x} \Rightarrow \\ \lambda \mathbf{v}e^{\lambda t} &= A\mathbf{v}e^{\lambda t}.\end{aligned}\tag{3.79}$$

For this to be true for all  $t$ , we have that

$$A\mathbf{v} = \lambda \mathbf{v}.\tag{3.80}$$

This is an eigenvalue problem.  $A$  is a  $2 \times 2$  matrix for our problem, but could easily be generalized to a system of  $n$  first order differential equations. We will confine our remarks for now to planar systems. However, we need to recall how to solve eigenvalue problems and then see how solutions of eigenvalue problems can be used to obtain solutions to our systems of differential equations.

### 3.5.1 Solving Constant Coefficient Systems in 2D

BEFORE PROCEEDING TO EXAMPLES, we first indicate the types of solutions that could result from the solution of a homogeneous, constant coefficient system of first order differential equations.

We begin with the linear system of differential equations in matrix form.

$$\frac{d\mathbf{x}}{dt} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mathbf{x} = A\mathbf{x}. \quad (3.81)$$

The type of behavior depends upon the eigenvalues of matrix  $A$ . The procedure is to determine the eigenvalues and eigenvectors and use them to construct the general solution.

If we have an initial condition,  $\mathbf{x}(t_0) = \mathbf{x}_0$ , we can determine the two arbitrary constants in the general solution in order to obtain the particular solution. Thus, if  $\mathbf{x}_1(t)$  and  $\mathbf{x}_2(t)$  are two linearly independent solutions<sup>4</sup>, then the general solution is given as

$$\mathbf{x}(t) = c_1\mathbf{x}_1(t) + c_2\mathbf{x}_2(t).$$

<sup>4</sup> Recall that linear independence means  $c_1\mathbf{x}_1(t) + c_2\mathbf{x}_2(t) = \mathbf{0}$  if and only if  $c_1, c_2 = 0$ . The reader should derive the condition on the  $\mathbf{x}_i$  for linear independence.

Then, setting  $t = 0$ , we get two linear equations for  $c_1$  and  $c_2$ :

$$c_1\mathbf{x}_1(0) + c_2\mathbf{x}_2(0) = \mathbf{x}_0.$$

The major work is in finding the linearly independent solutions. This depends upon the different types of eigenvalues that one obtains from solving the eigenvalue equation,  $\det(A - \lambda I) = 0$ . The nature of these roots indicate the form of the general solution. On the next page we summarize the classification of solutions in terms of the eigenvalues of the coefficient matrix. We first make some general remarks about the plausibility of these solutions and then provide examples in the following section to clarify the matrix methods for our two dimensional systems.

The construction of the general solution in Case I is straight forward. However, the other two cases need a little explanation.

**Classification of the Solutions for Two  
Linear First Order Differential Equations**

**1. Case I: Two real, distinct roots.**

Solve the eigenvalue problem  $A\mathbf{v} = \lambda\mathbf{v}$  for each eigenvalue obtaining two eigenvectors  $\mathbf{v}_1, \mathbf{v}_2$ . Then write the general solution as a linear combination  $\mathbf{x}(t) = c_1e^{\lambda_1 t}\mathbf{v}_1 + c_2e^{\lambda_2 t}\mathbf{v}_2$

**2. Case II: One Repeated Root**

Solve the eigenvalue problem  $A\mathbf{v} = \lambda\mathbf{v}$  for one eigenvalue  $\lambda$ , obtaining the first eigenvector  $\mathbf{v}_1$ . One then needs a second linearly independent solution. This is obtained by solving the nonhomogeneous problem  $A\mathbf{v}_2 - \lambda\mathbf{v}_2 = \mathbf{v}_1$  for  $\mathbf{v}_2$ .

The general solution is then given by  $\mathbf{x}(t) = c_1e^{\lambda t}\mathbf{v}_1 + c_2e^{\lambda t}(\mathbf{v}_2 + t\mathbf{v}_1)$ .

**3. Case III: Two complex conjugate roots.**

Solve the eigenvalue problem  $A\mathbf{x} = \lambda\mathbf{x}$  for one eigenvalue,  $\lambda = \alpha + i\beta$ , obtaining one eigenvector  $\mathbf{v}$ . Note that this eigenvector may have complex entries. Thus, one can write the vector

$$\mathbf{y}(t) = e^{\lambda t}\mathbf{v} = e^{\alpha t}(\cos \beta t + i \sin \beta t)\mathbf{v}.$$

Now, construct two linearly independent solutions to the problem using the real and imaginary parts of  $\mathbf{y}(t)$  :

$$\mathbf{y}_1(t) = \text{Re}(\mathbf{y}(t)) \text{ and } \mathbf{y}_2(t) = \text{Im}(\mathbf{y}(t)).$$

Then the general solution can be written as  $\mathbf{x}(t) = c_1\mathbf{y}_1(t) + c_2\mathbf{y}_2(t)$ .

Let's consider Case III. Note that since the original system of equations does not have any  $i$ 's, then we would expect real solutions. So, we look at the real and imaginary parts of the complex solution. We have that the complex solution satisfies the equation

$$\frac{d}{dt} [\text{Re}(\mathbf{y}(t)) + i\text{Im}(\mathbf{y}(t))] = A[\text{Re}(\mathbf{y}(t)) + i\text{Im}(\mathbf{y}(t))].$$

Differentiating the sum and splitting the real and imaginary parts of the equation, gives

$$\frac{d}{dt}\text{Re}(\mathbf{y}(t)) + i\frac{d}{dt}\text{Im}(\mathbf{y}(t)) = A[\text{Re}(\mathbf{y}(t))] + iA[\text{Im}(\mathbf{y}(t))].$$

Setting the real and imaginary parts equal, we have

$$\frac{d}{dt}\text{Re}(\mathbf{y}(t)) = A[\text{Re}(\mathbf{y}(t))],$$

and

$$\frac{d}{dt}\text{Im}(\mathbf{y}(t)) = A[\text{Im}(\mathbf{y}(t))].$$

Therefore, the real and imaginary parts each are linearly independent solutions of the system and the general solution can be written as a linear combination of these expressions.

We now turn to Case II. Writing the system of first order equations as a second order equation for  $x(t)$  with the sole solution of the characteristic equation,  $\lambda = \frac{1}{2}(a + d)$ , we have that the general solution takes the form

$$x(t) = (c_1 + c_2 t)e^{\lambda t}.$$

This suggests that the second linearly independent solution involves a term of the form  $\mathbf{v}te^{\lambda t}$ . It turns out that the guess that works is

$$\mathbf{x} = te^{\lambda t}\mathbf{v}_1 + e^{\lambda t}\mathbf{v}_2.$$

Inserting this guess into the system  $\mathbf{x}' = A\mathbf{x}$  yields

$$\begin{aligned} (te^{\lambda t}\mathbf{v}_1 + e^{\lambda t}\mathbf{v}_2)' &= A [te^{\lambda t}\mathbf{v}_1 + e^{\lambda t}\mathbf{v}_2]. \\ e^{\lambda t}\mathbf{v}_1 + \lambda te^{\lambda t}\mathbf{v}_1 + \lambda e^{\lambda t}\mathbf{v}_2 &= \lambda te^{\lambda t}\mathbf{v}_1 + e^{\lambda t}A\mathbf{v}_2. \\ e^{\lambda t}(\mathbf{v}_1 + \lambda\mathbf{v}_2) &= e^{\lambda t}A\mathbf{v}_2. \end{aligned} \quad (3.82)$$

Noting this is true for all  $t$ , we find that

$$\mathbf{v}_1 + \lambda\mathbf{v}_2 = A\mathbf{v}_2. \quad (3.83)$$

Therefore,

$$(A - \lambda I)\mathbf{v}_2 = \mathbf{v}_1.$$

We know everything except for  $\mathbf{v}_2$ . So, we just solve for it and obtain the second linearly independent solution.

### 3.5.2 Examples of the Matrix Method

HERE WE WILL GIVE SOME EXAMPLES for typical systems for the three cases mentioned in the last section.

**Example 3.10.**  $A = \begin{pmatrix} 4 & 2 \\ 3 & 3 \end{pmatrix}$ .

*Eigenvalues:* We first determine the eigenvalues.

$$0 = \begin{vmatrix} 4 - \lambda & 2 \\ 3 & 3 - \lambda \end{vmatrix} \quad (3.84)$$

Therefore,

$$\begin{aligned} 0 &= (4 - \lambda)(3 - \lambda) - 6 \\ 0 &= \lambda^2 - 7\lambda + 6 \\ 0 &= (\lambda - 1)(\lambda - 6) \end{aligned} \quad (3.85)$$

The eigenvalues are then  $\lambda = 1, 6$ . This is an example of Case I.

**Eigenvectors:** Next we determine the eigenvectors associated with each of these eigenvalues. We have to solve the system  $A\mathbf{v} = \lambda\mathbf{v}$  in each case.

Case  $\lambda = 1$ .

$$\begin{pmatrix} 4 & 2 \\ 3 & 3 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \quad (3.86)$$

$$\begin{pmatrix} 3 & 2 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (3.87)$$

This gives  $3v_1 + 2v_2 = 0$ . One possible solution yields an eigenvector of

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 2 \\ -3 \end{pmatrix}.$$

Case  $\lambda = 6$ .

$$\begin{pmatrix} 4 & 2 \\ 3 & 3 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = 6 \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \quad (3.88)$$

$$\begin{pmatrix} -2 & 2 \\ 3 & -3 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (3.89)$$

For this case we need to solve  $-2v_1 + 2v_2 = 0$ . This yields

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

**General Solution:** We can now construct the general solution.

$$\begin{aligned} \mathbf{x}(t) &= c_1 e^{\lambda_1 t} \mathbf{v}_1 + c_2 e^{\lambda_2 t} \mathbf{v}_2 \\ &= c_1 e^t \begin{pmatrix} 2 \\ -3 \end{pmatrix} + c_2 e^{6t} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 2c_1 e^t + c_2 e^{6t} \\ -3c_1 e^t + c_2 e^{6t} \end{pmatrix}. \end{aligned} \quad (3.90)$$

**Example 3.11.**  $A = \begin{pmatrix} 3 & -5 \\ 1 & -1 \end{pmatrix}$ .

**Eigenvalues:** Again, one solves the eigenvalue equation.

$$0 = \begin{vmatrix} 3 - \lambda & -5 \\ 1 & -1 - \lambda \end{vmatrix} \quad (3.91)$$

Therefore,

$$\begin{aligned}
0 &= (3 - \lambda)(-1 - \lambda) + 5 \\
0 &= \lambda^2 - 2\lambda + 2 \\
\lambda &= \frac{-(-2) \pm \sqrt{4 - 4(1)(2)}}{2} = 1 \pm i. \quad (3.92)
\end{aligned}$$

The eigenvalues are then  $\lambda = 1 + i, 1 - i$ . This is an example of Case III.

**Eigenvectors:** In order to find the general solution, we need only find the eigenvector associated with  $1 + i$ .

$$\begin{aligned}
\begin{pmatrix} 3 & -5 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} &= (1 + i) \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \\
\begin{pmatrix} 2 - i & -5 \\ 1 & -2 - i \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (3.93)
\end{aligned}$$

We need to solve  $(2 - i)v_1 - 5v_2 = 0$ . Thus,

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 2 + i \\ 1 \end{pmatrix}. \quad (3.94)$$

**Complex Solution:** In order to get the two real linearly independent solutions, we need to compute the real and imaginary parts of  $\mathbf{v}e^{\lambda t}$ .

$$\begin{aligned}
e^{\lambda t} \begin{pmatrix} 2 + i \\ 1 \end{pmatrix} &= e^{(1+i)t} \begin{pmatrix} 2 + i \\ 1 \end{pmatrix} \\
&= e^t (\cos t + i \sin t) \begin{pmatrix} 2 + i \\ 1 \end{pmatrix} \\
&= e^t \begin{pmatrix} (2 + i)(\cos t + i \sin t) \\ \cos t + i \sin t \end{pmatrix} \\
&= e^t \begin{pmatrix} (2 \cos t - \sin t) + i(\cos t + 2 \sin t) \\ \cos t + i \sin t \end{pmatrix} \\
&= e^t \begin{pmatrix} 2 \cos t - \sin t \\ \cos t \end{pmatrix} + ie^t \begin{pmatrix} \cos t + 2 \sin t \\ \sin t \end{pmatrix}.
\end{aligned}$$

**General Solution:** Now we can construct the general solution.

$$\begin{aligned}
\mathbf{x}(t) &= c_1 e^t \begin{pmatrix} 2 \cos t - \sin t \\ \cos t \end{pmatrix} + c_2 e^t \begin{pmatrix} \cos t + 2 \sin t \\ \sin t \end{pmatrix} \\
&= e^t \begin{pmatrix} c_1(2 \cos t - \sin t) + c_2(\cos t + 2 \sin t) \\ c_1 \cos t + c_2 \sin t \end{pmatrix}. \quad (3.95)
\end{aligned}$$

Note: This can be rewritten as

$$\mathbf{x}(t) = e^t \cos t \begin{pmatrix} 2c_1 + c_2 \\ c_1 \end{pmatrix} + e^t \sin t \begin{pmatrix} 2c_2 - c_1 \\ c_2 \end{pmatrix}.$$

**Example 3.12.**  $A = \begin{pmatrix} 7 & -1 \\ 9 & 1 \end{pmatrix}$ .

**Eigenvalues:**

$$0 = \begin{vmatrix} 7 - \lambda & -1 \\ 9 & 1 - \lambda \end{vmatrix} \quad (3.96)$$

Therefore,

$$\begin{aligned} 0 &= (7 - \lambda)(1 - \lambda) + 9 \\ 0 &= \lambda^2 - 8\lambda + 16 \\ 0 &= (\lambda - 4)^2. \end{aligned} \quad (3.97)$$

There is only one real eigenvalue,  $\lambda = 4$ . This is an example of Case II.

**Eigenvectors:** In this case we first solve for  $\mathbf{v}_1$  and then get the second linearly independent vector.

$$\begin{aligned} \begin{pmatrix} 7 & -1 \\ 9 & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} &= 4 \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \\ \begin{pmatrix} 3 & -1 \\ 9 & -3 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \end{aligned} \quad (3.98)$$

Therefore, we have

$$3v_1 - v_2 = 0, \quad \Rightarrow \quad \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

**Second Linearly Independent Solution:**

Now we need to solve  $A\mathbf{v}_2 - \lambda\mathbf{v}_2 = \mathbf{v}_1$ .

$$\begin{aligned} \begin{pmatrix} 7 & -1 \\ 9 & 1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} - 4 \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} &= \begin{pmatrix} 1 \\ 3 \end{pmatrix} \\ \begin{pmatrix} 3 & -1 \\ 9 & -3 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} &= \begin{pmatrix} 1 \\ 3 \end{pmatrix}. \end{aligned} \quad (3.99)$$

Expanding the matrix product, we obtain the system of equations

$$\begin{aligned} 3u_1 - u_2 &= 1 \\ 9u_1 - 3u_2 &= 3. \end{aligned} \quad (3.100)$$

The solution of this system is  $\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ .

**General Solution:** We construct the general solution as

$$\mathbf{y}(t) = c_1 e^{\lambda t} \mathbf{v}_1 + c_2 e^{\lambda t} (\mathbf{v}_2 + t\mathbf{v}_1).$$

$$\begin{aligned}
 &= c_1 e^{4t} \begin{pmatrix} 1 \\ 3 \end{pmatrix} + c_2 e^{4t} \left[ \begin{pmatrix} 1 \\ 2 \end{pmatrix} + t \begin{pmatrix} 1 \\ 3 \end{pmatrix} \right] \\
 &= e^{4t} \begin{pmatrix} c_1 + c_2(1+t) \\ 3c_1 + c_2(2+3t) \end{pmatrix}. \tag{3.101}
 \end{aligned}$$

### 3.5.3 Planar Systems - Summary

THE READER SHOULD HAVE NOTED by now that there is a connection between the behavior of the solutions of planar systems obtained in Chapter 2 and the eigenvalues found from the coefficient matrices in the previous examples. Here we summarize some of these cases.

Type	Eigenvalues	Stability
Node	Real $\lambda$ , same signs	$\lambda > 0$ , stable
Saddle	Real $\lambda$ opposite signs	Mostly Unstable
Center	$\lambda$ pure imaginary	—
Focus/Spiral	Complex $\lambda$ , $\text{Re}(\lambda) \neq 0$	$\text{Re}(\lambda > 0)$ , stable
Degenerate Node	Repeated roots	$\lambda > 0$ , stable
Line of Equilibria	One zero eigenvalue	$\lambda > 0$ , stable

Table 3.1: List of typical behaviors in planar systems.

The connection, as we have seen, is that the characteristic equation for the associated second order differential equation is the same as the eigenvalue equation of the coefficient matrix for the linear system. However, one should be a little careful in cases in which the coefficient matrix is not diagonalizable. In Table 3.2 are three examples of systems with repeated roots. The reader should look at these systems and look at the commonalities and differences in these systems and their solutions. In these cases one has unstable nodes, though they are degenerate in that there is only one accessible eigenvector.

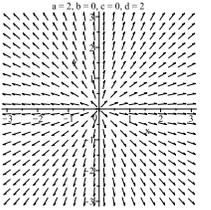
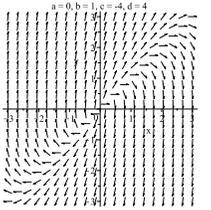
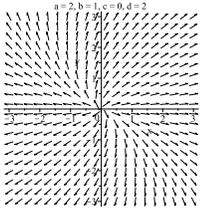
System 1	System 2	System 3
		
$\mathbf{x}' = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \mathbf{x}$	$\mathbf{x}' = \begin{pmatrix} 0 & 1 \\ -4 & 4 \end{pmatrix} \mathbf{x}$	$\mathbf{x}' = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \mathbf{x}$

Table 3.2: Three examples of systems with a repeated root of  $\lambda = 2$ .

### 3.6 Applications

In this section we will describe some simple applications leading to systems of differential equations which can be solved using the methods in this chapter. These systems are left for homework problems and the as the start of further explorations for student projects.

#### 3.6.1 Circuits

In the last chapter we investigated simple series LRC circuits. More complicated circuits are possible by looking at parallel connections, or other combinations, of resistors, capacitors and inductors. This will result in several equations for each loop in the circuit, leading to larger systems of differential equations. an example of another circuit setup is shown in Figure 3.5. This is not a problem that can be covered in the first year physics course.

There are two loops, indicated in Figure 3.6 as traversed clockwise. For each loop we need to apply Kirchoff's Loop Rule. There are three oriented currents, labeled  $I_i$ ,  $i = 1, 2, 3$ . Corresponding to each current is a changing charge,  $q_i$  such that

$$I_i = \frac{dq_i}{dt}, \quad i = 1, 2, 3.$$

For loop one we have

$$I_1 R_1 + \frac{q_2}{C} = V(t). \quad (3.102)$$

For loop two

$$I_3 R_2 + L \frac{dI_3}{dt} = \frac{q_2}{C}. \quad (3.103)$$

We have three unknown functions for the charge. Once we know the charge functions, differentiation will yield the currents. However, we only have two equations. We need a third equation. This is found from Kirchoff's Point (Junction) Rule. Consider the points A and B in Figure 3.6. Any charge (current) entering these junctions must be the same as the total charge (current) leaving the junctions. For point A we have

$$I_1 = I_2 + I_3, \quad (3.104)$$

or

$$\dot{q}_1 = \dot{q}_2 + \dot{q}_3. \quad (3.105)$$

Equations (3.102), (3.103), and (3.105) form a coupled system of differential equations for this problem. There are both first and second order derivatives involved. We can write the whole system in terms of charges as

$$R_1 \dot{q}_1 + \frac{q_2}{C} = V(t)$$

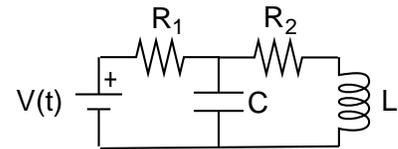


Figure 3.5: A circuit with two loops containing several different circuit elements.

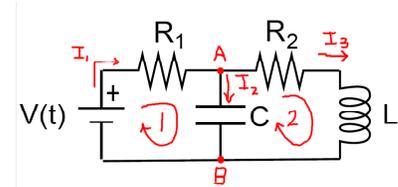


Figure 3.6: The previous parallel circuit with the directions indicated for traversing the loops in Kirchoff's Laws.

$$\begin{aligned} R_2\dot{q}_3 + L\ddot{q}_3 &= \frac{q_2}{C} \\ \dot{q}_1 &= \dot{q}_2 + \dot{q}_3. \end{aligned} \quad (3.106)$$

The question is whether, or not, we can write this as a system of first order differential equations. Since there is only one second order derivative, we can introduce the new variable  $q_4 = \dot{q}_3$ . The first equation can be solved for  $\dot{q}_1$ . The third equation can be solved for  $\dot{q}_2$  with appropriate substitutions for the other terms.  $\dot{q}_3$  is gotten from the definition of  $q_4$  and the second equation can be solved for  $\ddot{q}_3$  and substitutions made to obtain the system

$$\begin{aligned} \dot{q}_1 &= \frac{V}{R_1} - \frac{q_2}{R_1 C} \\ \dot{q}_2 &= \frac{V}{R_1} - \frac{q_2}{R_1 C} - q_4 \\ \dot{q}_3 &= q_4 \\ \dot{q}_4 &= \frac{q_2}{LC} - \frac{R_2}{L}q_4. \end{aligned}$$

So, we have a nonhomogeneous first order system of differential equations. In the last section we learned how to solve such systems.

### 3.6.2 Love Affairs

The next application is one that has been studied by several authors as a cute system involving relationships. One considers what happens to the affections that two people have for each other over time. Let  $R$  denote the affection that Romeo has for Juliet and  $J$  be the affection that Juliet has for Romeo. positive values indicate love and negative values indicate dislike.

One possible model is given by

$$\begin{aligned} \frac{dR}{dt} &= bJ \\ \frac{dJ}{dt} &= cR \end{aligned} \quad (3.107)$$

with  $b > 0$  and  $c < 0$ . In this case Romeo loves Juliet the more she likes him. But Juliet backs away when she finds his love for her increasing.

A typical system relating the combined changes in affection can be modeled as

$$\begin{aligned} \frac{dR}{dt} &= aR + bJ \\ \frac{dJ}{dt} &= cR + dJ. \end{aligned} \quad (3.108)$$

Several scenarios are possible for various choices of the constants. For example, if  $a > 0$  and  $b > 0$ , Romeo gets more and more excited by

Juliet's love for him. If  $c > 0$  and  $d < 0$ , Juliet is being cautious about her relationship with Romeo. For specific values of the parameters and initial conditions, one can explore this match of an overly zealous lover with a cautious lover.

### 3.6.3 Predator Prey Models

Another common model studied is that of competing species. For example, we could consider a population of rabbits and foxes. Left to themselves, rabbits would tend to multiply, thus

$$\frac{dR}{dt} = aR,$$

with  $a > 0$ . In such a model the rabbit population would grow exponentially. Similarly, a population of foxes would decay without the rabbits to feed on. So, we have that

$$\frac{dF}{dt} = -bF$$

for  $b > 0$ .

Now, if we put these populations together on a deserted island, they would interact. The more foxes, the rabbit population would decrease. However, the more rabbits, the foxes would have plenty to eat and the population would thrive. Thus, we could model the competing populations as

$$\begin{aligned} \frac{dR}{dt} &= aR - cF, \\ \frac{dF}{dt} &= -bF + dR, \end{aligned} \tag{3.109}$$

where all of the constants are positive numbers. Studying this coupled system would lead to a study of the dynamics of these populations. We will discuss other (nonlinear) systems in the next chapter.

### 3.6.4 Mixture Problems

There are many types of mixture problems. Such problems are standard in a first course on differential equations as examples of first order differential equations. Typically these examples consist of a tank of brine, water containing a specific amount of salt with pure water entering and the mixture leaving, or the flow of a pollutant into, or out of, a lake.

In general one has a rate of flow of some concentration of mixture entering a region and a mixture leaving the region. The goal is to determine how much stuff is in the region at a given time. This is governed by the equation

$$\text{Rate of change of substance} = \text{Rate In} - \text{Rate Out.}$$

This can be generalized to the case of two interconnected tanks. We provide some examples.

### Example 3.13. Single Tank Problem

A 50 gallon tank of pure water has a brine mixture with concentration of 2 pounds per gallon entering at the rate of 5 gallons per minute. [See Figure 3.7.] At the same time the well-mixed contents drain out at the rate of 5 gallons per minute. Find the amount of salt in the tank at time  $t$ . In all such problems one assumes that the solution is well mixed at each instant of time.

Let  $x(t)$  be the amount of salt at time  $t$ . Then the rate at which the salt in the tank increases is due to the amount of salt entering the tank less that leaving the tank. To figure out these rates, one notes that  $dx/dt$  has units of pounds per minute. The amount of salt entering per minute is given by the product of the entering concentration times the rate at which the brine enters. This gives the correct units:

$$\left(2 \frac{\text{pounds}}{\text{gal}}\right) \left(5 \frac{\text{gal}}{\text{min}}\right) = 10 \frac{\text{pounds}}{\text{min}}.$$

Similarly, one can determine the rate out as

$$\left(\frac{x \text{ pounds}}{50 \text{ gal}}\right) \left(5 \frac{\text{gal}}{\text{min}}\right) = \frac{x}{10} \frac{\text{pounds}}{\text{min}}.$$

Thus, we have

$$\frac{dx}{dt} = 10 - \frac{x}{10}.$$

This equation is easily solved using the methods for first order equations.

### Example 3.14. Double Tank Problem

One has two tanks connected together, labeled tank X and tank Y, as shown in Figure 3.8.

Let tank X initially have 100 gallons of brine made with 100 pounds of salt. Tank Y initially has 100 gallons of pure water. Now pure water is pumped into tank X at a rate of 2.0 gallons per minute. Some of the mixture of brine and pure water flows into tank Y at 3 gallons per minute. To keep the tank levels the same, one gallon of the Y mixture flows back into tank X at a rate of one gallon per minute and 2.0 gallons per minute drains out. Find the amount of salt at any given time in the tanks. What happens over a long period of time?

In this problem we set up two equations. Let  $x(t)$  be the amount of salt in tank X and  $y(t)$  the amount of salt in tank Y. Again, we carefully look at the rates into and out of each tank in order to set up the system of differential equations. We obtain the system

$$\begin{aligned} \frac{dx}{dt} &= \frac{y}{100} - \frac{3x}{100} \\ \frac{dy}{dt} &= \frac{3x}{100} - \frac{3y}{100}. \end{aligned} \quad (3.110)$$

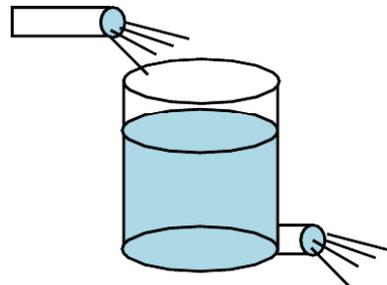


Figure 3.7: A typical mixing problem.

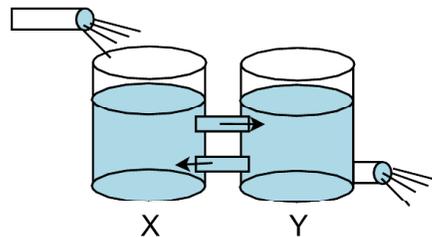
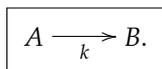


Figure 3.8: The two tank problem.

This is a linear, homogenous constant coefficient system of two first order equations, which we know how to solve.

### 3.6.5 Chemical Kinetics

There are many problems that come from studying chemical reactions. The simplest reaction is when a chemical  $A$  turns into chemical  $B$ . This happens at a certain rate,  $k > 0$ . This can be represented by the chemical formula

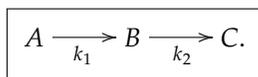


In this case we have that the rates of change of the concentrations of  $A$ ,  $[A]$ , and  $B$ ,  $[B]$ , are given by

$$\begin{aligned} \frac{d[A]}{dt} &= -k[A] \\ \frac{d[B]}{dt} &= k[A] \end{aligned} \quad (3.111)$$

Think about this as it is a key to understanding the next reactions.

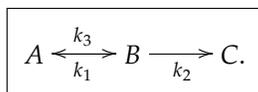
A more complicated reaction is given by



In this case we can add to the above equation the rates of change of concentrations  $[B]$  and  $[C]$ . The resulting system of equations is

$$\begin{aligned} \frac{d[A]}{dt} &= -k_1[A], \\ \frac{d[B]}{dt} &= k_1[A] - k_2[B], \\ \frac{d[C]}{dt} &= k_2[B]. \end{aligned} \quad (3.112)$$

One can further consider reactions in which a reverse reaction is possible. Thus, a further generalization occurs for the reaction



The resulting system of equations is

$$\begin{aligned} \frac{d[A]}{dt} &= -k_1[A] + k_3[B], \\ \frac{d[B]}{dt} &= k_1[A] - k_2[B] - k_3[B], \\ \frac{d[C]}{dt} &= k_2[B]. \end{aligned} \quad (3.113)$$

More complicated chemical reactions will be discussed at a later time.

### 3.6.6 Epidemics

Another interesting area of application of differential equation is in predicting the spread of disease. Typically, one has a population of susceptible people or animals. Several infected individuals are introduced into the population and one is interested in how the infection spreads and if the number of infected people drastically increases or dies off. Such models are typically nonlinear and we will look at what is called the SIR model in the next chapter. In this section we will model a simple linear model.

Let break the population into three classes. First,  $S(t)$  are the healthy people, who are susceptible to infection. Let  $I(t)$  be the number of infected people. Of these infected people, some will die from the infection and others recover. Let's assume that initially there is one infected person and the rest, say  $N$ , are obviously healthy. Can we predict how many deaths have occurred by time  $t$ ?

Let's try and model this problem using the compartmental analysis we had seen in the mixing problems. The total rate of change of any population would be due to those entering the group less those leaving the group. For example, the number of healthy people decreases due to infection and can increase when some of the infected group recovers. Let's assume that the rate of infection is proportional to the number of healthy people,  $aS$ . Also, we assume that the number who recover is proportional to the number of infected,  $rI$ . Thus, the rate of change of the healthy people is found as

$$\frac{dS}{dt} = -aS + rI.$$

Let the number of deaths be  $D(t)$ . Then, the death rate could be taken to be proportional to the number of infected people. So,

$$\frac{dD}{dt} = dI$$

Finally, the rate of change of infectives is due to healthy people getting infected and the infectives who either recover or die. Using the corresponding terms in the other equations, we can write

$$\frac{dI}{dt} = aS - rI - dI.$$

This linear system can be written in matrix form.

$$\frac{d}{dt} \begin{pmatrix} S \\ I \\ D \end{pmatrix} = \begin{pmatrix} -a & r & 0 \\ a & -d-r & 0 \\ 0 & d & 0 \end{pmatrix} \begin{pmatrix} S \\ I \\ D \end{pmatrix}. \quad (3.114)$$

The eigenvalue equation for this system is

$$\lambda \left[ \lambda^2 + (a + r + d)\lambda + ad \right] = 0.$$

The reader can find the solutions of this system and determine if this is a realistic model.

### 3.7 Rotations of Conics

EIGENVALUE PROBLEMS show up in applications other than the solution of differential equations. We will see applications of this later in the text. For now, we are content to deal with problems which can be cast into matrix form. One example is the transformation of a simple system through rotation into a more complicated appearing system simply do to the choice of coordinate system. In this section we will explore this through the study of the rotation of conics.

You may have seen the general form for the equation of a conic in Cartesian coordinates in your calculus class. It is given by

$$Ax^2 + 2Bxy + Cy^2 + Ex + Fy = D. \quad (3.115)$$

This equation can describe a variety of conics (ellipses, hyperbolae and parabolae) depending on the constants. The  $E$  and  $F$  terms result from a translation<sup>9</sup> of the origin and the  $B$  term is the result of a rotation of the coordinate system. We leave it to the reader to show that coordinate translations can be made to eliminate the linear terms. So, we will set  $E = F = 0$  in our discussion and only consider quadratic equations of the form

$$Ax^2 + 2Bxy + Cy^2 = D.$$

If  $B = 0$ , then the resulting equation could be an equation for the standard ellipse or hyperbola with center at the origin. In the case of an ellipse, the semimajor and semiminor axes lie along the coordinate axes. However, you could rotate the ellipse and that would introduce a  $B$  term, as we will see.

This conic equation can be written in matrix form. We note that

$$\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} A & B \\ B & C \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = Ax^2 + 2Bxy + Cy^2.$$

In short hand matrix form, we thus have for our equation

$$\mathbf{x}^T Q \mathbf{x} = D,$$

where  $Q$  is the matrix of coefficients  $A$ ,  $B$ , and  $C$ .

We want to determine the transformation that puts this conic into a coordinate system in which there is no  $B$  term. Our goal is to obtain an equation of the form

$$A'x'^2 + C'y'^2 = D'$$

<sup>9</sup> It is easy to see how such terms correspond to translations of conics. Consider the simple example  $x^2 + y^2 + 2x - 6y = 0$ . By completing the squares in both  $x$  and  $y$ , this equation can be written as  $(x + 1)^2 + (y - 3)^2 = 10$ . Now you recognize that this is a circle whose center has been translated from the origin to  $(-1, 3)$ .

in the new coordinates  $\mathbf{y}^T = (x', y')$ . The matrix form of this equation is given as

$$\mathbf{y}^T \begin{pmatrix} A' & 0 \\ 0 & C' \end{pmatrix} \mathbf{y} = D'.$$

We will denote the diagonal matrix by  $\Lambda$ .

So, we let

$$\mathbf{x} = R\mathbf{y},$$

where  $R$  is a rotation matrix. Inserting this transformation into our equation we find that

$$\begin{aligned} \mathbf{x}^T Q \mathbf{x} &= (R\mathbf{y})^T Q R \mathbf{y} \\ &= \mathbf{y}^T (R^T Q R) \mathbf{y}. \end{aligned} \quad (3.116)$$

Comparing this result to the desired form, we have

$$\Lambda = R^T Q R. \quad (3.117)$$

Recalling that the rotation matrix is an orthogonal matrix,  $R^T = R^{-1}$ , we have

$$\Lambda = R^{-1} Q R. \quad (3.118)$$

Thus, the problem reduces to that of trying to diagonalize the matrix  $Q$ . The eigenvalues of  $Q$  will lead to the constants in the rotated equation and the eigenvectors, as we will see, will give the directions of the principal axes (the semimajor and semiminor axes). We will first show this in an example.

**Example 3.15.** Determine the principle axes of the ellipse given by

$$13x^2 - 10xy + 13y^2 - 72 = 0.$$

A plot of this conic in Figure 3.9 shows that it is an ellipse. However, we might not know this without plotting it. (Actually, there are some conditions on the coefficients that do allow us to determine the conic. But you may not know this yet.) If the equation were in standard form, we could identify its general shape. So, we will use the method outlined above to find a coordinate system in which the ellipse appears in standard form.

The coefficient matrix for this equation is given by

$$Q = \begin{pmatrix} 13 & -5 \\ -5 & 13 \end{pmatrix}. \quad (3.119)$$

We seek a solution to the eigenvalue problem:  $Q\mathbf{v} = \lambda\mathbf{v}$ . Recall, the first step is to get the eigenvalue equation from  $\det(Q - \lambda I) = 0$ . For this problem we have

$$\begin{vmatrix} 13 - \lambda & -5 \\ -5 & 13 - \lambda \end{vmatrix} = 0. \quad (3.120)$$

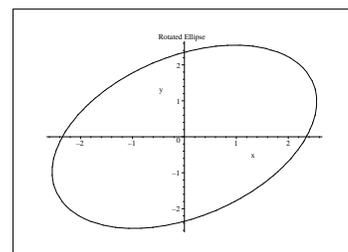


Figure 3.9: Plot of the ellipse given by  $13x^2 - 10xy + 13y^2 - 72 = 0$ .

So, we have to solve

$$(13 - \lambda)^2 - 25 = 0.$$

This is easily solved by taking square roots to get

$$\lambda - 13 = \pm 5,$$

or

$$\lambda = 13 \pm 5 = 18, 8.$$

Thus, the equation in the new system is

$$8x'^2 + 18y'^2 = 72.$$

Dividing out the 72 puts this into the standard form

$$\frac{x'^2}{9} + \frac{y'^2}{4} = 1.$$

Now we can identify the ellipse in the new system. We show the two ellipses in Figure 3.10. We note that the given ellipse is the new one rotated by some angle, which we still need to determine.

Next, we seek the eigenvectors corresponding to each eigenvalue.

**Eigenvalue 1:**  $\lambda = 8$

We insert the eigenvalue into the equation  $(Q - \lambda I)\mathbf{v} = 0$ . The system for the unknown eigenvector is

$$\begin{pmatrix} 13 - 8 & -5 \\ -5 & 13 - 8 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = 0. \quad (3.121)$$

The first equation is

$$5v_1 - 5v_2 = 0, \quad (3.122)$$

or  $v_1 = v_2$ . Thus, we can choose our eigenvector to be

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

**Eigenvalue 2:**  $\lambda = 18$

In the same way, we insert the eigenvalue into the equation  $(Q - \lambda I)\mathbf{v} = 0$  and obtain

$$\begin{pmatrix} 13 - 18 & -5 \\ -5 & 13 - 18 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = 0. \quad (3.123)$$

The first equation is

$$-5v_1 - 5v_2 = 0, \quad (3.124)$$

or  $v_1 = -v_2$ . Thus, we can choose our eigenvector to be

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

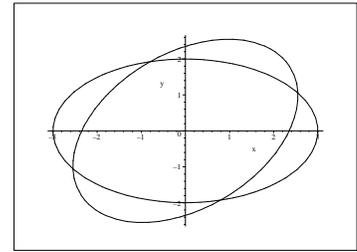


Figure 3.10: Plot of the ellipse given by  $13x^2 - 10xy + 13y^2 - 72 = 0$  and the ellipse  $\frac{x'^2}{9} + \frac{y'^2}{4} = 1$  showing that the first ellipse is a rotated version of the second ellipse.

In Figure 3.11 we superimpose the eigenvectors on our original ellipse. We see that the eigenvectors point in directions along the semimajor and semiminor axes and indicate the angle of rotation. Eigenvector one is at a  $45^\circ$  angle. Thus, our ellipse is a rotated version of one in standard position. Or, we could define new axes that are at  $45^\circ$  to the standard axes and then the ellipse would take the standard form in the new coordinate system.

A general rotation of any conic can be performed. Consider the general equation:

$$Ax^2 + 2Bxy + Cy^2 + Ex + Fy = D. \quad (3.125)$$

We would like to find a rotation that puts it in the form

$$\lambda_1 x'^2 + \lambda_2 y'^2 + E'x' + F'y' = D. \quad (3.126)$$

We use the rotation matrix

$$\hat{R}_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

and define  $\mathbf{x}' = \hat{R}_\theta^T \mathbf{x}$ , or  $\mathbf{x} = R_\theta \mathbf{x}'$ .

The general equation can be written in matrix form:

$$\mathbf{x}^T Q \mathbf{x} + \mathbf{f} \mathbf{x} = D, \quad (3.127)$$

where  $Q$  is the usual matrix of coefficients  $A$ ,  $B$ , and  $C$  and  $\mathbf{f} = (E, F)$ . Transforming this equation gives

$$\mathbf{x}'^T R_\theta^{-1} Q R_\theta \mathbf{x}' + \mathbf{f} R_\theta \mathbf{x}' = D. \quad (3.128)$$

The resulting equation is of the form

$$A'x'^2 + 2B'x'y' + C'y'^2 + E'x' + F'y' = D, \quad (3.129)$$

where

$$B' = 2(C - A) \sin \theta \cos \theta + 2B(2 \cos^2 \theta - 1). \quad (3.130)$$

(We only need  $B'$  for this discussion). If we want the nonrotated form, then we seek an angle  $\theta$  such that  $B' = 0$ . Noting that  $2 \sin \theta \cos \theta = \sin 2\theta$  and  $2 \cos^2 \theta - 1 = \cos 2\theta$ , this gives

$$\tan(2\theta) = \frac{A - C}{B}. \quad (3.131)$$

**Example 3.16.** So, in our previous example, with  $A = C = 13$  and  $B = -5$ , we have  $\tan(2\theta) = \infty$ . Thus,  $2\theta = \pi/2$ , or  $\theta = \pi/4$ .

Finally, we had noted that knowing the coefficients in the general quadratic is enough to determine the type of conic represented without doing any plotting. This is based on the fact that the determinant of

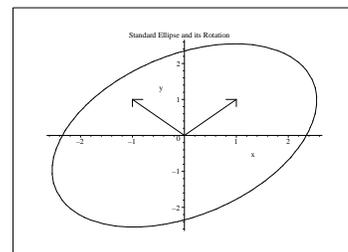


Figure 3.11: Plot of the ellipse given by  $13x^2 - 10xy + 13y^2 - 72 = 0$  and the eigenvectors. Note that they are along the semimajor and semiminor axes and indicate the angle of rotation.

the coefficient matrix is invariant under rotation. We see this from the equation for diagonalization

$$\begin{aligned} \det(\Lambda) &= \det(R_\theta^{-1}QR_\theta) \\ &= \det(R_\theta^{-1})\det(Q)\det(R_\theta) \\ &= \det(R_\theta^{-1}R_\theta)\det(Q) \\ &= \det(Q). \end{aligned} \tag{3.132}$$

Therefore, we have

$$\lambda_1\lambda_2 = AC - B^2.$$

Looking at Equation (3.126), we have three cases:

1. Ellipse  $\lambda_1\lambda_2 > 0$  or  $B^2 - AC < 0$ .
2. Hyperbola  $\lambda_1\lambda_2 < 0$  or  $B^2 - AC > 0$ .
3. Parabola  $\lambda_1\lambda_2 = 0$  or  $B^2 - AC = 0$ . and one eigenvalue is nonzero. Otherwise the equation degenerates to a linear equation.

**Example 3.17.** Consider the hyperbola  $xy = 6$ . We can see that this is a rotated hyperbola by plotting  $y = 6/x$ . A plot is shown in Figure 3.12. Determine the rotation need to put transform the hyperbola to new coordinates so that its equation will be in standard form.

The coefficient matrix for this equation is given by

$$A = \begin{pmatrix} 0 & -0.5 \\ 0.5 & 0 \end{pmatrix}. \tag{3.133}$$

The eigenvalue equation is

$$\begin{vmatrix} -\lambda & -0.5 \\ -0.5 & -\lambda \end{vmatrix} = 0. \tag{3.134}$$

Thus,

$$\lambda^2 - 0.25 = 0,$$

or  $\lambda = \pm 0.5$ .

Once again,  $\tan(2\theta) = \infty$ , so the new system is at  $45^\circ$  to the old. The equation in new coordinates is  $0.5x^2 + (-0.5)y^2 = 6$ , or  $x^2 - y^2 = 12$ . A plot is shown in Figure 3.13.

### 3.8 Appendix: Diagonalization and Linear Systems

AS WE HAVE SEEN, the matrix formulation for linear systems can be powerful, especially for  $n$  differential equations involving  $n$  unknown functions. Our ability to proceed towards solutions depended upon

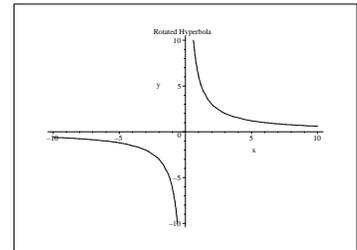


Figure 3.12: Plot of the hyperbola given by  $xy = 6$ .

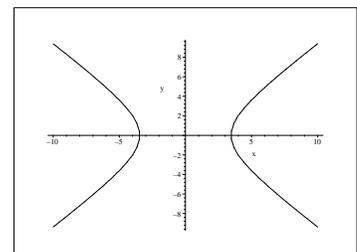


Figure 3.13: Plot of the rotated hyperbola given by  $x^2 - y^2 = 12$ .

the solution of eigenvalue problems. However, in the case of repeated eigenvalues we saw some additional complications. This all depends deeply on the background linear algebra. Namely, we relied on being able to diagonalize the given coefficient matrix. In this section we will discuss the limitations of diagonalization and introduce the Jordan canonical form.

We begin with the notion of similarity. Matrix  $A$  is *similar* to matrix  $B$  if and only if there exists a nonsingular matrix  $P$  such that

$$B = P^{-1}AP. \quad (3.135)$$

Recall that a nonsingular matrix has a nonzero determinant and is invertible.

We note that the similarity relation is an equivalence relation. Namely, it satisfies the following

1.  $A$  is similar to itself.
2. If  $A$  is similar to  $B$ , then  $B$  is similar to  $A$ .
3. If  $A$  is similar to  $B$  and  $B$  is similar to  $C$ , then  $A$  is similar to  $C$ .

Also, if  $A$  is similar to  $B$ , then they have the same eigenvalues. This follows from a simple computation of the eigenvalue equation. Namely,

$$\begin{aligned} 0 &= \det(B - \lambda I) \\ &= \det(P^{-1}AP - \lambda P^{-1}IP) \\ &= \det(P)^{-1} \det(A - \lambda I) \det(P) \\ &= \det(A - \lambda I). \end{aligned} \quad (3.136)$$

Therefore,  $\det(A - \lambda I) = 0$  and  $\lambda$  is an eigenvalue of both  $A$  and  $B$ .

An  $n \times n$  matrix  $A$  is *diagonalizable* if and only if  $A$  is similar to a diagonal matrix  $D$ ; i.e., there exists a nonsingular matrix  $P$  such that

$$D = P^{-1}AP. \quad (3.137)$$

One of the most important theorems in linear algebra is the Spectral Theorem. This theorem tells us when a matrix can be diagonalized. In fact, it goes beyond matrices to the diagonalization of linear operators. We learn in linear algebra that linear operators can be represented by matrices once we pick a particular representation basis. Diagonalization is simplest for finite dimensional vector spaces and requires some generalization for infinite dimensional vector spaces. Examples of operators to which the spectral theorem applies are self-adjoint operators (more generally normal operators on Hilbert spaces). We will explore

some of these ideas later in the course. The spectral theorem provides a canonical decomposition, called the spectral decomposition, or eigendecomposition, of the underlying vector space on which it acts.

The next theorem tells us how to diagonalize a matrix:

**Theorem 3.1.** *Let  $A$  be an  $n \times n$  matrix. Then  $A$  is diagonalizable if and only if  $A$  has  $n$  linearly independent eigenvectors. If so, then*

$$D = P^{-1}AP.$$

*If  $\{v_1, \dots, v_n\}$  are the eigenvectors of  $A$  and  $\{\lambda_1, \dots, \lambda_n\}$  are the corresponding eigenvalues, then  $v_j$  is the  $j$ th column of  $P$  and  $D_{jj} = \lambda_j$ .*

A simpler determination results by noting

**Theorem 3.2.** *Let  $A$  be an  $n \times n$  matrix with  $n$  real and distinct eigenvalues. Then  $A$  is diagonalizable.*

Therefore, we need only look at the eigenvalues and determine diagonalizability. In fact, one also has from linear algebra the following result.

**Theorem 3.3.** *Let  $A$  be an  $n \times n$  real symmetric matrix. Then  $A$  is diagonalizable.*

Recall that a symmetric matrix is one whose transpose is the same as the matrix, or  $A_{ij} = A_{ji}$ .

**Example 3.18.** *Consider the matrix*

$$A = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 3 & 0 \\ 2 & 0 & 3 \end{pmatrix}$$

*This is a real symmetric matrix. The characteristic polynomial is found to be*

$$\det(A - \lambda I) = -(\lambda - 5)(\lambda - 3)(\lambda + 1) = 0.$$

*As before, we can determine the corresponding eigenvectors (for  $\lambda = -1, 3, 5$ , respectively) as*

$$\begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

*We can use these to construct the diagonalizing matrix  $P$ . Namely, we have*

$$\begin{aligned} P^{-1}AP &= \begin{pmatrix} -2 & 0 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 2 & 2 \\ 2 & 3 & 0 \\ 2 & 0 & 3 \end{pmatrix} \begin{pmatrix} -2 & 0 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} -1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{pmatrix}. \end{aligned} \tag{3.138}$$

Now diagonalization is an important idea in solving linear systems of first order equations, as we have seen for simple systems. If our system is originally diagonal, that means our equations are completely uncoupled. Let our system take the form

$$\frac{d\mathbf{y}}{dt} = D\mathbf{y}, \quad (3.139)$$

where  $D$  is diagonal with entries  $\lambda_i$ ,  $i = 1, \dots, n$ . The system of equations,  $y'_i = \lambda_i y_i$ , has solutions

$$y_i(t) = c_i e^{\lambda_i t}.$$

Thus, it is easy to solve a diagonal system.

Let  $A$  be similar to this diagonal matrix. Then

$$\frac{d\mathbf{y}}{dt} = P^{-1}AP\mathbf{y}. \quad (3.140)$$

This can be rewritten as

$$\frac{dP\mathbf{y}}{dt} = AP\mathbf{y}.$$

Defining  $\mathbf{x} = P\mathbf{y}$ , we have

$$\frac{d\mathbf{x}}{dt} = A\mathbf{x}. \quad (3.141)$$

This simple derivation shows that if  $A$  is diagonalizable, then a transformation of the original system in  $\mathbf{x}$  to new *coordinates*, or a new basis, results in a simpler system in  $\mathbf{y}$ .

However, it is not always possible to diagonalize a given square matrix. This is because some matrices do not have enough linearly independent vectors, or we have repeated eigenvalues. However, we have the following theorem:

**Theorem 3.4.** *Every  $n \times n$  matrix  $A$  is similar to a matrix of the form*

$$J = \text{diag}[J_1, J_2, \dots, J_n],$$

where

$$J_i = \begin{pmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_i & 1 \\ 0 & 0 & \cdots & 0 & \lambda_i \end{pmatrix} \quad (3.142)$$

We will not go into the details of how one finds this **Jordan Canonical Form** or proving the theorem. In practice you can use a computer algebra system to determine this and the similarity matrix. However, we would still need to know how to use it to solve our system of differential equations.

**Example 3.19.** Let's consider a simple system with the  $3 \times 3$  Jordan block

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

The corresponding system of coupled first order differential equations takes the form

$$\begin{aligned} \frac{dx_1}{dt} &= 2x_1 + x_2, \\ \frac{dx_2}{dt} &= 2x_2 + x_3, \\ \frac{dx_3}{dt} &= 2x_3. \end{aligned} \quad (3.143)$$

The last equation is simple to solve, giving  $x_3(t) = c_3 e^{2t}$ . Inserting into the second equation, you have a

$$\frac{dx_2}{dt} = 2x_2 + c_3 e^{2t}.$$

Using the integrating factor,  $e^{-2t}$ , one can solve this equation to get  $x_2(t) = (c_2 + c_3 t)e^{2t}$ . Similarly, one can solve the first equation to obtain  $x_1(t) = (c_1 + c_2 t + \frac{1}{2}c_3 t^2)e^{2t}$ .

This should remind you of a problem we had solved earlier leading to the generalized eigenvalue problem in (3.83). This suggests that there is a more general theory when there are multiple eigenvalues and relating to Jordan canonical forms.

Let's write the solution we just obtained in vector form. We have

$$\mathbf{x}(t) = \left[ c_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} t \\ 1 \\ 0 \end{pmatrix} + c_3 \begin{pmatrix} \frac{1}{2}t^2 \\ t \\ 1 \end{pmatrix} \right] e^{2t}. \quad (3.144)$$

It looks like this solution is a linear combination of three linearly independent solutions,

$$\begin{aligned} \mathbf{x} &= \mathbf{v}_1 e^{2\lambda t} \\ \mathbf{x} &= (t\mathbf{v}_1 + \mathbf{v}_2) e^{\lambda t} \\ \mathbf{x} &= \left(\frac{1}{2}t^2\mathbf{v}_1 + t\mathbf{v}_2 + \mathbf{v}_3\right) e^{\lambda t}, \end{aligned} \quad (3.145)$$

where  $\lambda = 2$  and the vectors satisfy the equations

$$\begin{aligned} (A - \lambda I)\mathbf{v}_1 &= \mathbf{0}, \\ (A - \lambda I)\mathbf{v}_2 &= \mathbf{v}_1, \\ (A - \lambda I)\mathbf{v}_3 &= \mathbf{v}_2, \end{aligned} \quad (3.146)$$

and

$$\begin{aligned}(A - \lambda I)\mathbf{v}_1 &= 0, \\ (A - \lambda I)^2\mathbf{v}_2 &= 0, \\ (A - \lambda I)^3\mathbf{v}_3 &= 0.\end{aligned}\tag{3.147}$$

It is easy to generalize this result to build linearly independent solutions corresponding to multiple roots (eigenvalues) of the characteristic equation.

### Problems

1. Express the vector  $\mathbf{v} = (1, 2, 3)$  as a linear combination of the vectors  $\mathbf{a}_1 = (1, 1, 1)$ ,  $\mathbf{a}_2 = (1, 0, -1)$ , and  $\mathbf{a}_3 = (2, 1, 0)$ .

2. A symmetric matrix is one for which the transpose of the matrix is the same as the original matrix,  $A^T = A$ . An antisymmetric matrix is one which satisfies  $A^T = -A$ .

- Show that the diagonal elements of an  $n \times n$  antisymmetric matrix are all zero.
- Show that a general  $3 \times 3$  antisymmetric matrix has three independent off-diagonal elements.
- How many independent elements does a general  $3 \times 3$  symmetric matrix have?
- How many independent elements does a general  $n \times n$  symmetric matrix have?
- How many independent elements does a general  $n \times n$  antisymmetric matrix have?

3. Consider the matrix representations for two dimensional rotations of vectors by angles  $\alpha$  and  $\beta$ , denoted by  $R_\alpha$  and  $R_\beta$ , respectively.

- Find  $R_\alpha^{-1}$  and  $R_\alpha^T$ . How do they relate?
- Prove that  $R_{\alpha+\beta} = R_\alpha R_\beta = R_\beta R_\alpha$ .

4. The Pauli spin matrices in quantum mechanics are given by the matrices:  $\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ ,  $\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ , and  $\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ .

Show that

- $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = I$ .
- $\{\sigma_i, \sigma_j\} \equiv \sigma_i \sigma_j + \sigma_j \sigma_i = 2\delta_{ij}I$ , for  $i, j = 1, 2, 3$  and  $I$  the  $2 \times 2$  identity matrix.  $\{\cdot, \cdot\}$  is the anti-commutation operation.
- $[\sigma_1, \sigma_2] \equiv \sigma_1 \sigma_2 - \sigma_2 \sigma_1 = 2i\sigma_3$ , and similarly for the other pairs.  $[\cdot, \cdot]$  is the commutation operation.

- d. Show that an arbitrary  $2 \times 2$  matrix  $M$  can be written as a linear combination of Pauli matrices,  $M = a_0I + \sum_{j=1}^3 a_j\sigma_j$ , where the  $a_j$ 's are complex numbers.

5. Use Cramer's Rule to solve the system:

$$\begin{aligned} 2x - 5z &= 7 \\ x - 2y &= 1 \\ 3x - 5y - z &= 4. \end{aligned} \tag{3.148}$$

6. Find the eigenvalue(s) and eigenvector(s) for the following:

a.  $\begin{pmatrix} 4 & 2 \\ 3 & 3 \end{pmatrix}$

b.  $\begin{pmatrix} 3 & -5 \\ 1 & -1 \end{pmatrix}$

c.  $\begin{pmatrix} 4 & 1 \\ 0 & 4 \end{pmatrix}$

d.  $\begin{pmatrix} 1 & -1 & 4 \\ 3 & 2 & -1 \\ 2 & 1 & -1 \end{pmatrix}$

7. For the matrices in the last problem, compute the determinants and find the inverses, if they exist.

8. Consider the conic  $5x^2 - 4xy + 2y^2 = 30$ .

- Write the left side in matrix form.
- Diagonalize the coefficient matrix, finding the eigenvalues and eigenvectors.
- Construct the rotation matrix from the information in part b. What is the angle of rotation needed to bring the conic into standard form?
- What is the conic?

9. In Equation (3.76) the exponential of a matrix was defined.

- a. Let

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}.$$

Compute  $e^A$ .

- b. Give a definition of  $\cos A$  and compute  $\cos \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$  in simplest form.

- c. Using the definition of  $e^A$ , prove  $e^{PAP^{-1}} = Pe^AP^{-1}$  for general  $A$ .

**10.** Consider the following systems. For each system determine the coefficient matrix. When possible, solve the eigenvalue problem for each matrix and use the eigenvalues and eigenfunctions to provide solutions to the given systems. Finally, in the common cases which you investigated in Problem 17, make comparisons with your previous answers, such as what type of eigenvalues correspond to stable nodes.

a.

$$\begin{aligned}x' &= 3x - y \\y' &= 2x - 2y.\end{aligned}$$

b.

$$\begin{aligned}x' &= -y \\y' &= -5x.\end{aligned}$$

c.

$$\begin{aligned}x' &= x - y \\y' &= y.\end{aligned}$$

d.

$$\begin{aligned}x' &= 2x + 3y \\y' &= -3x + 2y.\end{aligned}$$

e.

$$\begin{aligned}x' &= -4x - y \\y' &= x - 2y.\end{aligned}$$

f.

$$\begin{aligned}x' &= x - y \\y' &= x + y.\end{aligned}$$

**11.** Add a third spring connected to mass two in the coupled system shown in Figure 2.17 to a wall on the far right. Assume that the masses are the same and the springs are the same.

- Model this system with a set of first order differential equations.
- If the masses are all 2.0 kg and the spring constants are all 10.0 N/m, then find the general solution for the system.

- c. Move mass one to the left (of equilibrium) 10.0 cm and mass two to the right 5.0 cm. Let them go. find the solution and plot it as a function of time. Where is each mass at 5.0 seconds?
12. Consider the series circuit in Figure 2.7 with  $L = 1.00$  H,  $R = 1.00 \times 10^2 \Omega$ ,  $C = 1.00 \times 10^{-4}$  F, and  $V_0 = 1.00 \times 10^3$  V.
- Set up the problem as a system of two first order differential equations for the charge and the current.
  - Suppose that no charge is present and no current is flowing at time  $t = 0$  when  $V_0$  is applied. Find the current and the charge on the capacitor as functions of time.
  - Plot your solutions and describe how the system behaves over time.
13. Consider the series circuit in Figure 3.5 with  $L = 1.00$  H,  $R_1 = R_2 = 1.00 \times 10^2 \Omega$ ,  $C = 1.00 \times 10^{-4}$  F, and  $V_0 = 1.00 \times 10^3$  V.
- Set up the problem as a system of first order differential equations for the charges and the currents in each loop.
  - Suppose that no charge is present and no current is flowing at time  $t = 0$  when  $V_0$  is applied. Find the current and the charge on the capacitor as functions of time.
  - Plot your solutions and describe how the system behaves over time.
14. Initially a 200 gallon tank is filled with pure water. At time  $t = 0$  a salt concentration with 3 pounds of salt per gallon is added to the container at the rate of 4 gallons per minute, and the well-stirred mixture is drained from the container at the same rate.
- Find the number of pounds of salt in the container as a function of time.
  - How many minutes does it take for the concentration to reach 2 pounds per gallon?
  - What does the concentration in the container approach for large values of time? Does this agree with your intuition?
  - Assuming that the tank holds much more than 200 gallons, and everything is the same except that the mixture is drained at 3 gallons per minute, what would the answers to parts a and b become?
15. You make two gallons of chili for a party. The recipe calls for two teaspoons of hot sauce per gallon, but you had accidentally put

in two tablespoons per gallon. You decide to feed your guests the chili anyway. Assume that the guests take 1 cup/min of chili and you replace what was taken with beans and tomatoes without any hot sauce. [1 gal = 16 cups and 1 Tb = 3 tsp.]

- a. Write down the differential equation and initial condition for the amount of hot sauce as a function of time in this mixture-type problem.
- b. Solve this initial value problem.
- c. How long will it take to get the chili back to the recipe's suggested concentration?

**16.** Consider the chemical reaction leading to the system in (3.113). Let the rate constants be  $k_1 = 0.20 \text{ ms}^{-1}$ ,  $k_2 = 0.05 \text{ ms}^{-1}$ , and  $k_3 = 0.10 \text{ ms}^{-1}$ . What do the eigenvalues of the coefficient matrix say about the behavior of the system? Find the solution of the system assuming  $[A](0) = A_0 = 1.0 \text{ } \mu\text{mol}$ ,  $[B](0) = 0$ , and  $[C](0) = 0$ . Plot the solutions for  $t = 0.0$  to  $50.0 \text{ ms}$  and describe what is happening over this time.

**17.** Consider the epidemic model leading to the system in (3.114). Choose the constants as  $a = 2.0 \text{ days}^{-1}$ ,  $d = 3.0 \text{ days}^{-1}$ , and  $r = 1.0 \text{ days}^{-1}$ . What are the eigenvalues of the coefficient matrix? Find the solution of the system assuming an initial population of 1,000 and one infected individual. Plot the solutions for  $t = 0.0$  to  $5.0 \text{ days}$  and describe what is happening over this time. Is this model realistic?