

## **Chapter 1**

# SYSTEMS OF EQUATIONS AND MATRICES

## 1.1 Systems of Linear Equations: Gaussian Elimination

Up until now, when we concerned ourselves with solving different types of equations there was only one equation to solve at a time. Given an equation  $f(x) = g(x)$ , we could check our solutions geometrically by finding where the graphs of  $y = f(x)$  and  $y = g(x)$  intersect. The  $x$ -coordinates of these intersection points correspond to the solutions to the equation  $f(x) = g(x)$ , and the  $y$ -coordinates were largely ignored. If we modify the problem and ask for the intersection points of the graphs of  $y = f(x)$  and  $y = g(x)$ , where both the solution to  $x$  and  $y$  are of interest, we have what is known as a *system of equations*, written as

$$\begin{cases} y = f(x) \\ y = g(x) \end{cases}$$

The ‘curly bracket’ notation means we are to find all *pairs* of points  $(x, y)$  which satisfy *both* equations. We assume the reader has some experience with systems of equations from high school algebra - specifically systems of linear equations comprised of two equations and two unknowns. We encourage the reader to read through Section ?? before proceeding if for no other reason than to refresh themselves on the basic mechanics and vocabulary involved. In order to move this section beyond a review of high school algebra, we define what is meant by a linear equation in  $n$  variables.

**DEFINITION 1.1.** A **linear equation in  $n$  variables**,  $x_1, x_2, \dots, x_n$ , is an equation of the form:

$$a_1x_1 + a_2x_2 + \dots + a_nx_n = c,$$

where  $a_1, a_2, \dots, a_n$  and  $c$  are real numbers and at least one of  $a_1, a_2, \dots, a_n$  is nonzero.

Instead of using more familiar variables like  $x, y$ , and even  $z$  and/or  $w$  in Definition 1.1, we use subscripts to distinguish the different variables. We have no idea how many variables may be involved, so we use numbers to distinguish them instead of letters. (There is an endless supply of distinct numbers.)

As an example, the linear equation  $3x_1 - x_2 = 4$  represents the same relationship between the variables  $x_1$  and  $x_2$  as the equation  $3x - y = 4$  does between the variables  $x$  and  $y$ . And, just as we cannot combine the terms in the expression  $3x - y$ , we cannot combine the terms in the expression  $3x_1 - x_2$ .

Coupling more than one linear equation in  $n$  variables results in a **system of linear equations in  $n$  variables**. When solving these systems, it becomes increasingly important to keep track of what operations are performed to which equations and to develop a strategy based on the kind of manipulations (substitution and elimination) taught in high school. To this end, we first remind ourselves of the maneuvers which can be applied to a system of linear equations that result in an equivalent system.<sup>1</sup>

**THEOREM 1.1.** Given a system of equations, the following moves will result in an equivalent system:

- Interchange the position of any two equations.
- Replace an equation with a nonzero multiple of itself.<sup>a</sup>
- Replace an equation with itself plus a nonzero multiple of another equation.

<sup>a</sup>That is, an equation which results from multiplying both sides of the equation by the same nonzero number.

<sup>1</sup>That is, a system with the same solution set.

The first move, while it obviously admits an equivalent system, seems silly to state, but our perception will change as we consider more equations and more variables in this, and later sections.

Consider the system of equations

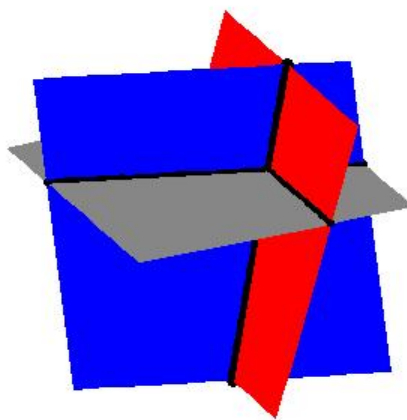
$$\begin{cases} x - \frac{1}{3}y + \frac{1}{2}z = 1 \\ y - \frac{1}{2}z = 4 \\ z = -1 \end{cases}$$

We have  $z = -1$ , so we substitute this into the second equation  $y - \frac{1}{2}(-1) = 4$  to obtain  $y = \frac{7}{2}$ . Substituting  $y = \frac{7}{2}$  and  $z = -1$  into the first equation we get  $x - \frac{1}{3}(\frac{7}{2}) + \frac{1}{2}(-1) = 1$ . This gives  $x = \frac{8}{3}$ . The reader can verify that these values of  $x$ ,  $y$  and  $z$  satisfy all three original equations.

It is tempting for us to write the solution to this system by extending the usual  $(x, y)$  notation to  $(x, y, z)$  and list our solution as  $(\frac{8}{3}, \frac{7}{2}, -1)$ . The question quickly becomes what does an 'ordered triple' like  $(\frac{8}{3}, \frac{7}{2}, -1)$  represent? Just as ordered pairs are used to locate points on the two-dimensional plane, ordered triples can be used to locate points in space.<sup>2</sup>

Moreover, just as equations involving the variables  $x$  and  $y$  describe graphs of one-dimensional lines and curves in the two-dimensional plane, equations involving variables  $x$ ,  $y$ , and  $z$  describe objects called *surfaces* in three-dimensional space. Each of the equations in the above system can be visualized as a plane situated in three-space. Geometrically, the system is trying to find the intersection, or common point, of all three planes. If you imagine three sheets of notebook paper each representing a portion of these planes, you will start to see the complexities involved in how three such planes can intersect.

Below is a sketch of the three planes. It turns out that any two of these planes intersect in a line,<sup>3</sup> so our intersection point is where all three of these lines meet.



<sup>2</sup>You were asked to think about this in Exercise ?? in Section ??.

<sup>3</sup>These lines are described by 'parametric solutions' to the systems formed by taking any two of these equations by themselves. (Again, see Section ??.) We'll see an example of this sort of solution in this section shortly.

Since the geometry for equations involving more than two variables is complicated, we will focus our efforts on the algebra. Returning to the system

$$\begin{cases} x - \frac{1}{3}y + \frac{1}{2}z = 1 \\ y - \frac{1}{2}z = 4 \\ z = -1 \end{cases}$$

we note the reason it was so easy to solve is because of its structure. The third equation is solved for  $z$  and the second equation involves only  $y$  and  $z$ . Since the coefficient of  $y$  is 1, it makes it easy to solve for  $y$  using our known value for  $z$ . Lastly, the coefficient of  $x$  in the first equation is 1 making it easy to substitute the known values of  $y$  and  $z$  and then solve for  $x$ .

We formalize this pattern below for the most general systems of linear equations. Again, we use subscripted variables to describe the general case. The variable with the smallest subscript in a given equation is typically called the *leading variable* of that equation.

**DEFINITION 1.2.** A system of linear equations with variables  $x_1, x_2, \dots, x_n$  is said to be in **triangular form** provided all of the following conditions hold:

1. The subscripts of the variables in each equation are always increasing from left to right.
2. The leading variable in each equation has coefficient 1.
3. The subscript on the leading variable in a given equation is greater than the subscript on the leading variable in the equation above it.
4. Any equation without variables<sup>a</sup> cannot be placed above an equation with variables.

<sup>a</sup>necessarily an identity or contradiction

In our previous system, if we make the obvious choices  $x = x_1$ ,  $y = x_2$ , and  $z = x_3$ , we see that the system is in triangular form.<sup>4</sup> An example of a more complicated system in triangular form is

$$\begin{cases} x_1 - 4x_3 + x_4 - x_6 = 6 \\ x_2 + 2x_3 = 1 \\ x_4 + 3x_5 - x_6 = 8 \\ x_5 + 9x_6 = 10 \end{cases}$$

Our goal henceforth will be to transform a given system of linear equations into triangular form using the moves in Theorem 1.1.

**EXAMPLE 1.1.1.** Use Theorem 1.1 to put the following systems into triangular form and then solve the system if possible. Classify each system as consistent independent, consistent dependent, or inconsistent.<sup>5</sup>

<sup>4</sup>If letters are used instead of subscripted variables, Definition 1.2 can be suitably modified using alphabetical order of the variables instead of numerical order on the subscripts of the variables.

<sup>5</sup>See Section ?? for a review of these terms.

$$1. \begin{cases} 3x - y + z = 3 \\ 2x - 4y + 3z = 16 \\ x - y + z = 5 \end{cases} \quad 2. \begin{cases} 2x + 3y - z = 1 \\ 10x - z = 2 \\ 4x - 9y + 2z = 5 \end{cases} \quad 3. \begin{cases} 3x_1 + x_2 + x_4 = 6 \\ 2x_1 + x_2 - x_3 = 4 \\ x_2 - 3x_3 - 2x_4 = 0 \end{cases}$$

**Solution.** For definitiveness, we label the topmost equation in each system  $E1$ , the equation beneath that  $E2$ , and so forth.

1. We put the system in triangular form using an algorithm known as *Gaussian Elimination*. Starting with  $x$ , we transform the system so that conditions 2 and 3 in Definition 1.2 are satisfied. Then we move on to the next variable, in this case  $y$ , and repeat.

Since the variables in all of the equations have a consistent ordering from left to right, our first move is to get an  $x$  in  $E1$ 's spot with a coefficient of 1. While there are many ways to do this, the easiest is to apply the first move listed in Theorem 1.1 and interchange  $E1$  and  $E3$ .

$$\begin{cases} (E1) & 3x - y + z = 3 \\ (E2) & 2x - 4y + 3z = 16 \\ (E3) & x - y + z = 5 \end{cases} \xrightarrow{\text{Switch } E1 \text{ and } E3} \begin{cases} (E1) & x - y + z = 5 \\ (E2) & 2x - 4y + 3z = 16 \\ (E3) & 3x - y + z = 3 \end{cases}$$

To satisfy Definition 1.2, we need to eliminate the  $x$ 's from  $E2$  and  $E3$ . We accomplish this by replacing each of them with a sum of themselves and a multiple of  $E1$ . To eliminate the  $x$  from  $E2$ , we need to multiply  $E1$  by  $-2$  then add; to eliminate the  $x$  from  $E3$ , we need to multiply  $E1$  by  $-3$  then add. Applying the third move listed in Theorem 1.1 twice, we get

$$\begin{cases} (E1) & x - y + z = 5 \\ (E2) & 2x - 4y + 3z = 16 \\ (E3) & 3x - y + z = 3 \end{cases} \xrightarrow[\text{Replace } E3 \text{ with } -3E1 + E3]{\text{Replace } E2 \text{ with } -2E1 + E2} \begin{cases} (E1) & x - y + z = 5 \\ (E2) & -2y + z = 6 \\ (E3) & 2y - 2z = -12 \end{cases}$$

Now we enforce the conditions stated in Definition 1.2 for the variable  $y$ . To that end we need to get the coefficient of  $y$  in  $E2$  equal to 1. We apply the second move listed in Theorem 1.1 and replace  $E2$  with itself times  $-\frac{1}{2}$ .

$$\begin{cases} (E1) & x - y + z = 5 \\ (E2) & -2y + z = 6 \\ (E3) & 2y - 2z = -12 \end{cases} \xrightarrow{\text{Replace } E2 \text{ with } -\frac{1}{2}E2} \begin{cases} (E1) & x - y + z = 5 \\ (E2) & y - \frac{1}{2}z = -3 \\ (E3) & 2y - 2z = -12 \end{cases}$$

To eliminate the  $y$  in  $E3$ , we add  $-2E2$  to it.

$$\begin{cases} (E1) & x - y + z = 5 \\ (E2) & y - \frac{1}{2}z = -3 \\ (E3) & 2y - 2z = -12 \end{cases} \xrightarrow{\text{Replace } E3 \text{ with } -2E2 + E3} \begin{cases} (E1) & x - y + z = 5 \\ (E2) & y - \frac{1}{2}z = -3 \\ (E3) & -z = -6 \end{cases}$$

Finally, we apply the second move from Theorem 1.1 one last time and multiply  $E3$  by  $-1$  to satisfy the conditions of Definition 1.2 for the variable  $z$ .

$$\left\{ \begin{array}{lcl} (E1) & x - y + z & = 5 \\ (E2) & y - \frac{1}{2}z & = -3 \\ (E3) & -z & = -6 \end{array} \right. \xrightarrow{\text{Replace } E3 \text{ with } -1E3} \left\{ \begin{array}{lcl} (E1) & x - y + z & = 5 \\ (E2) & y - \frac{1}{2}z & = -3 \\ (E3) & z & = 6 \end{array} \right.$$

Substituting  $z = 6$  into  $E2$  gives  $y - 3 = -3$  so that  $y = 0$ . With  $y = 0$  and  $z = 6$ ,  $E1$  becomes  $x - 0 + 6 = 5$ , or  $x = -1$ . Hence, our solution is  $(-1, 0, 6)$ . We leave it to the reader to check that substituting the respective values for  $x$ ,  $y$ , and  $z$  into the original system results in three identities.

Since there is a solution to the system, the system is classified as consistent. Since there are no free variables,<sup>6</sup> the system is classified as independent.

2. Proceeding as above, our first step is to get an equation with  $x$  in the  $E1$  position with 1 as its coefficient. Since there is no easy fix, we multiply  $E1$  by  $\frac{1}{2}$ .

$$\left\{ \begin{array}{lcl} (E1) & 2x + 3y - z & = 1 \\ (E2) & 10x - z & = 2 \\ (E3) & 4x - 9y + 2z & = 5 \end{array} \right. \xrightarrow{\text{Replace } E1 \text{ with } \frac{1}{2}E1} \left\{ \begin{array}{lcl} (E1) & x + \frac{3}{2}y - \frac{1}{2}z & = \frac{1}{2} \\ (E2) & 10x - z & = 2 \\ (E3) & 4x - 9y + 2z & = 5 \end{array} \right.$$

Now it's time to take care of the  $x$ 's in  $E2$  and  $E3$ .

$$\left\{ \begin{array}{lcl} (E1) & x + \frac{3}{2}y - \frac{1}{2}z & = \frac{1}{2} \\ (E2) & 10x - z & = 2 \\ (E3) & 4x - 9y + 2z & = 5 \end{array} \right. \xrightarrow[\text{Replace } E3 \text{ with } -4E1 + E3]{\text{Replace } E2 \text{ with } -10E1 + E2} \left\{ \begin{array}{lcl} (E1) & x + \frac{3}{2}y - \frac{1}{2}z & = \frac{1}{2} \\ (E2) & -15y + 4z & = -3 \\ (E3) & -15y + 4z & = 3 \end{array} \right.$$

Our next step is to get the coefficient of  $y$  in  $E2$  equal to 1. To that end, we have

$$\left\{ \begin{array}{lcl} (E1) & x + \frac{3}{2}y - \frac{1}{2}z & = \frac{1}{2} \\ (E2) & -15y + 4z & = -3 \\ (E3) & -15y + 4z & = 3 \end{array} \right. \xrightarrow{\text{Replace } E2 \text{ with } -\frac{1}{15}E2} \left\{ \begin{array}{lcl} (E1) & x + \frac{3}{2}y - \frac{1}{2}z & = \frac{1}{2} \\ (E2) & y - \frac{4}{15}z & = \frac{1}{5} \\ (E3) & -15y + 4z & = 3 \end{array} \right.$$

Finally, we rid  $E3$  of  $y$ .

$$\left\{ \begin{array}{lcl} (E1) & x + \frac{3}{2}y - \frac{1}{2}z & = \frac{1}{2} \\ (E2) & y - \frac{4}{15}z & = \frac{1}{5} \\ (E3) & -15y + 4z & = 3 \end{array} \right. \xrightarrow{\text{Replace } E3 \text{ with } 15E2 + E3} \left\{ \begin{array}{lcl} (E1) & x - y + z & = 5 \\ (E2) & y - \frac{1}{2}z & = -3 \\ (E3) & 0 & = 6 \end{array} \right.$$

The last equation,  $0 = 6$ , is a contradiction so the system has no solution. According to Theorem 1.1, since this system has no solutions, neither does the original, thus we have an inconsistent system.

<sup>6</sup>Again, see Section ?? for a review of this concept, if needed.

3. For our last system, we begin by multiplying  $E1$  by  $\frac{1}{3}$  to get a coefficient of 1 on  $x_1$ .

$$\left\{ \begin{array}{lcl} (E1) & 3x_1 + x_2 + x_4 & = 6 \\ (E2) & 2x_1 + x_2 - x_3 & = 4 \\ (E3) & x_2 - 3x_3 - 2x_4 & = 0 \end{array} \right. \xrightarrow{\text{Replace } E1 \text{ with } \frac{1}{3}E1} \left\{ \begin{array}{lcl} (E1) & x_1 + \frac{1}{3}x_2 + \frac{1}{3}x_4 & = 2 \\ (E2) & 2x_1 + x_2 - x_3 & = 4 \\ (E3) & x_2 - 3x_3 - 2x_4 & = 0 \end{array} \right.$$

Next we eliminate  $x_1$  from  $E2$

$$\left\{ \begin{array}{lcl} (E1) & x_1 + \frac{1}{3}x_2 + \frac{1}{3}x_4 & = 2 \\ (E2) & 2x_1 + x_2 - x_3 & = 4 \\ (E3) & x_2 - 3x_3 - 2x_4 & = 0 \end{array} \right. \xrightarrow[\text{with } -2E1 + E2]{\text{Replace } E2} \left\{ \begin{array}{lcl} (E1) & x_1 + \frac{1}{3}x_2 + \frac{1}{3}x_4 & = 2 \\ (E2) & \frac{1}{3}x_2 - x_3 - \frac{2}{3}x_4 & = 0 \\ (E3) & x_2 - 3x_3 - 2x_4 & = 0 \end{array} \right.$$

We switch  $E2$  and  $E3$  to get a coefficient of 1 for  $x_2$ .

$$\left\{ \begin{array}{lcl} (E1) & x_1 + \frac{1}{3}x_2 + \frac{1}{3}x_4 & = 2 \\ (E2) & \frac{1}{3}x_2 - x_3 - \frac{2}{3}x_4 & = 0 \\ (E3) & x_2 - 3x_3 - 2x_4 & = 0 \end{array} \right. \xrightarrow{\text{Switch } E2 \text{ and } E3} \left\{ \begin{array}{lcl} (E1) & x_1 + \frac{1}{3}x_2 + \frac{1}{3}x_4 & = 2 \\ (E2) & x_2 - 3x_3 - 2x_4 & = 0 \\ (E3) & \frac{1}{3}x_2 - x_3 - \frac{2}{3}x_4 & = 0 \end{array} \right.$$

Finally, we eliminate  $x_2$  in  $E3$ .

$$\left\{ \begin{array}{lcl} (E1) & x_1 + \frac{1}{3}x_2 + \frac{1}{3}x_4 & = 2 \\ (E2) & x_2 - 3x_3 - 2x_4 & = 0 \\ (E3) & \frac{1}{3}x_2 - x_3 - \frac{2}{3}x_4 & = 0 \end{array} \right. \xrightarrow[\text{with } -\frac{1}{3}E2 + E3]{\text{Replace } E3} \left\{ \begin{array}{lcl} (E1) & x_1 + \frac{1}{3}x_2 + \frac{1}{3}x_4 & = 2 \\ (E2) & x_2 - 3x_3 - 2x_4 & = 0 \\ (E3) & 0 & = 0 \end{array} \right.$$

Equation  $E3$  reduces to  $0 = 0$ , which is always true. Since we have no equations with  $x_3$  or  $x_4$  as leading variables, they are both 'free' variables so we have a consistent dependent system.

We 'parametrize' the solution set by letting  $x_3 = s$  and  $x_4 = t$ . From  $E2$ , we get  $x_2 = 3s + 2t$ . Substituting this and  $x_4 = t$  into  $E1$ , we have  $x_1 + \frac{1}{3}(3s + 2t) + \frac{1}{3}t = 2$  which gives  $x_1 = 2 - s - t$ . Our solution is the set  $\{(2 - s - t, 3s + 2t, s, t) \mid -\infty < s, t < \infty\}$ .<sup>7</sup> We leave it to the reader to verify that the substitutions  $x_1 = 2 - s - t$ ,  $x_2 = 3s + 2t$ ,  $x_3 = s$  and  $x_4 = t$  satisfy the equations in the original system regardless of the choices made for the parameters  $s$  and  $t$ .  $\square$

Like all algorithms, Gaussian Elimination has the advantage of always producing what we need, but it can also be inefficient at times. For example, when solving the second system in Example 1.1.1, it is clear after we eliminated the  $x$ 's in the second step to get the system

$$\left\{ \begin{array}{lcl} (E1) & x + \frac{3}{2}y - \frac{1}{2}z & = \frac{1}{2} \\ (E2) & -15y + 4z & = -3 \\ (E3) & -15y + 4z & = 3 \end{array} \right.$$

<sup>7</sup>Here, any choice of  $s$  and  $t$  determines a point in 4-dimensional space. Yeah, we have trouble visualizing that, too.

that equations  $E2$  and  $E3$ , taken together, produce a contradiction. (We have identical left hand sides and different right hand sides.) However, the algorithm takes an additional two steps to reach this conclusion. We also note that substitution in Gaussian Elimination is delayed until all the elimination is done, whence the name *back-substitution*. This may also be inefficient in many cases.

Lastly, we note that the last system in Example 1.1.1 is underdetermined,<sup>8</sup> and as it is consistent, we necessarily have free variables in our answer. We close this section with a standard ‘mixture’ type application of systems of linear equations which features an application of a consistent dependent system.

**EXAMPLE 1.1.2.** Lucas needs to create a 500 milliliters (mL) of a 40% acid solution. He has stock solutions of 30% and 90% acid as well as all of the distilled water he wants. Set-up and solve a system of linear equations which determines all of the possible combinations of the stock solutions and water which would produce the required solution.

**Solution.** We are after three unknowns, the amount (in mL) of the 30% stock solution (which we’ll call  $x$ ), the amount (in mL) of the 90% stock solution (which we’ll call  $y$ ) and the amount (in mL) of water (which we’ll call  $w$ ). We now need to determine some relationships between these variables.

Our goal is to produce 500 milliliters of a 40% acid solution. This product has two defining characteristics. First, it must be 500 mL; second, it must be 40% acid. We take each of these qualities in turn.

First, the total volume of 500 mL must be the sum of the volumes of the two stock solutions and the water:

$$\text{amount of 30\% stock solution} + \text{amount of 90\% stock solution} + \text{amount of water} = 500 \text{ mL}$$

Using our defined variables, this reduces to  $x + y + w = 500$ .

Next, we need to make sure the final solution is 40% acid. Since water contains no acid, the acid will come from the stock solutions only. We find 40% of 500 mL to be 200 mL which means the final solution must contain 200 mL of acid. We have

$$\text{amount of acid in 30\% stock solution} + \text{amount of acid 90\% stock solution} = 200 \text{ mL}$$

The amount of acid in  $x$  mL of 30% stock is  $0.30x$  and the amount of acid in  $y$  mL of 90% solution is  $0.90y$ . We have  $0.30x + 0.90y = 200$ . Converting to fractions,<sup>9</sup> our system of equations becomes

$$\begin{cases} x + y + w = 500 \\ \frac{3}{10}x + \frac{9}{10}y = 200 \end{cases}$$

We first eliminate the  $x$  from the second equation

$$\begin{cases} (E1) & x + y + w = 500 \\ (E2) & \frac{3}{10}x + \frac{9}{10}y = 200 \end{cases} \xrightarrow{\text{Replace } E2 \text{ with } -\frac{3}{10}E1 + E2} \begin{cases} (E1) & x + y + w = 500 \\ (E2) & \frac{3}{5}y - \frac{3}{10}w = 50 \end{cases}$$

Next, we get a coefficient of 1 on the leading variable in  $E2$

$$\begin{cases} (E1) & x + y + w = 500 \\ (E2) & \frac{3}{5}y - \frac{3}{10}w = 50 \end{cases} \xrightarrow{\text{Replace } E2 \text{ with } \frac{5}{3}E2} \begin{cases} (E1) & x + y + w = 500 \\ (E2) & y - \frac{1}{2}w = \frac{250}{3} \end{cases}$$

<sup>8</sup>Recall this means we have fewer equations than unknowns.

<sup>9</sup>We do this only because we believe students can use all of the practice with fractions they can get!



Notice that we have no equation to determine  $w$ , and as such,  $w$  is free. Setting  $w = t$  in  $E2$ , we get  $y = \frac{1}{2}t + \frac{250}{3}$ . Substituting for  $w$  and  $y$  in  $E1$  gives  $x + (\frac{1}{2}t + \frac{250}{3}) + t = 500$  so that  $x = -\frac{3}{2}t + \frac{1250}{3}$ .

This system is consistent, dependent and its solution set is  $\{(-\frac{3}{2}t + \frac{1250}{3}, \frac{1}{2}t + \frac{250}{3}, t) \mid -\infty < t < \infty\}$ .

While this answer checks algebraically, we have neglected to take into account that  $x$ ,  $y$  and  $w$ , being amounts of acid and water, need to be nonnegative. That is,  $x \geq 0$ ,  $y \geq 0$  and  $w \geq 0$ .

The constraint  $x \geq 0$  gives us  $-\frac{3}{2}t + \frac{1250}{3} \geq 0$ , or  $t \leq \frac{2500}{9}$ . From  $y \geq 0$ , we get  $\frac{1}{2}t + \frac{250}{3} \geq 0$  or  $t \geq -\frac{500}{3}$ .

The condition  $z \geq 0$  yields  $t \geq 0$ , and we see that when we take the set theoretic intersection of these intervals, we get  $0 \leq t \leq \frac{2500}{9}$ . This gives our final answer is  $\{(-\frac{3}{2}t + \frac{1250}{3}, \frac{1}{2}t + \frac{250}{3}, t) \mid 0 \leq t \leq \frac{2500}{9}\}$ .

Of what practical use is our answer? Suppose there is only 100 mL of the 90% solution remaining and it is due to expire. Can we use all of it to make our required solution? We would have  $y = 100$  so that  $\frac{1}{2}t + \frac{250}{3} = 100$ , and we get  $t = \frac{100}{3}$ . This means the amount of 30% solution required is  $x = -\frac{3}{2}t + \frac{1250}{3} = -\frac{3}{2}(\frac{100}{3}) + \frac{1250}{3} = \frac{1100}{3}$  mL, and for the water,  $w = t = \frac{100}{3}$  mL. The reader is invited to check that mixing these three amounts of our constituent solutions produces the required 40% acid mix.  $\square$

## 1.1.1 Exercises

(Review Exercises) In Exercises 1 - 8, take a trip down memory lane and solve the given system using substitution and/or elimination. Classify each system as consistent independent, consistent dependent, or inconsistent. Check your answers both algebraically and graphically.

$$1. \begin{cases} x + 2y = 5 \\ x = 6 \end{cases}$$

$$2. \begin{cases} 2y - 3x = 1 \\ y = -3 \end{cases}$$

$$3. \begin{cases} \frac{x+2y}{4} = -5 \\ \frac{3x-y}{2} = 1 \end{cases}$$

$$4. \begin{cases} \frac{2}{3}x - \frac{1}{5}y = 3 \\ \frac{1}{2}x + \frac{3}{4}y = 1 \end{cases}$$

$$5. \begin{cases} \frac{1}{2}x - \frac{1}{3}y = -1 \\ 2y - 3x = 6 \end{cases}$$

$$6. \begin{cases} x + 4y = 6 \\ \frac{1}{12}x + \frac{1}{3}y = \frac{1}{2} \end{cases}$$

$$7. \begin{cases} 3y - \frac{3}{2}x = -\frac{15}{2} \\ \frac{1}{2}x - y = \frac{3}{2} \end{cases}$$

$$8. \begin{cases} \frac{5}{6}x + \frac{5}{3}y = -\frac{7}{3} \\ -\frac{10}{3}x - \frac{20}{3}y = 10 \end{cases}$$

In Exercises 9 - 26, put each system of linear equations into triangular form and solve the system if possible. Classify each system as consistent independent, consistent dependent, or inconsistent.

$$9. \begin{cases} -5x + y = 17 \\ x + y = 5 \end{cases}$$

$$10. \begin{cases} x + y + z = 3 \\ 2x - y + z = 0 \\ -3x + 5y + 7z = 7 \end{cases}$$

$$11. \begin{cases} 4x - y + z = 5 \\ 2y + 6z = 30 \\ x + z = 5 \end{cases}$$

$$12. \begin{cases} 4x - y + z = 5 \\ 2y + 6z = 30 \\ x + z = 6 \end{cases}$$

$$13. \begin{cases} x + y + z = -17 \\ y - 3z = 0 \end{cases}$$

$$14. \begin{cases} x - 2y + 3z = 7 \\ -3x + y + 2z = -5 \\ 2x + 2y + z = 3 \end{cases}$$

$$15. \begin{cases} 3x - 2y + z = -5 \\ x + 3y - z = 12 \\ x + y + 2z = 0 \end{cases}$$

$$16. \begin{cases} 2x - y + z = -1 \\ 4x + 3y + 5z = 1 \\ 5y + 3z = 4 \end{cases}$$

$$17. \begin{cases} x - y + z = -4 \\ -3x + 2y + 4z = -5 \\ x - 5y + 2z = -18 \end{cases}$$

$$18. \begin{cases} 2x - 4y + z = -7 \\ x - 2y + 2z = -2 \\ -x + 4y - 2z = 3 \end{cases}$$

$$19. \begin{cases} 2x - y + z = 1 \\ 2x + 2y - z = 1 \\ 3x + 6y + 4z = 9 \end{cases}$$

$$20. \begin{cases} x - 3y - 4z = 3 \\ 3x + 4y - z = 13 \\ 2x - 19y - 19z = 2 \end{cases}$$

$$21. \begin{cases} x + y + z = 4 \\ 2x - 4y - z = -1 \\ x - y = 2 \end{cases}$$

$$22. \begin{cases} x - y + z = 8 \\ 3x + 3y - 9z = -6 \\ 7x - 2y + 5z = 39 \end{cases}$$

$$23. \begin{cases} 2x - 3y + z = -1 \\ 4x - 4y + 4z = -13 \\ 6x - 5y + 7z = -25 \end{cases}$$

$$24. \begin{cases} 2x_1 + x_2 - 12x_3 - x_4 = 16 \\ -x_1 + x_2 + 12x_3 - 4x_4 = -5 \\ 3x_1 + 2x_2 - 16x_3 - 3x_4 = 25 \\ x_1 + 2x_2 - 5x_4 = 11 \end{cases}$$

$$25. \begin{cases} x_1 - x_3 = -2 \\ 2x_2 - x_4 = 0 \\ x_1 - 2x_2 + x_3 = 0 \\ -x_3 + x_4 = 1 \end{cases}$$

$$26. \begin{cases} x_1 - x_2 - 5x_3 + 3x_4 = -1 \\ x_1 + x_2 + 5x_3 - 3x_4 = 0 \\ x_2 + 5x_3 - 3x_4 = 1 \\ x_1 - 2x_2 - 10x_3 + 6x_4 = -1 \end{cases}$$

27. Find two other forms of the parametric solution to Exercise 11 above by reorganizing the equations so that  $x$  or  $y$  can be the free variable.
28. A local buffet charges \$7.50 per person for the basic buffet and \$9.25 for the deluxe buffet (which includes crab legs.) If 27 diners went out to eat and the total bill was \$227.00 before taxes, how many chose the basic buffet and how many chose the deluxe buffet?
29. At The Old Home Fill'er Up and Keep on a-Truckin' Cafe, Mavis mixes two different types of coffee beans to produce a house blend. The first type costs \$3 per pound and the second costs \$8 per pound. How much of each type does Mavis use to make 50 pounds of a blend which costs \$6 per pound?
30. Skippy has a total of \$10,000 to split between two investments. One account offers 3% simple interest, and the other account offers 8% simple interest. For tax reasons, he can only earn \$500 in interest the entire year. How much money should Skippy invest in each account to earn \$500 in interest for the year?
31. A 10% salt solution is to be mixed with pure water to produce 75 gallons of a 3% salt solution. How much of each are needed?
32. At The Crispy Critter's Head Shop and Patchouli Emporium along with their dried up weeds, sunflower seeds and astrological postcards they sell an herbal tea blend. By weight, Type I herbal tea is 30% peppermint, 40% rose hips and 30% chamomile, Type II has percents 40%, 20% and 40%, respectively, and Type III has percents 35%, 30% and 35%, respectively. How much of each Type of tea is needed to make 2 pounds of a new blend of tea that is equal parts peppermint, rose hips and chamomile?
33. Discuss with your classmates how you would approach Exercise 32 above if they needed to use up a pound of Type I tea to make room on the shelf for a new canister.
34. If you were to try to make 100 mL of a 60% acid solution using stock solutions at 20% and 40%, respectively, what would the triangular form of the resulting system look like? Explain.

**1.1.2 Answers**

1. Consistent independent

Solution  $(6, -\frac{1}{2})$ 

2. Consistent independent

Solution  $(-\frac{7}{3}, -3)$ 

3. Consistent independent

Solution  $(-\frac{16}{7}, -\frac{62}{7})$ 

4. Consistent independent

Solution  $(\frac{49}{12}, -\frac{25}{18})$ 

5. Consistent dependent

Solution  $(t, \frac{3}{2}t + 3)$   
for all real numbers  $t$ 

6. Consistent dependent

Solution  $(6 - 4t, t)$   
for all real numbers  $t$ 

7. Inconsistent

No solution

8. Inconsistent

No solution

Because triangular form is not unique, we give only one possible answer to that part of the question. Yours may be different and still be correct.

$$9. \begin{cases} x + y = 5 \\ y = 7 \end{cases}$$

Consistent independent

Solution  $(-2, 7)$ 

$$10. \begin{cases} x - \frac{5}{3}y - \frac{7}{3}z = -\frac{7}{3} \\ y + \frac{5}{4}z = 2 \\ z = 0 \end{cases}$$

Consistent independent

Solution  $(1, 2, 0)$ 

$$11. \begin{cases} x - \frac{1}{4}y + \frac{1}{4}z = \frac{5}{4} \\ y + 3z = 15 \\ 0 = 0 \end{cases}$$

Consistent dependent

Solution  $(-t + 5, -3t + 15, t)$   
for all real numbers  $t$ 

$$12. \begin{cases} x - \frac{1}{4}y + \frac{1}{4}z = \frac{5}{4} \\ y + 3z = 15 \\ 0 = 1 \end{cases}$$

Inconsistent

No solution

$$13. \begin{cases} x + y + z = -17 \\ y - 3z = 0 \end{cases}$$

Consistent dependent

Solution  $(-4t - 17, 3t, t)$   
for all real numbers  $t$ 

$$14. \begin{cases} x - 2y + 3z = 7 \\ y - \frac{11}{5}z = -\frac{16}{5} \\ z = 1 \end{cases}$$

Consistent independent

Solution  $(2, -1, 1)$

15. 
$$\begin{cases} x + y + 2z = 0 \\ y - \frac{3}{2}z = 6 \\ z = -2 \end{cases}$$
 Consistent independent  
Solution  $(1, 3, -2)$
16. 
$$\begin{cases} x - \frac{1}{2}y + \frac{1}{2}z = -\frac{1}{2} \\ y + \frac{3}{5}z = \frac{3}{5} \\ 0 = 1 \end{cases}$$
 Inconsistent  
no solution
17. 
$$\begin{cases} x - y + z = -4 \\ y - 7z = 17 \\ z = -2 \end{cases}$$
 Consistent independent  
Solution  $(1, 3, -2)$
18. 
$$\begin{cases} x - 2y + 2z = -2 \\ y = \frac{1}{2} \\ z = 1 \end{cases}$$
 Consistent independent  
Solution  $(-3, \frac{1}{2}, 1)$
19. 
$$\begin{cases} x - \frac{1}{2}y + \frac{1}{2}z = \frac{1}{2} \\ y - \frac{2}{3}z = 0 \\ z = 1 \end{cases}$$
 Consistent independent  
Solution  $(\frac{1}{3}, \frac{2}{3}, 1)$
20. 
$$\begin{cases} x - 3y - 4z = 3 \\ y + \frac{11}{13}z = \frac{4}{13} \\ 0 = 0 \end{cases}$$
 Consistent dependent  
Solution  $(\frac{19}{13}t + \frac{51}{13}, -\frac{11}{13}t + \frac{4}{13}, t)$   
for all real numbers  $t$
21. 
$$\begin{cases} x + y + z = 4 \\ y + \frac{1}{2}z = \frac{3}{2} \\ 0 = 1 \end{cases}$$
 Inconsistent  
no solution
22. 
$$\begin{cases} x - y + z = 8 \\ y - 2z = -5 \\ z = 1 \end{cases}$$
 Consistent independent  
Solution  $(4, -3, 1)$
23. 
$$\begin{cases} x - \frac{3}{2}y + \frac{1}{2}z = -\frac{1}{2} \\ y + z = -\frac{11}{2} \\ 0 = 0 \end{cases}$$
 Consistent dependent  
Solution  $(-2t - \frac{35}{4}, -t - \frac{11}{2}, t)$   
for all real numbers  $t$
24. 
$$\begin{cases} x_1 + \frac{2}{3}x_2 - \frac{16}{3}x_3 - x_4 = \frac{25}{3} \\ x_2 + 4x_3 - 3x_4 = 2 \\ 0 = 0 \\ 0 = 0 \end{cases}$$
 Consistent dependent  
Solution  $(8s - t + 7, -4s + 3t + 2, s, t)$   
for all real numbers  $s$  and  $t$
25. 
$$\begin{cases} x_1 - x_3 = -2 \\ x_2 - \frac{1}{2}x_4 = 0 \\ x_3 - \frac{1}{2}x_4 = 1 \\ x_4 = 4 \end{cases}$$

	Consistent independent	Solution (1, 2, 3, 4)
26.	$\begin{cases} x_1 - x_2 - 5x_3 + 3x_4 = -1 \\ x_2 + 5x_3 - 3x_4 = \frac{1}{2} \\ 0 = 1 \\ 0 = 0 \end{cases}$	<p>Inconsistent</p> <p>No solution</p>

27. If  $x$  is the free variable then the solution is  $(t, 3t, -t + 5)$  and if  $y$  is the free variable then the solution is  $(\frac{1}{3}t, t, -\frac{1}{3}t + 5)$ .
28. 13 chose the basic buffet and 14 chose the deluxe buffet.
29. Mavis needs 20 pounds of \$3 per pound coffee and 30 pounds of \$8 per pound coffee.
30. Skippy needs to invest \$6000 in the 3% account and \$4000 in the 8% account.
31. 22.5 gallons of the 10% solution and 52.5 gallons of pure water.
32.  $\frac{4}{3} - \frac{1}{2}t$  pounds of Type I,  $\frac{2}{3} - \frac{1}{2}t$  pounds of Type II and  $t$  pounds of Type III where  $0 \leq t \leq \frac{4}{3}$ .

## 1.2 Systems of Linear Equations: Augmented Matrices

In Section 1.1 we introduced Gaussian Elimination as a means of transforming a system of linear equations into triangular form with the ultimate goal of producing an equivalent system of linear equations which is easier to solve. If we take a step back and study the process, we see that all of our moves are determined entirely by the *coefficients* of the variables involved, and not the variables themselves. Much the same thing happened when we studied long division in Section ???. Just as we developed synthetic division to streamline that process, in this section, we introduce a similar bookkeeping device to help us solve systems of linear equations, the *matrix*.

A *matrix* as a rectangular array of real numbers. We typically enclose matrices with ‘square brackets’ of the likes of ‘[’ and ‘]’, and we size matrices by the number of rows and columns they have. For example, the *size* (sometimes called the *dimension*) of

$$\begin{bmatrix} 3 & 0 & -1 \\ 2 & -5 & 10 \end{bmatrix}$$

is  $2 \times 3$  because it has 2 rows and 3 columns. The individual numbers in a matrix are called its *entries* and are usually labeled with double subscripts: the first tells which row the element is in and the second tells which column it is in. The rows are numbered from top to bottom and the columns are numbered from left to right. Matrices themselves are usually denoted by uppercase letters ( $A$ ,  $B$ ,  $C$ , etc.) while their entries are usually denoted by the corresponding letter. So, for instance, if we have

$$A = \begin{bmatrix} 3 & 0 & -1 \\ 2 & -5 & 10 \end{bmatrix}$$

then  $a_{11} = 3$ ,  $a_{12} = 0$ ,  $a_{13} = -1$ ,  $a_{21} = 2$ ,  $a_{22} = -5$ , and  $a_{23} = 10$ . We shall explore matrices as mathematical objects with their own algebra in Section 1.3 and introduce them here solely as a bookkeeping device. Consider the system of linear equations from number 2 in Example 1.1.1

$$\begin{cases} (E1) & 2x + 3y - z = 1 \\ (E2) & \quad 10x - z = 2 \\ (E3) & 4x - 9y + 2z = 5 \end{cases}$$

We encode this system into a matrix by assigning each equation to a corresponding row. Within that row, each variable and the constant gets its own column, and to separate the variables on the left hand side of the equation from the constants on the right hand side, we use a vertical bar,  $|$ . Note that in  $E2$ , since  $y$  is not present, we record its coefficient as 0. The matrix associated with this system is

$$\begin{array}{l} (E1) \rightarrow \\ (E2) \rightarrow \\ (E3) \rightarrow \end{array} \begin{array}{c} x \quad y \quad z \quad c \\ \left[ \begin{array}{ccc|c} 2 & 3 & -1 & 1 \\ 10 & 0 & -1 & 2 \\ 4 & -9 & 2 & 5 \end{array} \right] \end{array}$$

This matrix is called an *augmented matrix* because the column containing the constants is appended to the matrix containing the coefficients.<sup>1</sup> To solve this system, we can use the same kind operations on

<sup>1</sup>We shall study the coefficient and constant matrices separately in Section 1.3.

the *rows* of the matrix that we performed on the *equations* of the system. More specifically, we have the following analog of Theorem 1.1 below.

**THEOREM 1.2. Row Operations:** Given an augmented matrix for a system of linear equations, the following row operations produce an augmented matrix which corresponds to an equivalent system of linear equations.

- Interchange any two rows.
- Replace a row with a nonzero multiple of itself.<sup>a</sup>
- Replace a row with itself plus a nonzero multiple of another row.<sup>b</sup>

<sup>a</sup>That is, the row obtained by multiplying each entry in the row by the same nonzero number.

<sup>b</sup>Where we add entries in corresponding columns.

As a demonstration of the moves in Theorem 1.2, we revisit some of the steps that were used in solving the systems of linear equations in Example 1.1.1 of Section 1.1. The reader is encouraged to perform the indicated operations on the rows of the augmented matrix to see that the machinations are identical to what is done to the coefficients of the variables in the equations. We first see a demonstration of switching two rows using the first step of part 1 in Example 1.1.1.

$$\left\{ \begin{array}{lcl} (E1) & 3x - y + z & = 3 \\ (E2) & 2x - 4y + 3z & = 16 \\ (E3) & x - y + z & = 5 \end{array} \right. \xrightarrow{\text{Switch } E1 \text{ and } E3} \left\{ \begin{array}{lcl} (E1) & x - y + z & = 5 \\ (E2) & 2x - 4y + 3z & = 16 \\ (E3) & 3x - y + z & = 3 \end{array} \right.$$

$$\left[ \begin{array}{ccc|c} 3 & -1 & 1 & 3 \\ 2 & -4 & 3 & 16 \\ 1 & -1 & 1 & 5 \end{array} \right] \xrightarrow{\text{Switch } R1 \text{ and } R3} \left[ \begin{array}{ccc|c} 1 & -1 & 1 & 5 \\ 2 & -4 & 3 & 16 \\ 3 & -1 & 1 & 3 \end{array} \right]$$

Next, we have a demonstration of replacing a row with a nonzero multiple of itself using the first step of part 3 in Example 1.1.1.

$$\left\{ \begin{array}{lcl} (E1) & 3x_1 + x_2 + x_4 & = 6 \\ (E2) & 2x_1 + x_2 - x_3 & = 4 \\ (E3) & x_2 - 3x_3 - 2x_4 & = 0 \end{array} \right. \xrightarrow{\text{Replace } E1 \text{ with } \frac{1}{3}E1} \left\{ \begin{array}{lcl} (E1) & x_1 + \frac{1}{3}x_2 + \frac{1}{3}x_4 & = 2 \\ (E2) & 2x_1 + x_2 - x_3 & = 4 \\ (E3) & x_2 - 3x_3 - 2x_4 & = 0 \end{array} \right.$$

$$\left[ \begin{array}{cccc|c} 3 & 1 & 0 & 1 & 6 \\ 2 & 1 & -1 & 0 & 4 \\ 0 & 1 & -3 & -2 & 0 \end{array} \right] \xrightarrow{\text{Replace } R1 \text{ with } \frac{1}{3}R1} \left[ \begin{array}{cccc|c} 1 & \frac{1}{3} & 0 & \frac{1}{3} & 2 \\ 2 & 1 & -1 & 0 & 4 \\ 0 & 1 & -3 & -2 & 0 \end{array} \right]$$

Finally, we have an example of replacing a row with itself plus a multiple of another row using the second step from part 2 in Example 1.1.1.

$$\left\{ \begin{array}{lcl} (E1) & x + \frac{3}{2}y - \frac{1}{2}z & = \frac{1}{2} \\ (E2) & 10x - z & = 2 \\ (E3) & 4x - 9y + 2z & = 5 \end{array} \right. \xrightarrow{\begin{array}{l} \text{Replace } E2 \text{ with } -10E1 + E2 \\ \text{Replace } E3 \text{ with } -4E1 + E3 \end{array}} \left\{ \begin{array}{lcl} (E1) & x + \frac{3}{2}y - \frac{1}{2}z & = \frac{1}{2} \\ (E2) & -15y + 4z & = -3 \\ (E3) & -15y + 4z & = 3 \end{array} \right.$$



$$\left[ \begin{array}{ccc|c} 1 & \frac{3}{2} & -\frac{1}{2} & \frac{1}{2} \\ 10 & 0 & -1 & 2 \\ 4 & -9 & 2 & 5 \end{array} \right] \xrightarrow[\text{Replace } R3 \text{ with } -4R1 + R3]{\text{Replace } R2 \text{ with } -10R1 + R2} \left[ \begin{array}{ccc|c} 1 & \frac{3}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & -15 & 4 & -3 \\ 0 & -15 & 4 & 3 \end{array} \right]$$

The matrix equivalent of ‘triangular form’ is *row echelon form*. The reader is encouraged to refer to Definition 1.2 for comparison. Note that the analog of ‘leading variable’ of an equation is ‘leading entry’ of a row. Specifically, the first nonzero entry (if it exists) in a row is called the *leading entry* of that row.

**DEFINITION 1.3.** A matrix is said to be in **row echelon form** provided all of the following conditions hold:

1. The first nonzero entry in each row is 1.
2. The leading 1 of a given row must be to the right of the leading 1 of the row above it.
3. Any row of all zeros cannot be placed above a row with nonzero entries.

To solve a system of a linear equations using an augmented matrix, we encode the system into an augmented matrix and apply Gaussian Elimination to the rows to get the matrix into row-echelon form. We then decode the matrix and back substitute. The next example illustrates this nicely.

**EXAMPLE 1.2.1.** Use an augmented matrix to transform the following system of linear equations into triangular form. Solve the system.

$$\begin{cases} 3x - y + z = 8 \\ x + 2y - z = 4 \\ 2x + 3y - 4z = 10 \end{cases}$$

**Solution.** We first encode the system into an augmented matrix.

$$\begin{cases} 3x - y + z = 8 \\ x + 2y - z = 4 \\ 2x + 3y - 4z = 10 \end{cases} \xrightarrow{\text{Encode into the matrix}} \left[ \begin{array}{ccc|c} 3 & -1 & 1 & 8 \\ 1 & 2 & -1 & 4 \\ 2 & 3 & -4 & 10 \end{array} \right]$$

Thinking back to Gaussian Elimination at an equations level, our first order of business is to get  $x$  in  $E1$  with a coefficient of 1. At the matrix level, this means getting a leading 1 in  $R1$ . This is in accordance with the first criteria in Definition 1.3. To that end, we interchange  $R1$  and  $R2$ .

$$\left[ \begin{array}{ccc|c} 3 & -1 & 1 & 8 \\ 1 & 2 & -1 & 4 \\ 2 & 3 & -4 & 10 \end{array} \right] \xrightarrow{\text{Switch } R1 \text{ and } R2} \left[ \begin{array}{ccc|c} 1 & 2 & -1 & 4 \\ 3 & -1 & 1 & 8 \\ 2 & 3 & -4 & 10 \end{array} \right]$$

Our next step is to eliminate the  $x$ 's from  $E2$  and  $E3$ . From a matrix standpoint, this means we need 0's below the leading 1 in  $R1$ . This guarantees the leading 1 in  $R2$  will be to the right of the leading 1 in  $R1$  in accordance with the second requirement of Definition 1.3.

$$\left[ \begin{array}{ccc|c} 1 & 2 & -1 & 4 \\ 3 & -1 & 1 & 8 \\ 2 & 3 & -4 & 10 \end{array} \right] \xrightarrow[\text{Replace } R3 \text{ with } -2R1 + R3]{\text{Replace } R2 \text{ with } -3R1 + R2} \left[ \begin{array}{ccc|c} 1 & 2 & -1 & 4 \\ 0 & -7 & 4 & -4 \\ 0 & -1 & -2 & 2 \end{array} \right]$$

Now we repeat the above process for the variable  $y$ . That is, we need the leading entry in  $R2$  to be 1.

$$\left[ \begin{array}{ccc|c} 1 & 2 & -1 & 4 \\ 0 & -7 & 4 & -4 \\ 0 & -1 & -2 & 2 \end{array} \right] \xrightarrow{\text{Replace } R2 \text{ with } -\frac{1}{7}R2} \left[ \begin{array}{ccc|c} 1 & 2 & -1 & 4 \\ 0 & 1 & -\frac{4}{7} & \frac{4}{7} \\ 0 & -1 & -2 & 2 \end{array} \right]$$

To ensure the leading 1 in  $R3$  is to the right of the leading 1 in  $R2$ , we get a 0 in the second column of  $R3$ .

$$\left[ \begin{array}{ccc|c} 1 & 2 & -1 & 4 \\ 0 & 1 & -\frac{4}{7} & \frac{4}{7} \\ 0 & -1 & -2 & 2 \end{array} \right] \xrightarrow{\text{Replace } R3 \text{ with } R2 + R3} \left[ \begin{array}{ccc|c} 1 & 2 & -1 & 4 \\ 0 & 1 & -\frac{4}{7} & \frac{4}{7} \\ 0 & 0 & -\frac{18}{7} & \frac{18}{7} \end{array} \right]$$

Finally, we get the leading entry in  $R3$  to be 1.

$$\left[ \begin{array}{ccc|c} 1 & 2 & -1 & 4 \\ 0 & 1 & -\frac{4}{7} & \frac{4}{7} \\ 0 & 0 & -\frac{18}{7} & \frac{18}{7} \end{array} \right] \xrightarrow{\text{Replace } R3 \text{ with } -\frac{7}{18}R3} \left[ \begin{array}{ccc|c} 1 & 2 & -1 & 4 \\ 0 & 1 & -\frac{4}{7} & \frac{4}{7} \\ 0 & 0 & 1 & -1 \end{array} \right]$$

Decoding from the matrix gives a system in triangular form

$$\left[ \begin{array}{ccc|c} 1 & 2 & -1 & 4 \\ 0 & 1 & -\frac{4}{7} & \frac{4}{7} \\ 0 & 0 & 1 & -1 \end{array} \right] \xrightarrow{\text{Decode from the matrix}} \begin{cases} x + 2y - z = 4 \\ y - \frac{4}{7}z = \frac{4}{7} \\ z = -1 \end{cases}$$

We get  $z = -1$ ,  $y = \frac{4}{7}z + \frac{4}{7} = \frac{4}{7}(-1) + \frac{4}{7} = 0$  and  $x = -2y + z + 4 = -2(0) + (-1) + 4 = 3$  for a final answer of  $(3, 0, -1)$ . We leave it to the reader to check.  $\square$

As part of Gaussian Elimination, we used row operations to obtain 0's beneath each leading 1 to put the matrix into row echelon form. If we also require that 0's are the only numbers above a leading 1, we have what is known as the *reduced row echelon form* of the matrix.

**DEFINITION 1.4.** A matrix is said to be in **reduced row echelon form** provided both of the following conditions hold:

1. The matrix is in row echelon form.
2. The leading 1s are the only nonzero entry in their respective columns.

Of what significance is the reduced row echelon form of a matrix? To illustrate, let's take the row echelon form from Example 1.2.1 and perform the necessary steps to put into reduced row echelon form. We start by using the leading 1 in  $R3$  to zero out the numbers in the rows above it.

$$\left[ \begin{array}{ccc|c} 1 & 2 & -1 & 4 \\ 0 & 1 & -\frac{4}{7} & \frac{4}{7} \\ 0 & 0 & 1 & -1 \end{array} \right] \xrightarrow[\text{Replace } R2 \text{ with } \frac{4}{7}R3 + R2]{\text{Replace } R1 \text{ with } R3 + R1} \left[ \begin{array}{ccc|c} 1 & 2 & 0 & 3 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{array} \right]$$

Finally, we take care of the 2 in  $R1$  above the leading 1 in  $R2$ .

$$\left[ \begin{array}{ccc|c} 1 & 2 & 0 & 3 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{array} \right] \xrightarrow{\text{Replace } R1 \text{ with } -2R2 + R1} \left[ \begin{array}{ccc|c} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{array} \right]$$

To our surprise and delight, when we decode this matrix, we obtain the solution instantly without having to deal with any back-substitution at all.

$$\left[ \begin{array}{ccc|c} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{array} \right] \xrightarrow{\text{Decode from the matrix}} \begin{cases} x = 3 \\ y = 0 \\ z = -1 \end{cases}$$

Note that in the previous discussion, we could have started with  $R2$  and used it to get a zero above its leading 1 and then done the same for the leading 1 in  $R3$ . By starting with  $R3$ , however, we get more zeros first, and the more zeros there are, the faster the remaining calculations will be.<sup>2</sup> It is also worth noting that while a matrix has several<sup>3</sup> row echelon forms, it has only one reduced row echelon form. The process by which we have put a matrix into reduced row echelon form is called *Gauss-Jordan Elimination*.

EXAMPLE 1.2.2. Solve the following system using an augmented matrix. Use Gauss-Jordan Elimination to put the augmented matrix into reduced row echelon form.

$$\begin{cases} x_2 - 3x_1 + x_4 = 2 \\ 2x_1 + 4x_3 = 5 \\ 4x_2 - x_4 = 3 \end{cases}$$

**Solution.** We first encode the system into a matrix. (Pay attention to the subscripts!)

$$\begin{cases} x_2 - 3x_1 + x_4 = 2 \\ 2x_1 + 4x_3 = 5 \\ 4x_2 - x_4 = 3 \end{cases} \xrightarrow{\text{Encode into the matrix}} \left[ \begin{array}{cccc|c} -3 & 1 & 0 & 1 & 2 \\ 2 & 0 & 4 & 0 & 5 \\ 0 & 4 & 0 & -1 & 3 \end{array} \right]$$

Next, we get a leading 1 in the first column of  $R1$ .

$$\left[ \begin{array}{cccc|c} -3 & 1 & 0 & 1 & 2 \\ 2 & 0 & 4 & 0 & 5 \\ 0 & 4 & 0 & -1 & 3 \end{array} \right] \xrightarrow{\text{Replace } R1 \text{ with } -\frac{1}{3}R1} \left[ \begin{array}{cccc|c} 1 & -\frac{1}{3} & 0 & -\frac{1}{3} & -\frac{2}{3} \\ 2 & 0 & 4 & 0 & 5 \\ 0 & 4 & 0 & -1 & 3 \end{array} \right]$$

Now we eliminate the nonzero entry below our leading 1.

$$\left[ \begin{array}{cccc|c} 1 & -\frac{1}{3} & 0 & -\frac{1}{3} & -\frac{2}{3} \\ 2 & 0 & 4 & 0 & 5 \\ 0 & 4 & 0 & -1 & 3 \end{array} \right] \xrightarrow{\text{Replace } R2 \text{ with } -2R1 + R2} \left[ \begin{array}{cccc|c} 1 & -\frac{1}{3} & 0 & -\frac{1}{3} & -\frac{2}{3} \\ 0 & \frac{2}{3} & 4 & \frac{2}{3} & \frac{19}{3} \\ 0 & 4 & 0 & -1 & 3 \end{array} \right]$$

We proceed to get a leading 1 in  $R2$ .

<sup>2</sup>Carl also finds starting with  $R3$  to be more symmetric, in a purely poetic way.

<sup>3</sup>infinite, in fact

$$\left[ \begin{array}{cccc|c} 1 & -\frac{1}{3} & 0 & -\frac{1}{3} & -\frac{2}{3} \\ 0 & \frac{2}{3} & 4 & \frac{2}{3} & \frac{19}{3} \\ 0 & 4 & 0 & -1 & 3 \end{array} \right] \xrightarrow{\text{Replace } R2 \text{ with } \frac{3}{2}R2} \left[ \begin{array}{cccc|c} 1 & -\frac{1}{3} & 0 & -\frac{1}{3} & -\frac{2}{3} \\ 0 & 1 & 6 & 1 & \frac{19}{2} \\ 0 & 4 & 0 & -1 & 3 \end{array} \right]$$

We now zero out the entry below the leading 1 in  $R2$ .

$$\left[ \begin{array}{cccc|c} 1 & -\frac{1}{3} & 0 & -\frac{1}{3} & -\frac{2}{3} \\ 0 & 1 & 6 & 1 & \frac{19}{2} \\ 0 & 4 & 0 & -1 & 3 \end{array} \right] \xrightarrow{\text{Replace } R3 \text{ with } -4R2 + R3} \left[ \begin{array}{cccc|c} 1 & -\frac{1}{3} & 0 & -\frac{1}{3} & -\frac{2}{3} \\ 0 & 1 & 6 & 1 & \frac{19}{2} \\ 0 & 0 & -24 & -5 & -35 \end{array} \right]$$

Next, it's time for a leading 1 in  $R3$ .

$$\left[ \begin{array}{cccc|c} 1 & -\frac{1}{3} & 0 & -\frac{1}{3} & -\frac{2}{3} \\ 0 & 1 & 6 & 1 & \frac{19}{2} \\ 0 & 0 & -24 & -5 & -35 \end{array} \right] \xrightarrow{\text{Replace } R3 \text{ with } -\frac{1}{24}R3} \left[ \begin{array}{cccc|c} 1 & -\frac{1}{3} & 0 & -\frac{1}{3} & -\frac{2}{3} \\ 0 & 1 & 6 & 1 & \frac{19}{2} \\ 0 & 0 & 1 & \frac{5}{24} & \frac{35}{24} \end{array} \right]$$

The matrix is now in row echelon form. To get the reduced row echelon form, we start with the last leading 1 we produced and work to get 0's above it.

$$\left[ \begin{array}{cccc|c} 1 & -\frac{1}{3} & 0 & -\frac{1}{3} & -\frac{2}{3} \\ 0 & 1 & 6 & 1 & \frac{19}{2} \\ 0 & 0 & 1 & \frac{5}{24} & \frac{35}{24} \end{array} \right] \xrightarrow{\text{Replace } R2 \text{ with } -6R3 + R2} \left[ \begin{array}{cccc|c} 1 & -\frac{1}{3} & 0 & -\frac{1}{3} & -\frac{2}{3} \\ 0 & 1 & 0 & -\frac{1}{4} & \frac{3}{4} \\ 0 & 0 & 1 & \frac{5}{24} & \frac{35}{24} \end{array} \right]$$

Lastly, we get a 0 above the leading 1 of  $R2$ .

$$\left[ \begin{array}{cccc|c} 1 & -\frac{1}{3} & 0 & -\frac{1}{3} & -\frac{2}{3} \\ 0 & 1 & 0 & -\frac{1}{4} & \frac{3}{4} \\ 0 & 0 & 1 & \frac{5}{24} & \frac{35}{24} \end{array} \right] \xrightarrow{\text{Replace } R1 \text{ with } \frac{1}{3}R2 + R1} \left[ \begin{array}{cccc|c} 1 & 0 & 0 & -\frac{5}{12} & -\frac{5}{12} \\ 0 & 1 & 0 & -\frac{1}{4} & \frac{3}{4} \\ 0 & 0 & 1 & \frac{5}{24} & \frac{35}{24} \end{array} \right]$$

At last, we decode to get

$$\left[ \begin{array}{cccc|c} 1 & 0 & 0 & -\frac{5}{12} & -\frac{5}{12} \\ 0 & 1 & 0 & -\frac{1}{4} & \frac{3}{4} \\ 0 & 0 & 1 & \frac{5}{24} & \frac{35}{24} \end{array} \right] \xrightarrow{\text{Decode from the matrix}} \begin{cases} x_1 - \frac{5}{12}x_4 = -\frac{5}{12} \\ x_2 - \frac{1}{4}x_4 = \frac{3}{4} \\ x_3 + \frac{5}{24}x_4 = \frac{35}{24} \end{cases}$$

We see  $x_4$  is free and assign it the parameter  $t$ . We obtain  $x_3 = -\frac{5}{24}t + \frac{35}{24}$ ,  $x_2 = \frac{1}{4}t + \frac{3}{4}$ , and  $x_1 = \frac{5}{12}t - \frac{5}{12}$ . Our solution is  $\left\{ \left( \frac{5}{12}t - \frac{5}{12}, \frac{1}{4}t + \frac{3}{4}, -\frac{5}{24}t + \frac{35}{24}, t \right) : -\infty < t < \infty \right\}$  which we leave to the reader to check.  $\square$

Like all good algorithms, putting a matrix in row echelon or reduced row echelon form can easily be programmed into a calculator, and, doubtless, your graphing calculator has such a feature. We make use of this feature in next example to avoid tedious arithmetic.

EXAMPLE 1.2.3. Find the quadratic function passing through the points  $(-1, 3)$ ,  $(2, 4)$ ,  $(5, -2)$ .

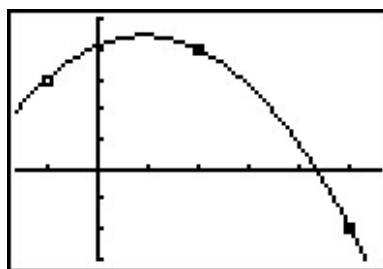
**Solution.** According to Definition ??, a quadratic function has the form  $f(x) = ax^2 + bx + c$  where  $a \neq 0$ . Our goal is to find  $a$ ,  $b$  and  $c$  so that the three given points are on the graph of  $f$ .

Since  $(-1, 3)$  is on the graph of  $f$ , we now  $f(-1) = 3$ . This gives  $a(-1)^2 + b(-1) + c = 3$ , or  $a - b + c = 3$ . This last equation is a linear equation with the variables  $a$ ,  $b$  and  $c$ . Similarly, if the point  $(2, 4)$  is on the graph of  $f$ , then  $f(2) = 4$ , so  $4a + 2b + c = 4$ . Lastly, the point  $(5, -2)$  is on the graph of  $f$  gives us  $25a + 5b + c = -2$ . Putting these together, we obtain a system of three linear equations, which we encode into the matrix below.

$$\begin{cases} a - b + c = 3 \\ 4a + 2b + c = 4 \\ 25a + 5b + c = -2 \end{cases} \xrightarrow{\text{Encode into the matrix}} \left[ \begin{array}{ccc|c} 1 & -1 & 1 & 3 \\ 4 & 2 & 1 & 4 \\ 25 & 5 & 1 & -2 \end{array} \right]$$

Using a calculator,<sup>4</sup> we find  $a = -\frac{7}{18}$ ,  $b = \frac{13}{18}$  and  $c = \frac{37}{9}$ . Hence, the one and only quadratic which fits the bill is  $f(x) = -\frac{7}{18}x^2 + \frac{13}{18}x + \frac{37}{9}$ . To verify this analytically, we see that  $f(-1) = 3$ ,  $f(2) = 4$ , and  $f(5) = -2$ . We can use the calculator to check our solution as well by plotting the three data points and the function  $f$ .

```
rref([A])>Frac
[[1 0 0 -7/18]
 [0 1 0 13/18]
 [0 0 1 37/9 1]]
```



The graph of  $f(x) = -\frac{7}{18}x^2 + \frac{13}{18}x + \frac{37}{9}$  with the points  $(-1, 3)$ ,  $(2, 4)$  and  $(5, -2)$

□

<sup>4</sup>We've tortured you enough already with fractions in this exposition!

## 1.2.1 Exercises

Exercise ideas: piggyback on Example 1.2.3: can you find a function...  $ax^3 + bx + c$ ,  $ax^{-1} + bx^2 + c$

Follow up in inverse determinant section: What about  $ax^m + bx^n + cx^p$  for these three points?

In Exercises 1 - 6, state whether the given matrix is in reduced row echelon form, row echelon form only or in neither of those forms.

$$1. \left[ \begin{array}{cc|c} 1 & 0 & 3 \\ 0 & 1 & 3 \end{array} \right]$$

$$2. \left[ \begin{array}{ccc|c} 3 & -1 & 1 & 3 \\ 2 & -4 & 3 & 16 \\ 1 & -1 & 1 & 5 \end{array} \right]$$

$$3. \left[ \begin{array}{ccc|c} 1 & 1 & 4 & 3 \\ 0 & 1 & 3 & 6 \\ 0 & 0 & 0 & 1 \end{array} \right]$$

$$4. \left[ \begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right]$$

$$5. \left[ \begin{array}{cccc|c} 1 & 0 & 4 & 3 & 0 \\ 0 & 1 & 3 & 6 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right]$$

$$6. \left[ \begin{array}{ccc|c} 1 & 1 & 4 & 3 \\ 0 & 1 & 3 & 6 \end{array} \right]$$

In Exercises 7 - 12, the following matrices are in reduced row echelon form. Determine the solution of the corresponding system of linear equations or state that the system is inconsistent.

$$7. \left[ \begin{array}{cc|c} 1 & 0 & -2 \\ 0 & 1 & 7 \end{array} \right]$$

$$8. \left[ \begin{array}{ccc|c} 1 & 0 & 0 & -3 \\ 0 & 1 & 0 & 20 \\ 0 & 0 & 1 & 19 \end{array} \right]$$

$$9. \left[ \begin{array}{cccc|c} 1 & 0 & 0 & 3 & 4 \\ 0 & 1 & 0 & 6 & -6 \\ 0 & 0 & 1 & 0 & 2 \end{array} \right]$$

$$10. \left[ \begin{array}{cccc|c} 1 & 0 & 0 & 3 & 0 \\ 0 & 1 & 2 & 6 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array} \right]$$

$$11. \left[ \begin{array}{cccc|c} 1 & 0 & -8 & 1 & 7 \\ 0 & 1 & 4 & -3 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right]$$

$$12. \left[ \begin{array}{ccc|c} 1 & 0 & 9 & -3 \\ 0 & 1 & -4 & 20 \\ 0 & 0 & 0 & 0 \end{array} \right]$$

In Exercises 13 - 26, solve the following systems of linear equations using the techniques discussed in this section. Compare and contrast these techniques with those you used to solve the systems in the Exercises in Section 1.1.

$$13. \begin{cases} -5x + y = 17 \\ x + y = 5 \end{cases}$$

$$14. \begin{cases} x + y + z = 3 \\ 2x - y + z = 0 \\ -3x + 5y + 7z = 7 \end{cases}$$

$$15. \begin{cases} 4x - y + z = 5 \\ 2y + 6z = 30 \\ x + z = 5 \end{cases}$$

$$16. \begin{cases} x - 2y + 3z = 7 \\ -3x + y + 2z = -5 \\ 2x + 2y + z = 3 \end{cases}$$

$$17. \begin{cases} 3x - 2y + z = -5 \\ x + 3y - z = 12 \\ x + y + 2z = 0 \end{cases}$$

$$18. \begin{cases} 2x - y + z = -1 \\ 4x + 3y + 5z = 1 \\ 5y + 3z = 4 \end{cases}$$

$$19. \begin{cases} x - y + z = -4 \\ -3x + 2y + 4z = -5 \\ x - 5y + 2z = -18 \end{cases}$$

$$20. \begin{cases} 2x - 4y + z = -7 \\ x - 2y + 2z = -2 \\ -x + 4y - 2z = 3 \end{cases}$$

$$21. \begin{cases} 2x - y + z = 1 \\ 2x + 2y - z = 1 \\ 3x + 6y + 4z = 9 \end{cases}$$

$$22. \begin{cases} x - 3y - 4z = 3 \\ 3x + 4y - z = 13 \\ 2x - 19y - 19z = 2 \end{cases}$$

$$23. \begin{cases} x + y + z = 4 \\ 2x - 4y - z = -1 \\ x - y = 2 \end{cases}$$

$$24. \begin{cases} x - y + z = 8 \\ 3x + 3y - 9z = -6 \\ 7x - 2y + 5z = 39 \end{cases}$$

$$25. \begin{cases} 2x - 3y + z = -1 \\ 4x - 4y + 4z = -13 \\ 6x - 5y + 7z = -25 \end{cases}$$

$$26. \begin{cases} x_1 - x_3 = -2 \\ 2x_2 - x_4 = 0 \\ x_1 - 2x_2 + x_3 = 0 \\ -x_3 + x_4 = 1 \end{cases}$$

27. It's time for another meal at our local buffet. This time, 22 diners (5 of whom were children) feasted for \$162.25, before taxes. If the kids buffet is \$4.50, the basic buffet is \$7.50, and the deluxe buffet (with crab legs) is \$9.25, find out how many diners chose the deluxe buffet.
28. Carl wants to make a party mix consisting of almonds (which cost \$7 per pound), cashews (which cost \$5 per pound), and peanuts (which cost \$2 per pound.) If he wants to make a 10 pound mix with a budget of \$35, what are the possible combinations almonds, cashews, and peanuts? (You may find it helpful to review Example 1.1.2 in Section 1.1.)
29. Find the quadratic function passing through the points  $(-2, 1)$ ,  $(1, 4)$ ,  $(3, -2)$
30. At 9 PM, the temperature was  $60^\circ\text{F}$ ; at midnight, the temperature was  $50^\circ\text{F}$ ; and at 6 AM, the temperature was  $70^\circ\text{F}$ . Use the technique in Example 1.2.3 to fit a quadratic function to these data with the temperature,  $T$ , measured in degrees Fahrenheit, as the dependent variable, and the number of hours after 9 PM,  $t$ , measured in hours, as the independent variable. What was the coldest temperature of the night? When did it occur?
31. The price for admission into the Stitz-Zeager Sasquatch Museum and Research Station is \$15 for adults and \$8 for kids 13 years old and younger. When the Zahlenreich family visits the museum their bill is \$38 and when the Nullsatz family visits their bill is \$39. One day both families went together and took an adult babysitter along to watch the kids and the total admission charge was \$92. Later that summer, the adults from both families went without the kids and the bill was \$45. Is that enough information to determine how many adults and children are in each family? If not, state whether the resulting system is inconsistent or consistent dependent. In the latter case, give at least two plausible solutions.
32. Use the technique in Example 1.2.3 to find the line between the points  $(-3, 4)$  and  $(6, 1)$ . How does your answer compare to the slope-intercept form of the line in Equation ???

33. With the help of your classmates, find at least two different row echelon forms for the matrix

$$\left[ \begin{array}{cc|c} 1 & 2 & 3 \\ 4 & 12 & 8 \end{array} \right]$$



**1.2.2 Answers**

1. Reduced row echelon form
2. Neither
3. Row echelon form only
4. Reduced row echelon form
5. Reduced row echelon form
6. Row echelon form only
7.  $(-2, 7)$
8.  $(-3, 20, 19)$
9.  $(-3t + 4, -6t - 6, 2, t)$   
for all real numbers  $t$
10. Inconsistent
11.  $(8s - t + 7, -4s + 3t + 2, s, t)$   
for all real numbers  $s$  and  $t$
12.  $(-9t - 3, 4t + 20, t)$   
for all real numbers  $t$
13.  $(-2, 7)$
14.  $(1, 2, 0)$
15.  $(-t + 5, -3t + 15, t)$   
for all real numbers  $t$
16.  $(2, -1, 1)$
17.  $(1, 3, -2)$
18. Inconsistent
19.  $(1, 3, -2)$
20.  $(-3, \frac{1}{2}, 1)$
21.  $(\frac{1}{3}, \frac{2}{3}, 1)$
22.  $(\frac{19}{13}t + \frac{51}{13}, -\frac{11}{13}t + \frac{4}{13}, t)$   
for all real numbers  $t$
23. Inconsistent
24.  $(4, -3, 1)$
25.  $(-2t - \frac{35}{4}, -t - \frac{11}{2}, t)$   
for all real numbers  $t$
26.  $(1, 2, 3, 4)$
27. This time, 7 diners chose the deluxe buffet.
28. If  $t$  represents the amount (in pounds) of peanuts, then we need  $1.5t - 7.5$  pounds of almonds and  $17.5 - 2.5t$  pounds of cashews. Since we can't have a negative amount of nuts,  $5 \leq t \leq 7$ .
29.  $f(x) = -\frac{4}{5}x^2 + \frac{1}{5}x + \frac{23}{5}$
30.  $T(t) = \frac{20}{27}t^2 - \frac{50}{9}t + 60$ . Lowest temperature of the evening  $\frac{595}{12} \approx 49.58^\circ\text{F}$  at 12:45 AM.

31. Let  $x_1$  and  $x_2$  be the numbers of adults and children, respectively, in the Zahlenreich family and let  $x_3$  and  $x_4$  be the numbers of adults and children, respectively, in the Nullsatz family. The system of equations determined by the given information is

$$\left\{ \begin{array}{rcl} 15x_1 + 8x_2 & = & 38 \\ 15x_3 + 8x_4 & = & 39 \\ 15x_1 + 8x_2 + 15x_3 + 8x_4 & = & 77 \\ 15x_1 + 15x_3 & = & 45 \end{array} \right.$$

We subtracted the cost of the babysitter in E3 so the constant is 77, not 92. This system is consistent dependent and its solution is  $(\frac{8}{15}t + \frac{2}{5}, -t + 4, -\frac{8}{15}t + \frac{13}{5}, t)$ . Our variables represent numbers of adults and children so they must be whole numbers. Running through the values  $t = 0, 1, 2, 3, 4$  yields only one solution where all four variables are whole numbers;  $t = 3$  gives us  $(2, 1, 1, 3)$ . Thus there are 2 adults and 1 child in the Zahlenreichs and 1 adult and 3 kids in the Nullsatzs.

## 1.3 Matrix Arithmetic

In Section 1.2, we used a special class of matrices, the augmented matrices, to assist us in solving systems of linear equations. In this section, we study matrices as mathematical objects of their own accord, temporarily divorced from systems of linear equations. To do so conveniently requires some more notation. When we write  $A = [a_{ij}]_{m \times n}$ , we mean  $A$  is an  $m$  by  $n$  matrix<sup>1</sup> and  $a_{ij}$  is the entry found in the  $i$ th row and  $j$ th column. Schematically, we have

$$A = \begin{matrix} & \begin{matrix} \xrightarrow{j \text{ counts columns} \\ \text{from left to right}} \end{matrix} & \\ \begin{matrix} \downarrow i \text{ counts rows} \\ \text{from top to bottom} \end{matrix} & \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} & \end{matrix}$$

With this new notation we can define what it means for two matrices to be equal.

**DEFINITION 1.5. Matrix Equality:** Two matrices are said to be **equal** if they are the same size and their corresponding entries are equal. More specifically, if  $A = [a_{ij}]_{m \times n}$  and  $B = [b_{ij}]_{p \times r}$ , we write  $A = B$  provided

1.  $m = p$  and  $n = r$
2.  $a_{ij} = b_{ij}$  for all  $1 \leq i \leq m$  and all  $1 \leq j \leq n$ .

Essentially, two matrices are equal if they are the same size and they have the same numbers in the same spots. For example, the two  $2 \times 3$  matrices below are, despite appearances, equal.

$$\begin{bmatrix} 0 & -2 & 9 \\ 25 & 117 & -3 \end{bmatrix} = \begin{bmatrix} \ln(1) & \sqrt[3]{-8} & e^{2\ln(3)} \\ 125^{2/3} & 3^2 \cdot 13 & \log(0.001) \end{bmatrix}$$

Now that we have an agreed upon understanding of what it means for two matrices to equal each other, we may begin defining arithmetic operations on matrices. Our first operation is addition.

**DEFINITION 1.6. Matrix Addition:** Given two matrices of the same size, the matrix obtained by adding the corresponding entries of the two matrices is called the **sum** of the two matrices. More specifically, if  $A = [a_{ij}]_{m \times n}$  and  $B = [b_{ij}]_{m \times n}$ , we define

$$A + B = [a_{ij}]_{m \times n} + [b_{ij}]_{m \times n} = [a_{ij} + b_{ij}]_{m \times n}$$

As an example, consider the sum below.

$$\begin{bmatrix} 2 & 3 \\ 4 & -1 \\ 0 & -7 \end{bmatrix} + \begin{bmatrix} -1 & 4 \\ -5 & -3 \\ 8 & 1 \end{bmatrix} = \begin{bmatrix} 2 + (-1) & 3 + 4 \\ 4 + (-5) & (-1) + (-3) \\ 0 + 8 & (-7) + 1 \end{bmatrix} = \begin{bmatrix} 1 & 7 \\ -1 & -4 \\ 8 & -6 \end{bmatrix}$$

<sup>1</sup>Recall that means  $A$  has  $m$  rows and  $n$  columns.

It is worth the reader's time to think what would have happened had we reversed the order of the summands above. As we would expect, we arrive at the same answer. In general,  $A + B = B + A$  for matrices  $A$  and  $B$ , provided they are the same size so that the sum is defined in the first place. This is the *commutative property* of matrix addition. To see why this is true in general, we appeal to the definition of matrix addition. Given  $A = [a_{ij}]_{m \times n}$  and  $B = [b_{ij}]_{m \times n}$ ,

$$A + B = [a_{ij}]_{m \times n} + [b_{ij}]_{m \times n} = [a_{ij} + b_{ij}]_{m \times n} = [b_{ij} + a_{ij}]_{m \times n} = [b_{ij}]_{m \times n} + [a_{ij}]_{m \times n} = B + A$$

where the second equality is the definition of  $A + B$ , the third equality holds by the commutative law of real number addition, and the fourth equality is the definition of  $B + A$ . In other words, matrix addition is commutative because real number addition is.

A similar argument shows the *associative property* of matrix addition also holds, inherited in turn from the associative law of real number addition. Specifically, for matrices  $A$ ,  $B$ , and  $C$  of the same size,  $(A + B) + C = A + (B + C)$ . In other words, when adding more than two matrices, it doesn't matter how they are grouped. This means that we can write  $A + B + C$  without parentheses and there is no ambiguity as to what this means.<sup>2</sup> These properties and more are summarized in the following theorem.

**THEOREM 1.3. Properties of Matrix Addition**

- **Commutative Property:** For all  $m \times n$  matrices,  $A + B = B + A$
- **Associative Property:** For all  $m \times n$  matrices,  $(A + B) + C = A + (B + C)$
- **Additive Identity:** If  $0_{m \times n}$  is the  $m \times n$  matrix whose entries are all 0, for all  $m \times n$  matrices  $A$

$$A + 0_{m \times n} = 0_{m \times n} + A = A$$

That is, the additive identity for a matrix is the matrix of the additive identity for each of its entries.

- **Additive Inverse:** For every given  $m \times n$  matrix  $A = [a_{ij}]_{m \times n}$ , the matrix  $B = [-a_{ij}]_{m \times n}$  satisfies

$$A + B = B + A = 0_{m \times n}$$

That is, the additive inverse of a matrix is the matrix of the additive inverses of each of its entries.

The identity property is easily verified by resorting to the definition of matrix addition; just as the number 0 is the additive identity for real numbers, the matrix comprised of all 0's does the same job for matrices. To establish the inverse property, we note that per the definition of matrix addition,

$$A + B = [a_{ij}]_{m \times n} + [-a_{ij}]_{m \times n} = [a_{ij} - a_{ij}]_{m \times n} = [0]_{m \times n} = 0_{m \times n}.$$

The fact that  $B + A = 0_{m \times n}$  as well comes from the commutative property of matrix addition.

More about the additive inverse is true. If a matrix  $C = [c_{ij}]_{m \times n}$  satisfies  $A + C = 0_{m \times n}$ , then once again by the definition of matrix addition, we must have  $a_{ij} + c_{ij} = 0$ , or  $c_{ij} = -a_{ij}$  for all  $i$  and  $j$ . This shows the

<sup>2</sup>We have seen this idea before in Sections ?? and ??.

matrix  $C$  must be the matrix  $B$  as described in Theorem 1.3 which shows the additive inverse of a matrix is *unique*. In general, we denote the additive inverse of a matrix  $A$  using the (suggestive) symbol  $-A$ .

With the concept of additive inverse well in hand, we may now discuss what is meant by subtracting matrices. You may remember from arithmetic that  $a - b = a + (-b)$ ; that is, subtraction is defined as ‘adding the opposite (inverse).’ We extend this concept to matrices. For two matrices  $A$  and  $B$  of the same size, we define  $A - B = A + (-B)$ . At the level of entries, this amounts to

$$A - B = A + (-B) = [a_{ij}]_{m \times n} + [-b_{ij}]_{m \times n} = [a_{ij} + (-b_{ij})]_{m \times n} = [a_{ij} - b_{ij}]_{m \times n}$$

Thus to subtract two matrices of equal size, we subtract their corresponding entries. Surprised?

Our next task is to define what it means to multiply a matrix by a real number. Thinking back to arithmetic, you may recall that multiplication, at least by a natural number, can be thought of as ‘rapid addition.’ For example,  $2 + 2 + 2 = 3 \cdot 2$ . We know from algebra<sup>3</sup> that  $3x = x + x + x$ , so it seems natural that given a matrix  $A$ , we define  $3A = A + A + A$ . If  $A = [a_{ij}]_{m \times n}$ , we have

$$3A = A + A + A = [a_{ij}]_{m \times n} + [a_{ij}]_{m \times n} + [a_{ij}]_{m \times n} = [a_{ij} + a_{ij} + a_{ij}]_{m \times n} = [3a_{ij}]_{m \times n}$$

In other words, multiplying the *matrix* in this fashion by 3 is the same as multiplying *each entry* by 3. This leads us to the following definition.

**DEFINITION 1.7. Scalar<sup>a</sup> Multiplication:** We define the product of a real number and a matrix to be the matrix obtained by multiplying each of its entries by said real number. More specifically, if  $k$  is a real number and  $A = [a_{ij}]_{m \times n}$ , we define

$$kA = k [a_{ij}]_{m \times n} = [ka_{ij}]_{m \times n}$$

<sup>a</sup>The word ‘scalar’ here refers to real numbers. ‘Scalar multiplication’ in this context means we are multiplying a matrix by a real number (a scalar). We will discuss this term momentarily.

The word ‘scalar’ means ‘scaling factor’ as we explain below. Every point  $P(x, y)$  in the plane can be represented by its position matrix,  $P$ :

$$(x, y) \leftrightarrow P = \begin{bmatrix} x \\ y \end{bmatrix}$$

Suppose we take the point  $(-2, 1)$  and multiply its position matrix by 3. We have

$$3P = 3 \begin{bmatrix} -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 3(-2) \\ 3(1) \end{bmatrix} = \begin{bmatrix} -6 \\ 3 \end{bmatrix}.$$

This new matrix corresponds to the point  $(-6, 3)$  which is the result of scaling both the horizontal and vertical directions by a factor of 3.

As did matrix addition, scalar multiplication inherits many properties from real number arithmetic. Below we summarize these properties.

<sup>3</sup>The Distributive Property, in particular.

**THEOREM 1.4. Properties of Scalar Multiplication**

- **Associative Property:** For every  $m \times n$  matrix  $A$  and scalars  $k$  and  $r$ ,  $(kr)A = k(rA)$ .
- **Identity Property:** For all  $m \times n$  matrices  $A$ ,  $1A = A$ .
- **Additive Inverse Property:** For all  $m \times n$  matrices  $A$ ,  $-A = (-1)A$ .
- **Distributive Property of Scalar Multiplication over Scalar Addition:**  
For every  $m \times n$  matrix  $A$  and scalars  $k$  and  $r$ ,

$$(k + r)A = kA + rA$$

- **Distributive Property of Scalar Multiplication over Matrix Addition:**  
For all  $m \times n$  matrices  $A$  and  $B$  scalars  $k$ ,

$$k(A + B) = kA + kB$$

- **Zero Product Property:** If  $A$  is an  $m \times n$  matrix and  $k$  is a scalar, then

$$kA = 0_{m \times n} \quad \text{if and only if} \quad k = 0 \quad \text{or} \quad A = 0_{m \times n}$$

As with the other results in this section, Theorem 1.4 can be proved using the definitions of scalar multiplication and matrix addition. For example, to prove that  $k(A + B) = kA + kB$  for a scalar  $k$  and  $m \times n$  matrices  $A$  and  $B$ , we start by adding  $A$  and  $B$ , then multiplying by  $k$  and seeing how that compares with the sum of  $kA$  and  $kB$ .

$$k(A + B) = k \left( [a_{ij}]_{m \times n} + [b_{ij}]_{m \times n} \right) = k [a_{ij} + b_{ij}]_{m \times n} = [k(a_{ij} + b_{ij})]_{m \times n} = [ka_{ij} + kb_{ij}]_{m \times n}$$

As for  $kA + kB$ , we have

$$kA + kB = k [a_{ij}]_{m \times n} + k [b_{ij}]_{m \times n} = [ka_{ij}]_{m \times n} + [kb_{ij}]_{m \times n} = [ka_{ij} + kb_{ij}]_{m \times n} \quad \checkmark$$

which establishes the property. The remaining proofs are similar and are left to the reader.

The properties in Theorems 1.3 and 1.4 establish an algebraic system that lets us treat matrices and scalars more or less as we would real numbers and variables. In the following example, we challenge the reader to justify each and every step of the calculations using either properties of matrix arithmetic.

**EXAMPLE 1.3.1.** Solve for the matrix  $A$ :

$$3A - \left( \begin{bmatrix} 2 & -1 \\ 3 & 5 \end{bmatrix} + 5A \right) = \begin{bmatrix} -4 & 2 \\ 6 & -2 \end{bmatrix} + \frac{1}{3} \begin{bmatrix} 9 & 12 \\ -3 & 39 \end{bmatrix}$$

**Solution.**

$$\begin{aligned}
3A - \left( \begin{bmatrix} 2 & -1 \\ 3 & 5 \end{bmatrix} + 5A \right) &= \begin{bmatrix} -4 & 2 \\ 6 & -2 \end{bmatrix} + \frac{1}{3} \begin{bmatrix} 9 & 12 \\ -3 & 39 \end{bmatrix} \\
3A + \left\{ - \left( \begin{bmatrix} 2 & -1 \\ 3 & 5 \end{bmatrix} + 5A \right) \right\} &= \begin{bmatrix} -4 & 2 \\ 6 & -2 \end{bmatrix} + \begin{bmatrix} \left(\frac{1}{3}\right)(9) & \left(\frac{1}{3}\right)(12) \\ \left(\frac{1}{3}\right)(-3) & \left(\frac{1}{3}\right)(39) \end{bmatrix} \\
3A + (-1) \left( \begin{bmatrix} 2 & -1 \\ 3 & 5 \end{bmatrix} + 5A \right) &= \begin{bmatrix} -4 & 2 \\ 6 & -2 \end{bmatrix} + \begin{bmatrix} 3 & 4 \\ -1 & 13 \end{bmatrix} \\
3A + \left\{ (-1) \begin{bmatrix} 2 & -1 \\ 3 & 5 \end{bmatrix} + (-1)(5A) \right\} &= \begin{bmatrix} -1 & 6 \\ 5 & 11 \end{bmatrix} \\
3A + (-1) \begin{bmatrix} 2 & -1 \\ 3 & 5 \end{bmatrix} + (-1)(5A) &= \begin{bmatrix} -1 & 6 \\ 5 & 11 \end{bmatrix} \\
3A + \begin{bmatrix} (-1)(2) & (-1)(-1) \\ (-1)(3) & (-1)(5) \end{bmatrix} + ((-1)(5))A &= \begin{bmatrix} -1 & 6 \\ 5 & 11 \end{bmatrix} \\
3A + \begin{bmatrix} -2 & 1 \\ -3 & -5 \end{bmatrix} + (-5)A &= \begin{bmatrix} -1 & 6 \\ 5 & 11 \end{bmatrix} \\
3A + (-5)A + \begin{bmatrix} -2 & 1 \\ -3 & -5 \end{bmatrix} &= \begin{bmatrix} -1 & 6 \\ 5 & 11 \end{bmatrix} \\
(3 + (-5))A + \begin{bmatrix} -2 & 1 \\ -3 & -5 \end{bmatrix} + \left( - \begin{bmatrix} -2 & 1 \\ -3 & -5 \end{bmatrix} \right) &= \begin{bmatrix} -1 & 6 \\ 5 & 11 \end{bmatrix} + \left( - \begin{bmatrix} -2 & 1 \\ -3 & -5 \end{bmatrix} \right) \\
(-2)A + 0_{2 \times 2} &= \begin{bmatrix} -1 & 6 \\ 5 & 11 \end{bmatrix} - \begin{bmatrix} -2 & 1 \\ -3 & -5 \end{bmatrix} \\
(-2)A &= \begin{bmatrix} -1 - (-2) & 6 - 1 \\ 5 - (-3) & 11 - (-5) \end{bmatrix} \\
(-2)A &= \begin{bmatrix} 1 & 5 \\ 8 & 16 \end{bmatrix} \\
\left(-\frac{1}{2}\right)((-2)A) &= -\frac{1}{2} \begin{bmatrix} 1 & 5 \\ 8 & 16 \end{bmatrix} \\
\left(\left(-\frac{1}{2}\right)(-2)\right)A &= \begin{bmatrix} \left(-\frac{1}{2}\right)(1) & \left(-\frac{1}{2}\right)(5) \\ \left(-\frac{1}{2}\right)(8) & \left(-\frac{1}{2}\right)(16) \end{bmatrix} \\
1A &= \begin{bmatrix} -\frac{1}{2} & -\frac{5}{2} \\ -4 & -\frac{16}{2} \end{bmatrix} \\
A &= \begin{bmatrix} -\frac{1}{2} & -\frac{5}{2} \\ -4 & -8 \end{bmatrix}
\end{aligned}$$

The reader is encouraged to check our answer in the original equation.

□

While the solution to the previous example is written in excruciating detail, in practice many of the steps above are omitted. The reader is encouraged to solve the equation in Example 1.3.1 as they would any other linear equation, for example:  $3a - (2 + 5a) = -4 + \frac{1}{3}(9)$ .

We now turn our attention to *matrix multiplication* - that is, multiplying a matrix by another matrix. Based on the 'no surprises' trend so far in the section, you may expect that in order to multiply two matrices, they must be of the same size and you find the product by multiplying the corresponding entries. While this kind of product is used in other areas of mathematics,<sup>4</sup> we define matrix multiplication to serve us in solving systems of linear equations.

To that end, we begin by defining the product of a row and a column. We motivate the general definition with an example. Consider the two matrices  $A$  and  $B$  below.

$$A = \begin{bmatrix} 2 & 0 & -1 \\ -10 & 3 & 5 \end{bmatrix} \quad B = \begin{bmatrix} 3 & 1 & 2 & -8 \\ 4 & 8 & -5 & 9 \\ 5 & 0 & -2 & -12 \end{bmatrix}$$

Let  $R1$  denote the first row of  $A$  and  $C1$  denote the first column of  $B$ . To find the 'product' of  $R1$  with  $C1$ , denoted  $R1 \cdot C1$ , we first find the product of the first entry in  $R1$  and the first entry in  $C1$ . Next, we add to that the product of the second entry in  $R1$  and the second entry in  $C1$ , and so on until we reach the last entry in  $R1$  and the last entry in  $C1$ .

Using entry notation,  $R1 \cdot C1 = a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} = (2)(3) + (0)(4) + (-1)(5) = 6 + 0 + (-5) = 1$ . We can visualize this schematically as follows

$$\begin{array}{c} \begin{bmatrix} 2 & 0 & -1 \\ -10 & 3 & 5 \end{bmatrix} \quad \begin{bmatrix} 3 & 1 & 2 & -8 \\ 4 & 8 & -5 & 9 \\ 5 & 0 & -2 & -12 \end{bmatrix} \\ \\ \underbrace{\begin{array}{c} \xrightarrow{\quad} \boxed{3} \\ \boxed{2} \quad 0 \quad -1 \quad \downarrow \\ 4 \\ 5 \end{array}}_{\substack{a_{11}b_{11} \\ (2)(3)}} \quad + \quad \underbrace{\begin{array}{c} \xrightarrow{\quad} \boxed{4} \\ 2 \quad \boxed{0} \quad -1 \quad \downarrow \\ 5 \end{array}}_{\substack{a_{12}b_{21} \\ (0)(4)}} \quad + \quad \underbrace{\begin{array}{c} \xrightarrow{\quad} \boxed{5} \\ 2 \quad 0 \quad \boxed{-1} \quad \downarrow \\ 5 \end{array}}_{\substack{a_{13}b_{31} \\ (-1)(5)}} \end{array}$$

To find  $R2 \cdot C3$  where  $R2$  denotes the second row of  $A$  and  $C3$  denotes the third column of  $B$ , we proceed similarly. We start with finding the product of the first entry of  $R2$  with the first entry in  $C3$  then add to it the product of the second entry in  $R2$  with the second entry in  $C3$ , and so forth. Using entry notation, we have  $R2 \cdot C3 = a_{21}b_{13} + a_{22}b_{23} + a_{23}b_{33} = (-10)(2) + (3)(-5) + (5)(-2) = -45$ . Schematically,

$$\begin{bmatrix} 2 & 0 & -1 \\ -10 & 3 & 5 \end{bmatrix} \quad \begin{bmatrix} 3 & 1 & 2 & -8 \\ 4 & 8 & -5 & 9 \\ 5 & 0 & -2 & -12 \end{bmatrix}$$

<sup>4</sup>See this article on the [Hadamard Product](#).



$$\begin{array}{c}
 \begin{array}{ccc} \xrightarrow{\quad} & & \boxed{2} \\ \boxed{-10} & 3 & 5 \\ & & \downarrow \\ & & \begin{array}{c} -5 \\ -2 \end{array} \end{array} & + & \begin{array}{ccc} \xrightarrow{\quad} & & 2 \\ -10 & \boxed{3} & 5 \\ & & \downarrow \\ & & \begin{array}{c} -5 \\ -2 \end{array} \end{array} & + & \begin{array}{ccc} \xrightarrow{\quad} & & 2 \\ -10 & 3 & \boxed{5} \\ & & \downarrow \\ & & \begin{array}{c} -5 \\ -2 \end{array} \end{array} \\
 \hline
 a_{21}b_{13} = (-10)(2) = -20 & + & a_{22}b_{23} = (3)(-5) = -15 & + & a_{23}b_{33} = (5)(-2) = -10
 \end{array}$$

Generalizing this process, we have the following definition.

**DEFINITION 1.8. Product of a Row and a Column:** Suppose  $A = [a_{ij}]_{m \times n}$  and  $B = [b_{ij}]_{n \times r}$ . Let  $R_i$  denote the  $i$ th row of  $A$  and let  $C_j$  denote the  $j$ th column of  $B$ . The **product of  $R_i$  and  $C_j$** , denoted  $R_i \cdot C_j$  is the real number defined by

$$R_i \cdot C_j = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{in}b_{nj}$$

Note that in order to multiply a row by a column, the number of entries in the row must match the number of entries in the column. We are now in the position to define matrix multiplication.

**DEFINITION 1.9. Matrix Multiplication:** Suppose  $A = [a_{ij}]_{m \times n}$  and  $B = [b_{ij}]_{n \times r}$ . Let  $R_i$  denote the  $i$ th row of  $A$  and let  $C_j$  denote the  $j$ th column of  $B$ . The **product of  $A$  and  $B$** , denoted  $AB$ , is the matrix

$$AB = [R_i \cdot C_j]_{m \times r}$$

that is

$$AB = \begin{bmatrix} R_1 \cdot C_1 & R_1 \cdot C_2 & \dots & R_1 \cdot C_r \\ R_2 \cdot C_1 & R_2 \cdot C_2 & \dots & R_2 \cdot C_r \\ \vdots & \vdots & & \vdots \\ R_m \cdot C_1 & R_m \cdot C_2 & \dots & R_m \cdot C_r \end{bmatrix}$$

There are a number of subtleties in Definition 1.9 which warrant closer inspection. First and foremost, Definition 1.9 tells us that the  $ij$ -entry of a matrix product  $AB$  is the  $i$ th row of  $A$  times the  $j$ th column of  $B$ . In order for this to be defined, the number of entries in the rows of  $A$  must match the number of entries in the columns of  $B$ . This means that the number of columns of  $A$  must match<sup>5</sup> the number of rows of  $B$ . In other words, to multiply  $A$  times  $B$ , the second dimension of  $A$  must match the first dimension of  $B$ , which is why in Definition 1.9,  $A_{m \times n}$  is being multiplied by a matrix  $B_{n \times r}$ .

Furthermore, the product matrix  $AB$  has as many rows as  $A$  and as many columns of  $B$ . As a result, when multiplying a matrix  $A_{m \times n}$  by a matrix  $B_{n \times r}$ , the result is the matrix  $AB_{m \times r}$ .

Returning to our example matrices below, we see that  $A$  is a  $2 \times 3$  matrix and  $B$  is a  $3 \times 4$  matrix. This means that the product matrix  $AB$  is defined and will be a  $2 \times 4$  matrix.

$$A = \begin{bmatrix} 2 & 0 & -1 \\ -10 & 3 & 5 \end{bmatrix} \quad B = \begin{bmatrix} 3 & 1 & 2 & -8 \\ 4 & 8 & -5 & 9 \\ 5 & 0 & -2 & -12 \end{bmatrix}$$

<sup>5</sup>The reader is encouraged to think this through carefully.

Using  $R_i$  to denote the  $i$ th row of  $A$  and  $C_j$  to denote the  $j$ th column of  $B$ , we form  $AB$  per to Definition 1.9:

$$AB = \begin{bmatrix} R_1 \cdot C_1 & R_1 \cdot C_2 & R_1 \cdot C_3 & R_1 \cdot C_4 \\ R_2 \cdot C_1 & R_2 \cdot C_2 & R_2 \cdot C_3 & R_2 \cdot C_4 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 6 & -4 \\ 7 & 14 & -45 & 47 \end{bmatrix}$$

Note that the product  $BA$  is not defined, since  $B$  is a  $3 \times 4$  matrix while  $A$  is a  $2 \times 3$  matrix;  $B$  has more columns than  $A$  has rows, and so it is not possible to multiply a row of  $B$  by a column of  $A$ .

Even when the dimensions of  $A$  and  $B$  are compatible such that  $AB$  and  $BA$  are both defined, the product  $AB$  and  $BA$  aren't necessarily equal.<sup>6</sup> In other words,  $AB$  may not equal  $BA$  which means matrix multiplication is not, in general, commutative. That being said, several other real number properties are inherited by matrix multiplication, as illustrated in our next theorem.

**THEOREM 1.5. Properties of Matrix Multiplication** Let  $A$ ,  $B$  and  $C$  be matrices such that all of the matrix products below are defined and let  $k$  be a real number.

- **Associative Property of Matrix Multiplication:**  $(AB)C = A(BC)$
- **Associative Property with Scalar Multiplication:**  $k(AB) = (kA)B = A(kB)$
- **Identity Property:**

For a natural number  $k$ , the  $k \times k$  **identity matrix**, denoted  $I_k$ , is defined by  $I_k = [d_{ij}]_{k \times k}$  where

$$d_{ij} = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{otherwise} \end{cases}$$

For all  $m \times n$  matrices,  $I_m A = A I_n = A$ .

- **Distributive Property of Matrix Multiplication over Matrix Addition:**

$$A(B \pm C) = AB \pm AC \text{ and } (A \pm B)C = AC \pm BC$$

The one property in Theorem 1.5 which begs further investigation is, without doubt, the multiplicative identity. The entries in a matrix where  $i = j$  comprise what is called the *main diagonal* of the matrix. The identity matrix has 1's along its main diagonal and 0's everywhere else. A few examples of the matrix  $I_k$  mentioned in Theorem 1.5 are given below. The reader is encouraged to see how they match the definition of the identity matrix presented there.

$$\begin{array}{cccc} [1] & \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} & \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\ I_1 & I_2 & I_3 & I_4 \end{array}$$

The identity matrix is an example of what is called a *square matrix* as it has the same number of rows as columns. Note that to in order to verify that the identity matrix acts as a multiplicative identity, some care

<sup>6</sup>And may not even have the same dimensions. For example, if  $A$  is a  $2 \times 3$  matrix and  $B$  is a  $3 \times 2$  matrix, then  $AB$  is defined and is a  $2 \times 2$  matrix while  $BA$  is also defined... but is a  $3 \times 3$  matrix!

must be taken depending on the order of the multiplication. For example, take the matrix  $2 \times 3$  matrix  $A$ :

$$A = \begin{bmatrix} 2 & 0 & -1 \\ -10 & 3 & 5 \end{bmatrix}$$

In order for the product  $I_k A$  to be defined,  $k = 2$ ; similarly, for  $A I_k$  to be defined,  $k = 3$ . We leave it to the reader to show  $I_2 A = A$  and  $A I_3 = A$ . In other words,

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & -1 \\ -10 & 3 & 5 \end{bmatrix} = \begin{bmatrix} 2 & 0 & -1 \\ -10 & 3 & 5 \end{bmatrix}$$

and

$$\begin{bmatrix} 2 & 0 & -1 \\ -10 & 3 & 5 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 & -1 \\ -10 & 3 & 5 \end{bmatrix}$$

While the proofs of the properties in Theorem 1.5 are computational in nature, the notation becomes quite involved very quickly, so they are left to a course in Linear Algebra. The following example provides some practice with matrix multiplication and its properties. As usual, some valuable lessons are to be learned.

EXAMPLE 1.3.2.

1. Find  $AB$  for  $A = \begin{bmatrix} -23 & -1 & 17 \\ 46 & 2 & -34 \end{bmatrix}$  and  $B = \begin{bmatrix} -3 & 2 \\ 1 & 5 \\ -4 & 3 \end{bmatrix}$ .
2. Find  $C^2 - 5C + 10I_2$  for  $C = \begin{bmatrix} 1 & -2 \\ 3 & 4 \end{bmatrix}$ .
3. Suppose  $M$  is a  $4 \times 4$  matrix. Use Theorem 1.5 to expand  $(M - 2I_4)(M + 3I_4)$ .

**Solution.**

$$1. \text{ We have } AB = \begin{bmatrix} -23 & -1 & 17 \\ 46 & 2 & -34 \end{bmatrix} \begin{bmatrix} -3 & 2 \\ 1 & 5 \\ -4 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

2. Just as  $x^2$  means  $x$  times itself,  $C^2$  denotes the matrix  $C$  times itself. We get

$$\begin{aligned} C^2 - 5C + 10I_2 &= \begin{bmatrix} 1 & -2 \\ 3 & 4 \end{bmatrix}^2 - 5 \begin{bmatrix} 1 & -2 \\ 3 & 4 \end{bmatrix} + 10 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & -2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} -5 & 10 \\ -15 & -20 \end{bmatrix} + \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix} \\ &= \begin{bmatrix} -5 & -10 \\ 15 & 10 \end{bmatrix} + \begin{bmatrix} 5 & 10 \\ -15 & -10 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

3. We expand  $(M - 2I_4)(M + 3I_4)$  with the same pedantic zeal we showed in Example 1.3.1. The reader is encouraged to determine which property of matrix arithmetic justifies each step.

$$\begin{aligned}
 (M - 2I_4)(M + 3I_4) &= (M - 2I_4)M + (M - 2I_4)(3I_4) \\
 &= MM - (2I_4)M + M(3I_4) - (2I_4)(3I_4) \\
 &= M^2 - 2(I_4M) + 3(MI_4) - 2(I_4(3I_4)) \\
 &= M^2 - 2M + 3M - 2(3(I_4I_4)) \\
 &= M^2 + M - 6I_4
 \end{aligned}$$

□

Example 1.3.2 illustrates some interesting features of matrix multiplication. First note that in the first problem, neither  $A$  nor  $B$  is the zero matrix, yet the product  $AB$  is the zero matrix. Hence, the zero product property enjoyed by real numbers and scalar multiplication does not hold for matrix multiplication.

The second and third problems introduce us to polynomials involving matrices. The reader is encouraged to step back and compare our expansion of the matrix product  $(M - 2I_4)(M + 3I_4)$  in third problem with the product  $(x - 2)(x + 3)$  from real number algebra. The exercises explore this kind of parallel further.

As we mentioned earlier, a point  $P(x, y)$  in the  $xy$ -plane can be represented as a  $2 \times 1$  position matrix. We now show that matrix multiplication can be used to rotate these points, and hence graphs of equations.

EXAMPLE 1.3.3. Let  $R = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}$ .

1. Plot  $P(2, -2)$ ,  $Q(4, 0)$ ,  $S(0, 3)$ , and  $T(-3, -3)$  in the plane as well as the points  $RP$ ,  $RQ$ ,  $RS$ , and  $RT$ . Plot the lines  $y = x$  and  $y = -x$  as guides. What does  $R$  appear to be doing to these points?
2. If a point  $P$  is on the hyperbola  $x^2 - y^2 = 4$ , show that the point  $RP$  is on the curve  $y = \frac{2}{x}$ .

**Solution.**

1. For  $P(2, -2)$ , the position matrix is  $P = \begin{bmatrix} 2 \\ -2 \end{bmatrix}$ , and

$$\begin{aligned}
 RP &= \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} 2 \\ -2 \end{bmatrix} \\
 &= \begin{bmatrix} 2\sqrt{2} \\ 0 \end{bmatrix}
 \end{aligned}$$

We have that  $R$  takes  $(2, -2)$  to  $(2\sqrt{2}, 0)$ . Similarly, we find  $(4, 0)$  is moved to  $(2\sqrt{2}, 2\sqrt{2})$ ,  $(0, 3)$  is moved to  $(-\frac{3\sqrt{2}}{2}, \frac{3\sqrt{2}}{2})$ , and  $(-3, -3)$  is moved to  $(0, -3\sqrt{2})$ . We plot these points below on the left along with the lines  $y = x$  and  $y = -x$ . We see that the matrix  $R$  is rotating these points counterclockwise by  $45^\circ$ .

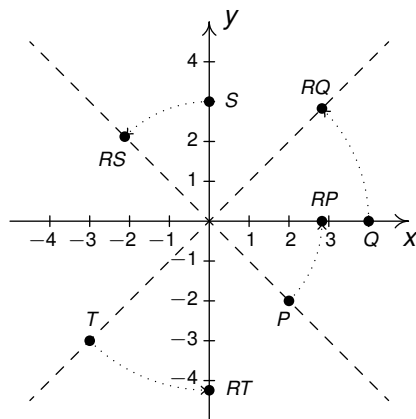
2. For a generic point  $P(x, y)$  on the hyperbola  $x^2 - y^2 = 4$ , we have

$$\begin{aligned} RP &= \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ &= \begin{bmatrix} \frac{\sqrt{2}}{2}x - \frac{\sqrt{2}}{2}y \\ \frac{\sqrt{2}}{2}x + \frac{\sqrt{2}}{2}y \end{bmatrix} \end{aligned}$$

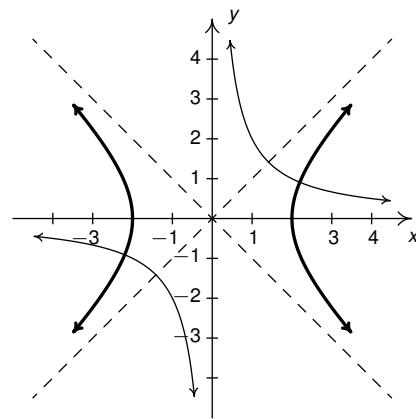
which means  $R$  takes  $(x, y)$  to  $\left(\frac{\sqrt{2}}{2}x - \frac{\sqrt{2}}{2}y, \frac{\sqrt{2}}{2}x + \frac{\sqrt{2}}{2}y\right)$ . To show that this point is on the curve  $y = \frac{2}{x}$ , we replace  $x$  with  $\frac{\sqrt{2}}{2}x - \frac{\sqrt{2}}{2}y$  and  $y$  with  $\frac{\sqrt{2}}{2}x + \frac{\sqrt{2}}{2}y$  and simplify.

$$\begin{aligned} y &= \frac{2}{x} \\ \frac{\sqrt{2}}{2}x + \frac{\sqrt{2}}{2}y &\stackrel{?}{=} \frac{2}{\frac{\sqrt{2}}{2}x - \frac{\sqrt{2}}{2}y} \\ \left(\frac{\sqrt{2}}{2}x - \frac{\sqrt{2}}{2}y\right)\left(\frac{\sqrt{2}}{2}x + \frac{\sqrt{2}}{2}y\right) &\stackrel{?}{=} \left(\frac{2}{\frac{\sqrt{2}}{2}x - \frac{\sqrt{2}}{2}y}\right)\left(\frac{\sqrt{2}}{2}x - \frac{\sqrt{2}}{2}y\right) \\ \left(\frac{\sqrt{2}}{2}x\right)^2 - \left(\frac{\sqrt{2}}{2}y\right)^2 &\stackrel{?}{=} 2 \\ \frac{x^2}{2} - \frac{y^2}{2} &\stackrel{?}{=} 2 \\ x^2 - y^2 &\stackrel{\checkmark}{=} 4 \end{aligned}$$

Since  $(x, y)$  is on the hyperbola  $x^2 - y^2 = 4$ , we know that this last equation is true. Since all of our steps are reversible, this last equation is equivalent to our original equation, showing the graph of  $y = \frac{2}{x}$  is none other than the hyperbola  $x^2 - y^2 = 4$  when rotated counterclockwise by  $45^\circ$ . Below on the right are the graphs of  $x^2 - y^2 = 4$  (thicker line) and  $y = \frac{2}{x}$  for comparison.



Plotting  $P$  and  $RP$ .



Graphing  $x^2 - y^2 = 4$  and  $y = \frac{2}{x}$ .

When we started this section, we mentioned that we would temporarily consider matrices as their own entities, but that the algebra developed here would ultimately allow us to solve systems of linear equations. To that end, consider the system

$$\begin{cases} 3x - y + z = 8 \\ x + 2y - z = 4 \\ 2x + 3y - 4z = 10 \end{cases}$$

In Section 1.2, we encoded this system into the augmented matrix

$$\left[ \begin{array}{ccc|c} 3 & -1 & 1 & 8 \\ 1 & 2 & -1 & 4 \\ 2 & 3 & -4 & 10 \end{array} \right]$$

Recall that the entries to the left of the vertical line come from the coefficients of the variables in the system, while those on the right comprise the associated constants. For that reason, we may form the *coefficient matrix*  $A$ , the *unknowns matrix*  $X$  and the *constant matrix*  $B$  as below

$$A = \begin{bmatrix} 3 & -1 & 1 \\ 1 & 2 & -1 \\ 2 & 3 & -4 \end{bmatrix} \quad X = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad B = \begin{bmatrix} 8 \\ 4 \\ 10 \end{bmatrix}$$

We now consider the matrix equation  $AX = B$ .

$$AX = B$$

$$\begin{bmatrix} 3 & -1 & 1 \\ 1 & 2 & -1 \\ 2 & 3 & -4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \\ 10 \end{bmatrix}$$

$$\begin{bmatrix} 3x - y + z \\ x + 2y - z \\ 2x + 3y - 4z \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \\ 10 \end{bmatrix}$$

We see that finding a solution  $(x, y, z)$  to the original system corresponds to finding a solution  $X$  for the matrix equation  $AX = B$ . If we think about solving the real number equation  $ax = b$ , we would simply ‘divide’ both sides by  $a$ . Is it possible to ‘divide’ both sides of the matrix equation  $AX = B$  by the matrix  $A$ ? This is the central topic of Section 1.4.

## 1.3.1 Exercises

For each pair of matrices  $A$  and  $B$  in Exercises 1 - 7, find the following, if defined

- $3A$
- $-B$
- $A^2$
- $A - 2B$
- $AB$
- $BA$

$$1. A = \begin{bmatrix} 2 & -3 \\ 1 & 4 \end{bmatrix}, B = \begin{bmatrix} 5 & -2 \\ 4 & 8 \end{bmatrix}$$

$$2. A = \begin{bmatrix} -1 & 5 \\ -3 & 6 \end{bmatrix}, B = \begin{bmatrix} 2 & 10 \\ -7 & 1 \end{bmatrix}$$

$$3. A = \begin{bmatrix} -1 & 3 \\ 5 & 2 \end{bmatrix}, B = \begin{bmatrix} 7 & 0 & 8 \\ -3 & 1 & 4 \end{bmatrix}$$

$$4. A = \begin{bmatrix} 2 & 4 \\ 6 & 8 \end{bmatrix}, B = \begin{bmatrix} -1 & 3 & -5 \\ 7 & -9 & 11 \end{bmatrix}$$

$$5. A = \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix}, B = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$$

$$6. A = \begin{bmatrix} 1 & -2 \\ -3 & 4 \\ 5 & -6 \end{bmatrix}, B = \begin{bmatrix} -5 & 1 & 8 \end{bmatrix}$$

$$7. A = \begin{bmatrix} 2 & -3 & 5 \\ 3 & 1 & -2 \\ -7 & 1 & -1 \end{bmatrix}, B = \begin{bmatrix} 1 & 2 & 1 \\ 17 & 33 & 19 \\ 10 & 19 & 11 \end{bmatrix}$$

In Exercises 8 - 21, use the matrices

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad B = \begin{bmatrix} 0 & -3 \\ -5 & 2 \end{bmatrix} \quad C = \begin{bmatrix} 10 & -\frac{11}{2} & 0 \\ \frac{3}{5} & 5 & 9 \end{bmatrix}$$

$$D = \begin{bmatrix} 7 & -13 \\ -\frac{4}{3} & 0 \\ 6 & 8 \end{bmatrix} \quad E = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & -9 \\ 0 & 0 & -5 \end{bmatrix}$$

to compute the following or state that the indicated operation is undefined.

- 8.  $7B - 4A$
- 9.  $AB$
- 10.  $BA$
- 11.  $E + D$
- 12.  $ED$
- 13.  $CD + 2I_2A$
- 14.  $A - 4I_2$
- 15.  $A^2 - B^2$
- 16.  $(A + B)(A - B)$
- 17.  $A^2 - 5A - 2I_2$
- 18.  $E^2 + 5E - 36I_3$
- 19.  $EDC$
- 20.  $CDE$
- 21.  $ABCEDI_2$
- 22. Let  $A = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}$   $E_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$   $E_2 = \begin{bmatrix} 5 & 0 \\ 0 & 1 \end{bmatrix}$   $E_3 = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}$

Compute  $E_1A$ ,  $E_2A$  and  $E_3A$ . What effect did each of the  $E_i$  matrices have on the rows of  $A$ ? Create  $E_4$  so that its effect on  $A$  is to multiply the bottom row by  $-6$ . How would you extend this idea to matrices with more than two rows?

In Exercises 23 - 29, consider the following scenario. In the small village of Pedimaxus in the country of Sasquatchia, all 150 residents get one of the two local newspapers. Market research has shown that in any given week, 90% of those who subscribe to the Pedimaxus Tribune want to keep getting it, but 10% want to switch to the Sasquatchia Picayune. Of those who receive the Picayune, 80% want to continue with it and 20% want switch to the Tribune. We can express this situation using matrices. Specifically, let  $X$  be the 'state matrix' given by

$$X = \begin{bmatrix} T \\ P \end{bmatrix}$$

where  $T$  is the number of people who get the Tribune and  $P$  is the number of people who get the Picayune in a given week. Let  $Q$  be the 'transition matrix' given by

$$Q = \begin{bmatrix} 0.90 & 0.20 \\ 0.10 & 0.80 \end{bmatrix}$$

such that  $QX$  will be the state matrix for the next week.

23. Let's assume that when Pedimaxus was founded, all 150 residents got the Tribune. (Let's call this Week 0.) This would mean

$$X = \begin{bmatrix} 150 \\ 0 \end{bmatrix}$$

Since 10% of that 150 want to switch to the Picayune, we should have that for Week 1, 135 people get the Tribune and 15 people get the Picayune. Show that  $QX$  in this situation is indeed

$$QX = \begin{bmatrix} 135 \\ 15 \end{bmatrix}$$

24. Assuming that the percentages stay the same, we can get to the subscription numbers for Week 2 by computing  $Q^2X$ . How many people get each paper in Week 2?
25. Explain why the transition matrix does what we want it to do.
26. If the conditions do not change from week to week, then  $Q$  remains the same and we have what's known as a **Stochastic Process**<sup>7</sup> because Week  $n$ 's numbers are found by computing  $Q^nX$ . Choose a few values of  $n$  and, with the help of your classmates and calculator, find out how many people get each paper for that week. You should start to see a pattern as  $n \rightarrow \infty$ .
27. If you didn't see the pattern, we'll help you out. Let

$$X_s = \begin{bmatrix} 100 \\ 50 \end{bmatrix}.$$

Show that  $QX_s = X_s$ . This is called the **steady state** because the number of people who get each paper didn't change for the next week. Show that  $Q^nX \rightarrow X_s$  as  $n \rightarrow \infty$ .

---

<sup>7</sup>More specifically, we have a Markov Chain, which is a special type of stochastic process.



28. Now let

$$S = \begin{bmatrix} \frac{2}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

Show that  $Q^n \rightarrow S$  as  $n \rightarrow \infty$ .

29. Show that  $SY = X_s$  for any matrix  $Y$  of the form

$$Y = \begin{bmatrix} y \\ 150 - y \end{bmatrix}$$

This means that no matter how the distribution starts in Pedimaxus, if  $Q$  is applied often enough, we always end up with 100 people getting the Tribune and 50 people getting the Picayune.

30. Let  $z = a + bi$  and  $w = c + di$  be arbitrary complex numbers. Associate  $z$  and  $w$  with the matrices

$$Z = \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \quad \text{and} \quad W = \begin{bmatrix} c & d \\ -d & c \end{bmatrix}$$

Show that complex number addition, subtraction and multiplication are mirrored by the associated *matrix* arithmetic. That is, show that  $Z + W$ ,  $Z - W$  and  $ZW$  produce matrices which can be associated with the complex numbers  $z + w$ ,  $z - w$  and  $zw$ , respectively.

31. Let

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & -3 \\ -5 & 2 \end{bmatrix}$$

Compare  $(A + B)^2$  to  $A^2 + 2AB + B^2$ . Discuss with your classmates what constraints must be placed on two arbitrary matrices  $A$  and  $B$  so that both  $(A + B)^2$  and  $A^2 + 2AB + B^2$  exist. When will  $(A + B)^2 = A^2 + 2AB + B^2$ ? In general, what is the correct formula for  $(A + B)^2$ ?

In Exercises 32 - 36, consider the following definitions. A square matrix is said to be an **upper triangular matrix** if all of its entries below the main diagonal are zero and it is said to be a **lower triangular matrix** if all of its entries above the main diagonal are zero. For example,

$$E = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & -9 \\ 0 & 0 & -5 \end{bmatrix}$$

from Exercises 8 - 21 above is an upper triangular matrix whereas

$$F = \begin{bmatrix} 1 & 0 \\ 3 & 0 \end{bmatrix}$$

is a lower triangular matrix. (Zeros are allowed on the main diagonal.) Discuss the following questions with your classmates.

32. Give an example of a matrix which is neither upper triangular nor lower triangular.
33. Is the product of two  $n \times n$  upper triangular matrices always upper triangular?
34. Is the product of two  $n \times n$  lower triangular matrices always lower triangular?
35. Given the matrix

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

write  $A$  as  $LU$  where  $L$  is a lower triangular matrix and  $U$  is an upper triangular matrix?

36. Are there any matrices which are simultaneously upper and lower triangular?

**1.3.2 Answers**

1. For  $A = \begin{bmatrix} 2 & -3 \\ 1 & 4 \end{bmatrix}$  and  $B = \begin{bmatrix} 5 & -2 \\ 4 & 8 \end{bmatrix}$

$$\bullet 3A = \begin{bmatrix} 6 & -9 \\ 3 & 12 \end{bmatrix}$$

$$\bullet -B = \begin{bmatrix} -5 & 2 \\ -4 & -8 \end{bmatrix}$$

$$\bullet A^2 = \begin{bmatrix} 1 & -18 \\ 6 & 13 \end{bmatrix}$$

$$\bullet A - 2B = \begin{bmatrix} -8 & 1 \\ -7 & -12 \end{bmatrix}$$

$$\bullet AB = \begin{bmatrix} -2 & -28 \\ 21 & 30 \end{bmatrix}$$

$$\bullet BA = \begin{bmatrix} 8 & -23 \\ 16 & 20 \end{bmatrix}$$

2. For  $A = \begin{bmatrix} -1 & 5 \\ -3 & 6 \end{bmatrix}$  and  $B = \begin{bmatrix} 2 & 10 \\ -7 & 1 \end{bmatrix}$

$$\bullet 3A = \begin{bmatrix} -3 & 15 \\ -9 & 18 \end{bmatrix}$$

$$\bullet -B = \begin{bmatrix} -2 & -10 \\ 7 & -1 \end{bmatrix}$$

$$\bullet A^2 = \begin{bmatrix} -14 & 25 \\ -15 & 21 \end{bmatrix}$$

$$\bullet A - 2B = \begin{bmatrix} -5 & -15 \\ 11 & 4 \end{bmatrix}$$

$$\bullet AB = \begin{bmatrix} -37 & -5 \\ -48 & -24 \end{bmatrix}$$

$$\bullet BA = \begin{bmatrix} -32 & 70 \\ 4 & -29 \end{bmatrix}$$

3. For  $A = \begin{bmatrix} -1 & 3 \\ 5 & 2 \end{bmatrix}$  and  $B = \begin{bmatrix} 7 & 0 & 8 \\ -3 & 1 & 4 \end{bmatrix}$

$$\bullet 3A = \begin{bmatrix} -3 & 9 \\ 15 & 6 \end{bmatrix}$$

$$\bullet -B = \begin{bmatrix} -7 & 0 & -8 \\ 3 & -1 & -4 \end{bmatrix}$$

$$\bullet A^2 = \begin{bmatrix} 16 & 3 \\ 5 & 19 \end{bmatrix}$$

$$\bullet A - 2B \text{ is not defined}$$

$$\bullet AB = \begin{bmatrix} -16 & 3 & 4 \\ 29 & 2 & 48 \end{bmatrix}$$

$$\bullet BA \text{ is not defined}$$

4. For  $A = \begin{bmatrix} 2 & 4 \\ 6 & 8 \end{bmatrix}$  and  $B = \begin{bmatrix} -1 & 3 & -5 \\ 7 & -9 & 11 \end{bmatrix}$

$$\bullet 3A = \begin{bmatrix} 6 & 12 \\ 18 & 24 \end{bmatrix}$$

$$\bullet -B = \begin{bmatrix} 1 & -3 & 5 \\ -7 & 9 & -11 \end{bmatrix}$$

$$\bullet A^2 = \begin{bmatrix} 28 & 40 \\ 60 & 88 \end{bmatrix}$$

$$\bullet A - 2B \text{ is not defined}$$

$$\bullet AB = \begin{bmatrix} 26 & -30 & 34 \\ 50 & -54 & 58 \end{bmatrix}$$

$$\bullet BA \text{ is not defined}$$

5. For  $A = \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix}$  and  $B = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$

•  $3A = \begin{bmatrix} 21 \\ 24 \\ 27 \end{bmatrix}$

•  $A^2$  is not defined

•  $AB = \begin{bmatrix} 7 & 14 & 21 \\ 8 & 16 & 24 \\ 9 & 18 & 27 \end{bmatrix}$

•  $-B = \begin{bmatrix} -1 & -2 & -3 \end{bmatrix}$

•  $A - 2B$  is not defined

•  $BA = [50]$

6. For  $A = \begin{bmatrix} 1 & -2 \\ -3 & 4 \\ 5 & -6 \end{bmatrix}$  and  $B = \begin{bmatrix} -5 & 1 & 8 \end{bmatrix}$

•  $3A = \begin{bmatrix} 3 & -6 \\ -9 & 12 \\ 15 & -18 \end{bmatrix}$

•  $A^2$  is not defined

•  $AB$  is not defined

•  $-B = \begin{bmatrix} 5 & -1 & -8 \end{bmatrix}$

•  $A - 2B$  is not defined

•  $BA = \begin{bmatrix} 32 & -34 \end{bmatrix}$

7. For  $A = \begin{bmatrix} 2 & -3 & 5 \\ 3 & 1 & -2 \\ -7 & 1 & -1 \end{bmatrix}$  and  $B = \begin{bmatrix} 1 & 2 & 1 \\ 17 & 33 & 19 \\ 10 & 19 & 11 \end{bmatrix}$

•  $3A = \begin{bmatrix} 6 & -9 & 15 \\ 9 & 3 & -6 \\ -21 & 3 & -3 \end{bmatrix}$

•  $A^2 = \begin{bmatrix} -40 & -4 & 11 \\ 23 & -10 & 15 \\ -4 & 21 & -36 \end{bmatrix}$

•  $AB = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

•  $-B = \begin{bmatrix} -1 & -2 & -1 \\ -17 & -33 & -19 \\ -10 & -19 & -11 \end{bmatrix}$

•  $A - 2B = \begin{bmatrix} 0 & -7 & 3 \\ -31 & -65 & -40 \\ -27 & -37 & -23 \end{bmatrix}$

•  $BA = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

8.  $7B - 4A = \begin{bmatrix} -4 & -29 \\ -47 & -2 \end{bmatrix}$

9.  $AB = \begin{bmatrix} -10 & 1 \\ -20 & -1 \end{bmatrix}$

$$10. BA = \begin{bmatrix} -9 & -12 \\ 1 & -2 \end{bmatrix}$$

$$12. ED = \begin{bmatrix} \frac{67}{3} & 11 \\ -\frac{178}{3} & -72 \\ -30 & -40 \end{bmatrix}$$

$$14. A - 4I_2 = \begin{bmatrix} -3 & 2 \\ 3 & 0 \end{bmatrix}$$

$$16. (A + B)(A - B) = \begin{bmatrix} -7 & 3 \\ 46 & 2 \end{bmatrix}$$

$$18. E^2 + 5E - 36I_3 = \begin{bmatrix} -30 & 20 & -15 \\ 0 & 0 & -36 \\ 0 & 0 & -36 \end{bmatrix}$$

$$20. CDE \text{ is undefined}$$

$$22. E_1A = \begin{bmatrix} d & e & f \\ a & b & c \end{bmatrix} \quad E_1 \text{ interchanged } R1 \text{ and } R2 \text{ of } A.$$

$$E_2A = \begin{bmatrix} 5a & 5b & 5c \\ d & e & f \end{bmatrix} \quad E_2 \text{ multiplied } R1 \text{ of } A \text{ by } 5.$$

$$E_3A = \begin{bmatrix} a - 2d & b - 2e & c - 2f \\ d & e & f \end{bmatrix} \quad E_3 \text{ replaced } R1 \text{ in } A \text{ with } R1 - 2R2.$$

$$E_4 = \begin{bmatrix} 1 & 0 \\ 0 & -6 \end{bmatrix}$$

$$11. E + D \text{ is undefined}$$

$$13. CD + 2I_2A = \begin{bmatrix} \frac{238}{3} & -126 \\ \frac{863}{15} & \frac{361}{5} \end{bmatrix}$$

$$15. A^2 - B^2 = \begin{bmatrix} -8 & 16 \\ 25 & 3 \end{bmatrix}$$

$$17. A^2 - 5A - 2I_2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$19. EDC = \begin{bmatrix} \frac{3449}{15} & -\frac{407}{6} & 99 \\ -\frac{9548}{15} & -\frac{101}{3} & -648 \\ -324 & -35 & -360 \end{bmatrix}$$

$$21. ABCEDI_2 = \begin{bmatrix} -\frac{90749}{15} & -\frac{28867}{5} \\ -\frac{156601}{15} & -\frac{47033}{5} \end{bmatrix}$$

## 1.4 Systems of Linear Equations: Matrix Inverses

We concluded Section 1.3 by showing how we can rewrite a system of linear equations as the matrix equation  $AX = B$  where  $A$  and  $B$  are known matrices and the solution matrix  $X$  of the equation corresponds to the solution of the system. In this section, we develop the method for solving such an equation. To that end, consider the system

$$\begin{cases} 2x - 3y = 16 \\ 3x + 4y = 7 \end{cases}$$

To write this as a matrix equation, we follow the procedure outlined on page 38. We find the coefficient matrix  $A$ , the unknowns matrix  $X$  and constant matrix  $B$  to be

$$A = \begin{bmatrix} 2 & -3 \\ 3 & 4 \end{bmatrix} \quad X = \begin{bmatrix} x \\ y \end{bmatrix} \quad B = \begin{bmatrix} 16 \\ 7 \end{bmatrix}$$

In order to motivate how we solve a matrix equation like  $AX = B$ , we revisit solving a similar equation involving real numbers. For instance, to solve  $3x = 5$ , we divide both sides by 3 and obtain  $x = \frac{5}{3}$ . How can we go about defining an analogous process for matrices? To answer this question, we solve  $3x = 5$  again, but this time, we pay attention to the properties of real numbers being used at each step. Recall that dividing by 3 is the same as multiplying by  $\frac{1}{3} = 3^{-1}$ , the so-called *multiplicative inverse*<sup>1</sup> of 3.

$$\begin{array}{ll} 3x & = 5 \\ 3^{-1}(3x) & = 3^{-1}(5) \quad \text{Multiply by the (multiplicative) inverse of 3} \\ (3^{-1} \cdot 3)x & = 3^{-1}(5) \quad \text{Associative property of multiplication} \\ 1 \cdot x & = 3^{-1}(5) \quad \text{Inverse property} \\ x & = 3^{-1}(5) \quad \text{Multiplicative Identity} \end{array}$$

If we wish to check our answer, we substitute  $x = 3^{-1}(5)$  into the original equation

$$\begin{array}{ll} 3x & \stackrel{?}{=} 5 \\ 3(3^{-1}(5)) & \stackrel{?}{=} 5 \\ (3 \cdot 3^{-1})(5) & \stackrel{?}{=} 5 \quad \text{Associative property of multiplication} \\ 1 \cdot 5 & \stackrel{?}{=} 5 \quad \text{Inverse property} \\ 5 & \stackrel{\checkmark}{=} 5 \quad \text{Multiplicative Identity} \end{array}$$

Thinking back to Theorem 1.5, we know that matrix multiplication enjoys both an associative property and a multiplicative identity. What's missing from the mix is a multiplicative inverse for the coefficient matrix  $A$ . Assuming we can find such a beast, we can mimic our solution (and check) to  $3x = 5$  as follows

<sup>1</sup>Every nonzero real number  $a$  has a multiplicative inverse, denoted  $a^{-1}$ , such that  $a^{-1} \cdot a = a \cdot a^{-1} = 1$ .

Solving $AX = B$	Checking our answer
$AX = B$	$AX \stackrel{?}{=} B$
$A^{-1}(AX) = A^{-1}B$	$A(A^{-1}B) \stackrel{?}{=} B$
$(A^{-1}A)X = A^{-1}B$	$(AA^{-1})B \stackrel{?}{=} B$
$I_2X = A^{-1}B$	$I_2B \stackrel{?}{=} B$
$X = A^{-1}B$	$B \stackrel{\checkmark}{=} B$

The matrix  $A^{-1}$  is read ‘A-inverse’ and we will define it formally later in the section. At this stage, we have no idea if such a matrix  $A^{-1}$  exists, but that won’t deter us from trying to find it.<sup>2</sup>

We want  $A^{-1}$  to satisfy two equations,  $A^{-1}A = I_2$  and  $AA^{-1} = I_2$ , making  $A^{-1}$  necessarily a  $2 \times 2$  matrix.<sup>3</sup> Hence,  $A^{-1}$  has the form

$$A^{-1} = \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix}$$

for real numbers  $x_1, x_2, x_3$  and  $x_4$ . For reasons which will become clear later, we focus our attention on the equation  $AA^{-1} = I_2$ . We have

$$\begin{aligned}
 AA^{-1} &= I_2 \\
 \begin{bmatrix} 2 & -3 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix} &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\
 \begin{bmatrix} 2x_1 - 3x_3 & 2x_2 - 3x_4 \\ 3x_1 + 4x_3 & 3x_2 + 4x_4 \end{bmatrix} &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}
 \end{aligned}$$

This gives rise to *two* more systems of equations

$$\begin{cases} 2x_1 - 3x_3 = 1 \\ 3x_1 + 4x_3 = 0 \end{cases} \quad \begin{cases} 2x_2 - 3x_4 = 0 \\ 3x_2 + 4x_4 = 1 \end{cases}$$

At this point, it may seem absurd to continue with this venture. After all, the intent was to solve *one* system of equations, and in doing so, we have produced *two* more to solve. Remember, the objective of this discussion is to develop a general *method* which, when used in the correct scenarios, allows us to do far more than just solve a system of equations.

If we set about solving these systems using augmented matrices using the techniques in Section 1.2, we see that not only do both systems have the same coefficient matrix, this coefficient matrix is none other than the matrix  $A$  itself! (We will come back to this observation in a moment.)

<sup>2</sup>Much like Carl’s quest to find Sasquatch.

<sup>3</sup>Since matrix multiplication isn’t necessarily commutative, at this stage, these are two different equations.

$$\begin{cases} 2x_1 - 3x_3 = 1 \\ 3x_1 + 4x_3 = 0 \end{cases} \xrightarrow{\text{Encode into a matrix}} \left[ \begin{array}{cc|c} 2 & -3 & 1 \\ 3 & 4 & 0 \end{array} \right]$$

$$\begin{cases} 2x_2 - 3x_4 = 0 \\ 3x_2 + 4x_4 = 1 \end{cases} \xrightarrow{\text{Encode into a matrix}} \left[ \begin{array}{cc|c} 2 & -3 & 0 \\ 3 & 4 & 1 \end{array} \right]$$

To solve these two systems, we use Gauss-Jordan Elimination to put the augmented matrices into reduced row echelon form. (We leave the details to the reader.) For the first system, we get

$$\left[ \begin{array}{cc|c} 2 & -3 & 1 \\ 3 & 4 & 0 \end{array} \right] \xrightarrow{\text{Gauss Jordan Elimination}} \left[ \begin{array}{cc|c} 1 & 0 & \frac{4}{17} \\ 0 & 1 & -\frac{3}{17} \end{array} \right]$$

which gives  $x_1 = \frac{4}{17}$  and  $x_3 = -\frac{3}{17}$ . To solve the second system, we use the exact same row operations, in the same order, and we obtain

$$\left[ \begin{array}{cc|c} 2 & -3 & 0 \\ 3 & 4 & 1 \end{array} \right] \xrightarrow{\text{Gauss Jordan Elimination}} \left[ \begin{array}{cc|c} 1 & 0 & \frac{3}{17} \\ 0 & 1 & \frac{2}{17} \end{array} \right]$$

which means  $x_2 = \frac{3}{17}$  and  $x_4 = \frac{2}{17}$ . Hence,

$$A^{-1} = \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix} = \begin{bmatrix} \frac{4}{17} & \frac{3}{17} \\ -\frac{3}{17} & \frac{2}{17} \end{bmatrix}$$

We can check to see that  $A^{-1}$  behaves as it should by computing  $AA^{-1}$

$$AA^{-1} = \begin{bmatrix} 2 & -3 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} \frac{4}{17} & \frac{3}{17} \\ -\frac{3}{17} & \frac{2}{17} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2 \checkmark$$

As an added bonus,

$$A^{-1}A = \begin{bmatrix} \frac{4}{17} & \frac{3}{17} \\ -\frac{3}{17} & \frac{2}{17} \end{bmatrix} \begin{bmatrix} 2 & -3 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2 \checkmark$$

We can now return to the problem at hand. From our discussion on page 47, we know

$$X = A^{-1}B = \begin{bmatrix} \frac{4}{17} & \frac{3}{17} \\ -\frac{3}{17} & \frac{2}{17} \end{bmatrix} \begin{bmatrix} 16 \\ 7 \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \end{bmatrix}$$

so that our final solution to the system is  $(x, y) = (5, -2)$ .

As mentioned, the point of this exercise was not just to solve the system of linear equations, but to develop a general method for finding  $A^{-1}$ . To that end, we analyze the foregoing discussion in a more general context. To find  $A^{-1}$ , we used two augmented matrices, both containing the entries as  $A$ :



$$\left[ \begin{array}{cc|c} 2 & -3 & 1 \\ 3 & 4 & 0 \end{array} \right] = \left[ \begin{array}{c|c} A & \begin{matrix} 1 \\ 0 \end{matrix} \end{array} \right]$$

$$\left[ \begin{array}{cc|c} 2 & -3 & 0 \\ 3 & 4 & 1 \end{array} \right] = \left[ \begin{array}{c|c} A & \begin{matrix} 0 \\ 1 \end{matrix} \end{array} \right]$$

We also note that the reduced row echelon forms of these augmented matrices can be written as

$$\left[ \begin{array}{cc|c} 1 & 0 & \frac{4}{17} \\ 0 & 1 & -\frac{3}{17} \end{array} \right] = \left[ \begin{array}{c|c} I_2 & \begin{matrix} x_1 \\ x_3 \end{matrix} \end{array} \right]$$

$$\left[ \begin{array}{cc|c} 1 & 0 & \frac{3}{17} \\ 0 & 1 & \frac{2}{17} \end{array} \right] = \left[ \begin{array}{c|c} I_2 & \begin{matrix} x_2 \\ x_4 \end{matrix} \end{array} \right]$$

where we have identified the entries to the left of the vertical bar as the identity  $I_2$  and the entries to the right of the vertical bar as the solutions to our systems. The long and short of the solution process can be summarized as

$$\left[ \begin{array}{c|c} A & \begin{matrix} 1 \\ 0 \end{matrix} \end{array} \right] \xrightarrow{\text{Gauss Jordan Elimination}} \left[ \begin{array}{c|c} I_2 & \begin{matrix} x_1 \\ x_3 \end{matrix} \end{array} \right]$$

$$\left[ \begin{array}{c|c} A & \begin{matrix} 0 \\ 1 \end{matrix} \end{array} \right] \xrightarrow{\text{Gauss Jordan Elimination}} \left[ \begin{array}{c|c} I_2 & \begin{matrix} x_2 \\ x_4 \end{matrix} \end{array} \right]$$

Since the row operations for both processes are the same, all of the arithmetic on the left hand side of the vertical bar is identical in both problems. The only difference between the two processes is what happens to the constants to the right of the vertical bar. As long as we keep these separated into columns, we can combine our efforts into one ‘super-sized’ augmented matrix and describe the above process as

$$\left[ \begin{array}{c|cc} A & \begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix} \end{array} \right] \xrightarrow{\text{Gauss Jordan Elimination}} \left[ \begin{array}{c|cc} I_2 & \begin{matrix} x_1 & x_2 \\ x_3 & x_4 \end{matrix} \end{array} \right]$$

We have the identity matrix  $I_2$  appearing as the right hand side of the first super-sized augmented matrix and the left hand side of the second super-sized augmented matrix. To our surprise and delight, the elements on the right hand side of the second super-sized augmented matrix are none other than those which comprise  $A^{-1}$ . Hence, we have

$$\left[ \begin{array}{c|c} A & I_2 \end{array} \right] \xrightarrow{\text{Gauss Jordan Elimination}} \left[ \begin{array}{c|c} I_2 & A^{-1} \end{array} \right]$$

In other words, the process of finding  $A^{-1}$  for a matrix  $A$  can be viewed as performing a series of row operations which transform  $A$  into the identity matrix of the same dimension. We can view this process as follows. In trying to find  $A^{-1}$ , we are trying to ‘undo’ multiplication by the matrix  $A$ . The identity matrix in the super-sized augmented matrix  $[A|I]$  keeps a running memory of all of the moves required to ‘undo’  $A$ . This results in exactly what we want,  $A^{-1}$ . We are now ready to formalize and generalize the foregoing discussion. We begin with the formal definition of an invertible matrix.<sup>4</sup>

<sup>4</sup>If this rhetoric sounds familiar, it should. See section ??.

**DEFINITION 1.10.** An  $n \times n$  matrix  $A$  is said to be **invertible** if there exists a matrix  $A^{-1}$ , read ‘ $A$  inverse’, such that  $A^{-1}A = AA^{-1} = I_n$ .

Note that, as a consequence of our definition, invertible matrices are square, and as such, the conditions in Definition 1.10 force the matrix  $A^{-1}$  to be same dimensions as  $A$ , that is,  $n \times n$ . Since not all matrices are square, not all matrices are invertible. However, just because a matrix is square doesn’t guarantee it is invertible. (See the exercises.)

Our first result summarizes some of the important characteristics of invertible matrices and their inverses.

**THEOREM 1.6.** Suppose  $A$  is an  $n \times n$  matrix.

1. If  $A$  is invertible then  $A^{-1}$  is unique.
2.  $A$  is invertible if and only if  $AX = B$  has a unique solution for every  $n \times r$  matrix  $B$ .

The proofs of the properties in Theorem 1.6 rely on a healthy mix of definition and matrix arithmetic.

To establish the first property, we assume that  $A$  is invertible and suppose the matrices  $B$  and  $C$  act as inverses for  $A$ . That is,  $BA = AB = I_n$  and  $CA = AC = I_n$ . We need to show that  $B$  and  $C$  are, in fact, the same matrix. To see this, we note that  $B = I_n B = (CA)B = C(AB) = CI_n = C$ . Hence, any two matrices that act like  $A^{-1}$  are, in fact, the same matrix.<sup>5</sup>

To prove the second property of Theorem 1.6, we note that if  $A$  is invertible then the discussion on page 47 shows the solution to  $AX = B$  to be  $X = A^{-1}B$ , and since  $A^{-1}$  is unique, so is  $A^{-1}B$ .

Conversely, if  $AX = B$  has a unique solution for every  $n \times r$  matrix  $B$ , then, in particular, there is a unique solution  $X_0$  to the equation  $AX = I_n$ . The solution matrix  $X_0$  is our candidate for  $A^{-1}$ . We have  $AX_0 = I_n$  by definition, but we need to also show  $X_0A = I_n$ . To that end, we note that  $A(X_0A) = (AX_0)A = I_nA = A$ . In other words, the matrix  $X_0A$  is a solution to the equation  $AX = A$ . Clearly,  $X = I_n$  is also a solution to the equation  $AX = A$ , and since we are assuming every such equation has a *unique* solution, we must have  $X_0A = I_n$ . Hence, we have  $X_0A = AX_0 = I_n$ , so that  $X_0 = A^{-1}$  and  $A$  is invertible.

The foregoing discussion justifies our quest to find  $A^{-1}$  using our super-sized augmented matrix approach

$$\left[ A \mid I_n \right] \xrightarrow{\text{Gauss Jordan Elimination}} \left[ I_n \mid A^{-1} \right]$$

We are, in essence, trying to find the unique solution to the equation  $AX = I_n$  using row operations.

What does all of this mean for a system of linear equations? Theorem 1.6 tells us that if we write the system in the form  $AX = B$ , then if the coefficient matrix  $A$  is invertible, there is only one solution to the system — that is, if  $A$  is invertible, the system is consistent and independent.<sup>6</sup>

We also know that the process by which we find  $A^{-1}$  is determined completely by  $A$ , and not by the constants in  $B$ . This answers the question as to why we would bother doing row operations on a super-sized augmented matrix to find  $A^{-1}$  instead of an ordinary augmented matrix to solve a system; by finding

<sup>5</sup>If this proof sounds familiar, it should. See the discussion following Theorem ??.

<sup>6</sup>It can be shown that a matrix is invertible if and only if when it serves as a coefficient matrix for a system of equations, the system is always consistent independent. It amounts to the second property in Theorem 1.6 where the matrices  $B$  are restricted to being  $n \times 1$  matrices. We note that, owing to how matrix multiplication is defined, being able to find unique solutions to  $AX = B$  for  $n \times 1$  matrices  $B$  gives you the same statement about solving such equations for  $n \times r$  matrices — since we can find a unique solution to them one column at a time.

$A^{-1}$  we have done all of the row operations we ever need to do, once and for all, since we can quickly solve *any* equation  $AX = B$  using *one* multiplication,  $A^{-1}B$ .

EXAMPLE 1.4.1. Let  $A = \begin{bmatrix} 3 & 1 & 2 \\ 0 & -1 & 5 \\ 2 & 1 & 4 \end{bmatrix}$

1. Use row operations to find  $A^{-1}$ . Check your answer by finding  $A^{-1}A$  and  $AA^{-1}$ .
2. Use  $A^{-1}$  to solve the following systems of equations

$$(a) \begin{cases} 3x + y + 2z = 26 \\ -y + 5z = 39 \\ 2x + y + 4z = 117 \end{cases} \quad (b) \begin{cases} 3x + y + 2z = 4 \\ -y + 5z = 2 \\ 2x + y + 4z = 5 \end{cases} \quad (c) \begin{cases} 3x + y + 2z = 1 \\ -y + 5z = 0 \\ 2x + y + 4z = 0 \end{cases}$$

**Solution.**

1. We begin with a super-sized augmented matrix and proceed with Gauss-Jordan elimination.

$$\begin{aligned} & \left[ \begin{array}{ccc|ccc} 3 & 1 & 2 & 1 & 0 & 0 \\ 0 & -1 & 5 & 0 & 1 & 0 \\ 2 & 1 & 4 & 0 & 0 & 1 \end{array} \right] \xrightarrow[\text{with } \frac{1}{3}R1]{\text{Replace } R1} \left[ \begin{array}{ccc|ccc} 1 & \frac{1}{3} & \frac{2}{3} & \frac{1}{3} & 0 & 0 \\ 0 & -1 & 5 & 0 & 1 & 0 \\ 2 & 1 & 4 & 0 & 0 & 1 \end{array} \right] \\ & \left[ \begin{array}{ccc|ccc} 1 & \frac{1}{3} & \frac{2}{3} & \frac{1}{3} & 0 & 0 \\ 0 & -1 & 5 & 0 & 1 & 0 \\ 2 & 1 & 4 & 0 & 0 & 1 \end{array} \right] \xrightarrow[-2R1 + R3]{\text{Replace } R3 \text{ with}} \left[ \begin{array}{ccc|ccc} 1 & \frac{1}{3} & \frac{2}{3} & \frac{1}{3} & 0 & 0 \\ 0 & -1 & 5 & 0 & 1 & 0 \\ 0 & \frac{1}{3} & \frac{8}{3} & -\frac{2}{3} & 0 & 1 \end{array} \right] \\ & \left[ \begin{array}{ccc|ccc} 1 & \frac{1}{3} & \frac{2}{3} & \frac{1}{3} & 0 & 0 \\ 0 & -1 & 5 & 0 & 1 & 0 \\ 0 & \frac{1}{3} & \frac{8}{3} & -\frac{2}{3} & 0 & 1 \end{array} \right] \xrightarrow[\text{with } (-1)R2]{\text{Replace } R2} \left[ \begin{array}{ccc|ccc} 1 & \frac{1}{3} & \frac{2}{3} & \frac{1}{3} & 0 & 0 \\ 0 & 1 & -5 & 0 & -1 & 0 \\ 0 & \frac{1}{3} & \frac{8}{3} & -\frac{2}{3} & 0 & 1 \end{array} \right] \\ & \left[ \begin{array}{ccc|ccc} 1 & \frac{1}{3} & \frac{2}{3} & \frac{1}{3} & 0 & 0 \\ 0 & 1 & -5 & 0 & -1 & 0 \\ 0 & \frac{1}{3} & \frac{8}{3} & -\frac{2}{3} & 0 & 1 \end{array} \right] \xrightarrow[-\frac{1}{3}R2 + R3]{\text{Replace } R3 \text{ with}} \left[ \begin{array}{ccc|ccc} 1 & \frac{1}{3} & \frac{2}{3} & \frac{1}{3} & 0 & 0 \\ 0 & 1 & -5 & 0 & -1 & 0 \\ 0 & 0 & \frac{13}{3} & -\frac{2}{3} & \frac{1}{3} & 1 \end{array} \right] \\ & \left[ \begin{array}{ccc|ccc} 1 & \frac{1}{3} & \frac{2}{3} & \frac{1}{3} & 0 & 0 \\ 0 & 1 & -5 & 0 & -1 & 0 \\ 0 & 0 & \frac{13}{3} & -\frac{2}{3} & \frac{1}{3} & 1 \end{array} \right] \xrightarrow[\text{with } \frac{3}{13}R3]{\text{Replace } R3} \left[ \begin{array}{ccc|ccc} 1 & \frac{1}{3} & \frac{2}{3} & \frac{1}{3} & 0 & 0 \\ 0 & 1 & -5 & 0 & -1 & 0 \\ 0 & 0 & 1 & -\frac{2}{13} & \frac{1}{13} & \frac{3}{13} \end{array} \right] \\ & \left[ \begin{array}{ccc|ccc} 1 & \frac{1}{3} & \frac{2}{3} & \frac{1}{3} & 0 & 0 \\ 0 & 1 & -5 & 0 & -1 & 0 \\ 0 & 0 & 1 & -\frac{2}{13} & \frac{1}{13} & \frac{3}{13} \end{array} \right] \xrightarrow[\text{Replace } R2 \text{ with } 5R3 + R2]{\text{Replace } R1 \text{ with } -\frac{2}{3}R3 + R1} \left[ \begin{array}{ccc|ccc} 1 & \frac{1}{3} & 0 & \frac{17}{39} & -\frac{2}{39} & -\frac{2}{13} \\ 0 & 1 & 0 & -\frac{10}{13} & -\frac{8}{13} & \frac{15}{13} \\ 0 & 0 & 1 & -\frac{2}{13} & \frac{1}{13} & \frac{3}{13} \end{array} \right] \\ & \left[ \begin{array}{ccc|ccc} 1 & \frac{1}{3} & 0 & \frac{17}{39} & -\frac{2}{39} & -\frac{2}{13} \\ 0 & 1 & 0 & -\frac{10}{13} & -\frac{8}{13} & \frac{15}{13} \\ 0 & 0 & 1 & -\frac{2}{13} & \frac{1}{13} & \frac{3}{13} \end{array} \right] \xrightarrow[-\frac{1}{3}R2 + R1]{\text{Replace } R1 \text{ with}} \left[ \begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{9}{13} & \frac{2}{13} & -\frac{7}{13} \\ 0 & 1 & 0 & -\frac{10}{13} & -\frac{8}{13} & \frac{15}{13} \\ 0 & 0 & 1 & -\frac{2}{13} & \frac{1}{13} & \frac{3}{13} \end{array} \right] \end{aligned}$$

We find  $A^{-1} = \begin{bmatrix} \frac{9}{13} & \frac{2}{13} & -\frac{7}{13} \\ -\frac{10}{13} & -\frac{8}{13} & \frac{15}{13} \\ -\frac{2}{13} & \frac{1}{13} & \frac{3}{13} \end{bmatrix}$ . To check our answer, we compute

$$A^{-1}A = \begin{bmatrix} \frac{9}{13} & \frac{2}{13} & -\frac{7}{13} \\ -\frac{10}{13} & -\frac{8}{13} & \frac{15}{13} \\ -\frac{2}{13} & \frac{1}{13} & \frac{3}{13} \end{bmatrix} \begin{bmatrix} 3 & 1 & 2 \\ 0 & -1 & 5 \\ 2 & 1 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_3 \checkmark$$

and

$$AA^{-1} = \begin{bmatrix} 3 & 1 & 2 \\ 0 & -1 & 5 \\ 2 & 1 & 4 \end{bmatrix} \begin{bmatrix} \frac{9}{13} & \frac{2}{13} & -\frac{7}{13} \\ -\frac{10}{13} & -\frac{8}{13} & \frac{15}{13} \\ -\frac{2}{13} & \frac{1}{13} & \frac{3}{13} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_3 \checkmark$$

2. Each of the systems in this part has  $A$  as its coefficient matrix. The only difference between the systems is the constants which is the matrix  $B$  in the associated matrix equation  $AX = B$ . We solve each of them using the formula  $X = A^{-1}B$ .

$$(a) X = A^{-1}B = \begin{bmatrix} \frac{9}{13} & \frac{2}{13} & -\frac{7}{13} \\ -\frac{10}{13} & -\frac{8}{13} & \frac{15}{13} \\ -\frac{2}{13} & \frac{1}{13} & \frac{3}{13} \end{bmatrix} \begin{bmatrix} 26 \\ 39 \\ 117 \end{bmatrix} = \begin{bmatrix} -39 \\ 91 \\ 26 \end{bmatrix}. \text{ Our solution is } (-39, 91, 26).$$

$$(b) X = A^{-1}B = \begin{bmatrix} \frac{9}{13} & \frac{2}{13} & -\frac{7}{13} \\ -\frac{10}{13} & -\frac{8}{13} & \frac{15}{13} \\ -\frac{2}{13} & \frac{1}{13} & \frac{3}{13} \end{bmatrix} \begin{bmatrix} 4 \\ 2 \\ 5 \end{bmatrix} = \begin{bmatrix} \frac{5}{13} \\ \frac{19}{13} \\ \frac{9}{13} \end{bmatrix}. \text{ We get } (\frac{5}{13}, \frac{19}{13}, \frac{9}{13}).$$

$$(c) X = A^{-1}B = \begin{bmatrix} \frac{9}{13} & \frac{2}{13} & -\frac{7}{13} \\ -\frac{10}{13} & -\frac{8}{13} & \frac{15}{13} \\ -\frac{2}{13} & \frac{1}{13} & \frac{3}{13} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{9}{13} \\ -\frac{10}{13} \\ -\frac{2}{13} \end{bmatrix}. \text{ We find } (\frac{9}{13}, -\frac{10}{13}, -\frac{2}{13}).^7$$

□

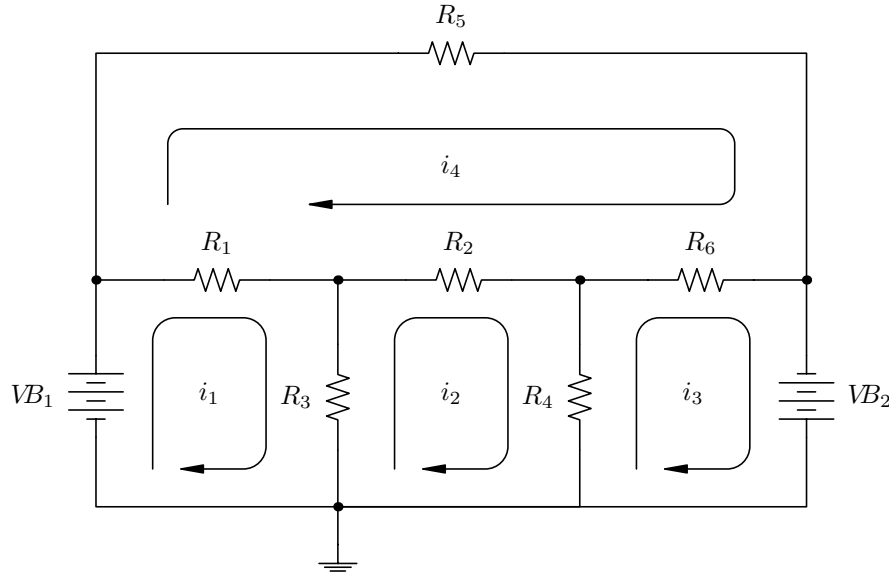
In Example 1.4.1, we see that finding one inverse matrix can enable us to solve an entire family of systems of linear equations. There are many examples of where this comes in handy ‘in the wild’, and we chose our example for this section from the field of electronics. We also take this opportunity to introduce the student to how we can compute inverse matrices using the calculator.

EXAMPLE 1.4.2. Consider the circuit diagram below.<sup>8</sup> We have two batteries with source voltages  $VB_1$  and  $VB_2$ , measured in volts  $V$ , along with six resistors with resistances  $R_1$  through  $R_6$ , measured in kilohms,  $k\Omega$ . Using [Ohm’s Law](#) and [Kirchhoff’s Voltage Law](#), we can relate the voltage supplied to the circuit by the two batteries to the voltage drops across the six resistors in order to find the four ‘mesh’ currents:  $i_1$ ,  $i_2$ ,  $i_3$  and  $i_4$ , measured in milliamps,  $mA$ . If we think of electrons flowing through the circuit, we can think of the

<sup>7</sup>Note that the solution is the first column of the  $A^{-1}$ . The reader is encouraged to meditate on this ‘coincidence’.

<sup>8</sup>The authors wish to thank Don Anthan of Lakeland Community College for the design of this example.

voltage sources as providing the 'push' which makes the electrons move, the resistors as obstacles for the electrons to overcome, and the mesh current as a net rate of flow of electrons around the indicated loops.



The system of linear equations associated with this circuit is

$$\begin{cases} (R_1 + R_3) i_1 - R_3 i_2 - R_1 i_4 = VB_1 \\ -R_3 i_1 + (R_2 + R_3 + R_4) i_2 - R_4 i_3 - R_2 i_4 = 0 \\ -R_4 i_2 + (R_4 + R_6) i_3 - R_6 i_4 = -VB_2 \\ -R_1 i_1 - R_2 i_2 - R_6 i_3 + (R_1 + R_2 + R_5 + R_6) i_4 = 0 \end{cases}$$

Assuming the resistances are all  $1\text{ k}\Omega$ , find the mesh currents if the battery voltages are

- $VB_1 = 10\text{ V}$  and  $VB_2 = 5\text{ V}$
- $VB_1 = 10\text{ V}$  and  $VB_2 = 0\text{ V}$
- $VB_1 = 0\text{ V}$  and  $VB_2 = 10\text{ V}$
- $VB_1 = 10\text{ V}$  and  $VB_2 = 10\text{ V}$

**Solution.**

Substituting the resistance values into our system of equations, we get

$$\begin{cases} 2i_1 - i_2 - i_4 = VB_1 \\ -i_1 + 3i_2 - i_3 - i_4 = 0 \\ -i_2 + 2i_3 - i_4 = -VB_2 \\ -i_1 - i_2 - i_3 + 4i_4 = 0 \end{cases}$$

This corresponds to the matrix equation  $AX = B$  where

$$A = \begin{bmatrix} 2 & -1 & 0 & -1 \\ -1 & 3 & -1 & -1 \\ 0 & -1 & 2 & -1 \\ -1 & -1 & -1 & 4 \end{bmatrix} \quad X = \begin{bmatrix} i_1 \\ i_2 \\ i_3 \\ i_4 \end{bmatrix} \quad B = \begin{bmatrix} VB_1 \\ 0 \\ -VB_2 \\ 0 \end{bmatrix}$$

When we input the matrix  $A$  into the calculator, we find

$$[A]^{-1} = \begin{bmatrix} 1.625 & 1.25 & 1.125 & 1 \\ 1.25 & 1.5 & 1.25 & 1 \\ 1.125 & 1.25 & 1.625 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

$$[A]^{-1} = \begin{bmatrix} 1.625 & 1.25 & 1.125 & 1 \\ 1.25 & 1.5 & 1.25 & 1 \\ 1.125 & 1.25 & 1.625 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

from which we have  $A^{-1} = \begin{bmatrix} 1.625 & 1.25 & 1.125 & 1 \\ 1.25 & 1.5 & 1.25 & 1 \\ 1.125 & 1.25 & 1.625 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$ .

To solve the four systems given to us, we find  $X = A^{-1}B$  where the value of  $B$  is determined by the given values of  $VB_1$  and  $VB_2$

$$B = \begin{bmatrix} 10 \\ 0 \\ -5 \\ 0 \end{bmatrix}, \quad B = \begin{bmatrix} 10 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ -10 \\ 0 \end{bmatrix}, \quad B = \begin{bmatrix} 10 \\ 0 \\ 10 \\ 0 \end{bmatrix}$$

- For  $VB_1 = 10V$  and  $VB_2 = 5V$ , the calculator gives  $i_1 = 10.625 \text{ mA}$ ,  $i_2 = 6.25 \text{ mA}$ ,  $i_3 = 3.125 \text{ mA}$ , and  $i_4 = 5 \text{ mA}$ . We include a calculator screenshot below for this part (and this part only!) for reference.

$$[A]^{-1}[B] = \begin{bmatrix} 10.625 \\ 6.25 \\ 3.125 \\ 5 \end{bmatrix}$$

- By keeping  $VB_1 = 10V$  and setting  $VB_2 = 0V$ , we are removing the effect of the second battery. We get  $i_1 = 16.25 \text{ mA}$ ,  $i_2 = 12.5 \text{ mA}$ ,  $i_3 = 11.25 \text{ mA}$ , and  $i_4 = 10 \text{ mA}$ .
- Here, we are zeroing out  $VB_1$  and making  $VB_2 = 10$ , essentially removing the effect of the first battery. We find  $i_1 = -11.25 \text{ mA}$ ,  $i_2 = -12.5 \text{ mA}$ ,  $i_3 = -16.25 \text{ mA}$ , and  $i_4 = -10 \text{ mA}$ , where the negatives indicate that the current is flowing in the opposite direction as is indicated on the diagram. The reader is encouraged to study the symmetry here, and if need be, hold up a mirror to the diagram to literally 'see' what is happening.
- For  $VB_1 = 10V$  and  $VB_2 = 10V$ , we get  $i_1 = 5 \text{ mA}$ ,  $i_2 = 0 \text{ mA}$ ,  $i_3 = -5 \text{ mA}$ , and  $i_4 = 0 \text{ mA}$ . The mesh currents  $i_2$  and  $i_4$  being zero is a consequence of both batteries 'pushing' in equal but opposite directions, causing the net flow of electrons in these two regions to cancel out.  $\square$

## 1.4.1 Exercises

In Exercises 1 - 8, find the inverse of the matrix or state that the matrix is not invertible.

1.  $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

2.  $B = \begin{bmatrix} 12 & -7 \\ -5 & 3 \end{bmatrix}$

3.  $C = \begin{bmatrix} 6 & 15 \\ 14 & 35 \end{bmatrix}$

4.  $D = \begin{bmatrix} 2 & -1 \\ 16 & -9 \end{bmatrix}$

5.  $E = \begin{bmatrix} 3 & 0 & 4 \\ 2 & -1 & 3 \\ -3 & 2 & -5 \end{bmatrix}$

6.  $F = \begin{bmatrix} 4 & 6 & -3 \\ 3 & 4 & -3 \\ 1 & 2 & 6 \end{bmatrix}$

7.  $G = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 11 \\ 3 & 4 & 19 \end{bmatrix}$

8.  $H = \begin{bmatrix} 1 & 0 & -3 & 0 \\ 2 & -2 & 8 & 7 \\ -5 & 0 & 16 & 0 \\ 1 & 0 & 4 & 1 \end{bmatrix}$

In Exercises 9 - 11, use one matrix inverse to solve the following systems of linear equations.

9.  $\begin{cases} 3x + 7y = 26 \\ 5x + 12y = 39 \end{cases}$

10.  $\begin{cases} 3x + 7y = 0 \\ 5x + 12y = -1 \end{cases}$

11.  $\begin{cases} 3x + 7y = -7 \\ 5x + 12y = 5 \end{cases}$

In Exercises 12 - 14, use the inverse of  $E$  from Exercise 5 above to solve the following systems of linear equations.

12.  $\begin{cases} 3x + 4z = 1 \\ 2x - y + 3z = 0 \\ -3x + 2y - 5z = 0 \end{cases}$

13.  $\begin{cases} 3x + 4z = 0 \\ 2x - y + 3z = 1 \\ -3x + 2y - 5z = 0 \end{cases}$

14.  $\begin{cases} 3x + 4z = 0 \\ 2x - y + 3z = 0 \\ -3x + 2y - 5z = 1 \end{cases}$

15. This exercise is a continuation of Example 1.3.3 in Section 1.3 and gives another application of matrix inverses. Recall that given the position matrix  $P$  for a point in the plane, the matrix  $RP$  corresponds to a point rotated  $45^\circ$  counterclockwise from  $P$  where

$$R = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}$$

- Find  $R^{-1}$ .
- If  $RP$  rotates a point counterclockwise  $45^\circ$ , what should  $R^{-1}P$  do? Check your answer by finding  $R^{-1}P$  for various points on the coordinate axes and the lines  $y = \pm x$ .
- Find  $R^{-1}P$  where  $P$  corresponds to a generic point  $P(x, y)$ . Verify that this takes points on the curve  $y = \frac{2}{x}$  to points on the curve  $x^2 - y^2 = 4$ .

16. A Sasquatch's diet consists of three primary foods: Ippizuti Fish, Misty Mushrooms, and Sun Berries. Each serving of Ippizuti Fish is 500 calories, contains 40 grams of protein, and has no Vitamin X. Each serving of Misty Mushrooms is 50 calories, contains 1 gram of protein, and 5 milligrams of Vitamin X. Finally, each serving of Sun Berries is 80 calories, contains no protein, but has 15 milligrams of Vitamin X.<sup>9</sup>
- If an adult male Sasquatch requires 3200 calories, 130 grams of protein, and 275 milligrams of Vitamin X daily, use a matrix inverse to find how many servings each of Ippizuti Fish, Misty Mushrooms, and Sun Berries he needs to eat each day.
  - An adult female Sasquatch requires 3100 calories, 120 grams of protein, and 300 milligrams of Vitamin X daily. Use the matrix inverse you found in part (a) to find how many servings each of Ippizuti Fish, Misty Mushrooms, and Sun Berries she needs to eat each day.
  - An adolescent Sasquatch requires 5000 calories, 400 grams of protein daily, but no Vitamin X daily.<sup>10</sup> Use the matrix inverse you found in part (a) to find how many servings each of Ippizuti Fish, Misty Mushrooms, and Sun Berries she needs to eat each day.
17. Matrices can be used in cryptography. Suppose we wish to encode the message 'BIGFOOT LIVES'. We start by assigning a number to each letter of the alphabet, say  $A = 1$ ,  $B = 2$  and so on. We reserve 0 to act as a space. Hence, our message 'BIGFOOT LIVES' corresponds to the string of numbers '2, 9, 7, 6, 15, 15, 20, 0, 12, 9, 22, 5, 19.' To encode this message, we use an invertible matrix. Any invertible matrix will do, but for this exercise, we choose

$$A = \begin{bmatrix} 2 & -3 & 5 \\ 3 & 1 & -2 \\ -7 & 1 & -1 \end{bmatrix}$$

Since  $A$  is  $3 \times 3$  matrix, we encode our message string into a matrix  $M$  with 3 rows. To do this, we take the first three numbers, 2 9 7, and make them our first column, the next three numbers, 6 15 15, and make them our second column, and so on. We put 0's to round out the matrix.

$$M = \begin{bmatrix} 2 & 6 & 20 & 9 & 19 \\ 9 & 15 & 0 & 22 & 0 \\ 7 & 15 & 12 & 5 & 0 \end{bmatrix}$$

To encode the message, we find the product  $AM$

$$AM = \begin{bmatrix} 2 & -3 & 5 \\ 3 & 1 & -2 \\ -7 & 1 & -1 \end{bmatrix} \begin{bmatrix} 2 & 6 & 20 & 9 & 19 \\ 9 & 15 & 0 & 22 & 0 \\ 7 & 15 & 12 & 5 & 0 \end{bmatrix} = \begin{bmatrix} 12 & 42 & 100 & -23 & 38 \\ 1 & 3 & 36 & 39 & 57 \\ -12 & -42 & -152 & -46 & -133 \end{bmatrix}$$

<sup>9</sup>Misty Mushrooms and Sun Berries are the only known fictional sources of Vitamin X.

<sup>10</sup>Vitamin X is needed to sustain Sasquatch longevity only.



So our coded message is '12, 1, -12, 42, 3, -42, 100, 36, -152, -23, 39, -46, 38, 57, -133.' To decode this message, we start with this string of numbers, construct a message matrix as we did earlier (we should get the matrix  $AM$  again) and then multiply by  $A^{-1}$ .

- (a) Find  $A^{-1}$ .
  - (b) Use  $A^{-1}$  to decode the message and check this method actually works.
  - (c) Decode the message '14, 37, -76, 128, 21, -151, 31, 65, -140'
  - (d) Choose another invertible matrix and encode and decode your own messages.
18. Using the matrices  $A$  from Exercise 1,  $B$  from Exercise 2 and  $D$  from Exercise 4, show  $AB = D$  and  $D^{-1} = B^{-1}A^{-1}$ . That is, show that  $(AB)^{-1} = B^{-1}A^{-1}$ .
19. Let  $M$  and  $N$  be invertible  $n \times n$  matrices. Show that  $(MN)^{-1} = N^{-1}M^{-1}$  and compare your work to Exercise ?? in Section ??.

## 1.4.2 Answers

$$1. A^{-1} = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}$$

$$2. B^{-1} = \begin{bmatrix} 3 & 7 \\ 5 & 12 \end{bmatrix}$$

3.  $C$  is not invertible

$$4. D^{-1} = \begin{bmatrix} \frac{9}{2} & -\frac{1}{2} \\ 8 & -1 \end{bmatrix}$$

$$5. E^{-1} = \begin{bmatrix} -1 & 8 & 4 \\ 1 & -3 & -1 \\ 1 & -6 & -3 \end{bmatrix}$$

$$6. F^{-1} = \begin{bmatrix} -\frac{5}{2} & \frac{7}{2} & \frac{1}{2} \\ \frac{7}{4} & -\frac{9}{4} & -\frac{1}{4} \\ -\frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{bmatrix}$$

7.  $G$  is not invertible

$$8. H^{-1} = \begin{bmatrix} 16 & 0 & 3 & 0 \\ -90 & -\frac{1}{2} & -\frac{35}{2} & \frac{7}{2} \\ 5 & 0 & 1 & 0 \\ -36 & 0 & -7 & 1 \end{bmatrix}$$

The coefficient matrix is  $B^{-1}$  from Exercise 2 above so the inverse we need is  $(B^{-1})^{-1} = B$ .

$$9. \begin{bmatrix} 12 & -7 \\ -5 & 3 \end{bmatrix} \begin{bmatrix} 26 \\ 39 \end{bmatrix} = \begin{bmatrix} 39 \\ -13 \end{bmatrix} \text{ So } x = 39 \text{ and } y = -13.$$

$$10. \begin{bmatrix} 12 & -7 \\ -5 & 3 \end{bmatrix} \begin{bmatrix} 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 7 \\ -3 \end{bmatrix} \text{ So } x = 7 \text{ and } y = -3.$$

$$11. \begin{bmatrix} 12 & -7 \\ -5 & 3 \end{bmatrix} \begin{bmatrix} -7 \\ 5 \end{bmatrix} = \begin{bmatrix} -119 \\ 50 \end{bmatrix} \text{ So } x = -119 \text{ and } y = 50.$$

The coefficient matrix is  $E = \begin{bmatrix} 3 & 0 & 4 \\ 2 & -1 & 3 \\ -3 & 2 & -5 \end{bmatrix}$  from Exercise 5, so  $E^{-1} = \begin{bmatrix} -1 & 8 & 4 \\ 1 & -3 & -1 \\ 1 & -6 & -3 \end{bmatrix}$

$$12. \begin{bmatrix} -1 & 8 & 4 \\ 1 & -3 & -1 \\ 1 & -6 & -3 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \text{ So } x = -1, y = 1 \text{ and } z = 1.$$

$$13. \begin{bmatrix} -1 & 8 & 4 \\ 1 & -3 & -1 \\ 1 & -6 & -3 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 8 \\ -3 \\ -6 \end{bmatrix} \text{ So } x = 8, y = -3 \text{ and } z = -6.$$

$$14. \begin{bmatrix} -1 & 8 & 4 \\ 1 & -3 & -1 \\ 1 & -6 & -3 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \\ -3 \end{bmatrix} \text{ So } x = 4, y = -1 \text{ and } z = -3.$$

16. (a) The adult male Sasquatch needs: 3 servings of Ippizuti Fish, 10 servings of Misty Mushrooms, and 15 servings of Sun Berries daily.
- (b) The adult female Sasquatch needs: 3 servings of Ippizuti Fish and 20 servings of Sun Berries daily. (No Misty Mushrooms are needed!)
- (c) The adolescent Sasquatch requires 10 servings of Ippizuti Fish daily. (No Misty Mushrooms or Sun Berries are needed!)

17. (a)  $A^{-1} = \begin{bmatrix} 1 & 2 & 1 \\ 17 & 33 & 19 \\ 10 & 19 & 11 \end{bmatrix}$

(b)  $\begin{bmatrix} 1 & 2 & 1 \\ 17 & 33 & 19 \\ 10 & 19 & 11 \end{bmatrix} \begin{bmatrix} 12 & 42 & 100 & -23 & 38 \\ 1 & 3 & 36 & 39 & 57 \\ -12 & -42 & -152 & -46 & -133 \end{bmatrix} = \begin{bmatrix} 2 & 6 & 20 & 9 & 19 \\ 9 & 15 & 0 & 22 & 0 \\ 7 & 15 & 12 & 5 & 0 \end{bmatrix} \quad \checkmark$

(c) 'LOGS RULE'

## 1.5 Determinants and Cramer's Rule

### 1.5.1 Definition and Properties of the Determinant

In this section we assign to each square matrix  $A$  a real number, called the *determinant* of  $A$ , which will eventually lead us to yet another technique for solving consistent independent systems of linear equations. The determinant is defined recursively, that is, we define it for  $1 \times 1$  matrices and give a rule by which we can reduce determinants of  $n \times n$  matrices to a sum of determinants of  $(n - 1) \times (n - 1)$  matrices.<sup>1</sup> This means we will be able to evaluate the determinant of a  $2 \times 2$  matrix as a sum of the determinants of  $1 \times 1$  matrices; the determinant of a  $3 \times 3$  matrix as a sum of the determinants of  $2 \times 2$  matrices, and so forth. To explain how we will take an  $n \times n$  matrix and distill an  $(n - 1) \times (n - 1)$  matrix from it, we use introduce the following notation.

DEFINITION 1.11. Given an  $n \times n$  matrix  $A$  where  $n > 1$ , the matrix  $A_{ij}$  is the  $(n - 1) \times (n - 1)$  matrix formed by deleting the  $i$ th row of  $A$  and the  $j$ th column of  $A$ .

For the matrix  $A$  below, we obtain  $A_{23}$  by deleting the second row and third column of  $A$ :

$$A = \begin{bmatrix} 3 & 1 & 2 \\ 0 & -1 & 5 \\ 2 & 1 & 4 \end{bmatrix} \xrightarrow{\text{Delete } R2 \text{ and } C3} A_{23} = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}$$

We are now in the position to define the determinant of a matrix.

DEFINITION 1.12. Given an  $n \times n$  matrix  $A$  the **determinant of  $A$** , denoted  $\det(A)$ , is defined as follows

- If  $n = 1$ , then  $A = [a_{11}]$  and  $\det(A) = \det([a_{11}]) = a_{11}$ .
- If  $n > 1$ , then  $A = [a_{ij}]_{n \times n}$  and

$$\det(A) = \det([a_{ij}]_{n \times n}) = a_{11} \det(A_{11}) - a_{12} \det(A_{12}) + \dots + (-1)^{1+n} a_{1n} \det(A_{1n})$$

There are two commonly used notations for the determinant of a matrix  $A$ : ' $\det(A)$ ' and ' $|A|$ '. We have chosen to use the notation  $\det(A)$  as opposed to  $|A|$  because we find that the latter is often confused with absolute value, especially in the context of a  $1 \times 1$  matrix.

In the expansion  $a_{11} \det(A_{11}) - a_{12} \det(A_{12}) + \dots + (-1)^{1+n} a_{1n} \det(A_{1n})$ , the notation ' $+ \dots + (-1)^{1+n} a_{1n}$ ' means that the signs alternate and the final sign is dictated by the sign of the quantity  $(-1)^{1+n}$ .

Since the entries  $a_{11}$ ,  $a_{12}$  and so forth up through  $a_{1n}$  comprise the first row of  $A$ , we say we are finding the determinant of  $A$  by 'expanding along the first row'. Later in the section, we will develop a formula for  $\det(A)$  which allows us to find it by expanding along any row.

Applying Definition 1.12 to the matrix  $A = \begin{bmatrix} 4 & -3 \\ 2 & 1 \end{bmatrix}$  we get

<sup>1</sup>We will talk more about the term 'recursively' in Section ??.

$$\begin{aligned}
\det(A) &= \det \left( \begin{bmatrix} 4 & -3 \\ 2 & 1 \end{bmatrix} \right) \\
&= 4 \det(A_{11}) - (-3) \det(A_{12}) \\
&= 4 \det([1]) + 3 \det([2]) \\
&= 4(1) + 3(2) \\
&= 10
\end{aligned}$$

For a generic  $2 \times 2$  matrix  $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$  we get

$$\begin{aligned}
\det(A) &= \det \left( \begin{bmatrix} a & b \\ c & d \end{bmatrix} \right) \\
&= a \det(A_{11}) - b \det(A_{12}) \\
&= a \det([d]) - b \det([c]) \\
&= ad - bc
\end{aligned}$$

This formula is worth remembering

EQUATION 1.1. For a  $2 \times 2$  matrix,

$$\det \left( \begin{bmatrix} a & b \\ c & d \end{bmatrix} \right) = ad - bc$$

Applying Definition 1.12 to the  $3 \times 3$  matrix  $A = \begin{bmatrix} 3 & 1 & 2 \\ 0 & -1 & 5 \\ 2 & 1 & 4 \end{bmatrix}$  we obtain

$$\begin{aligned}
\det(A) &= \det \left( \begin{bmatrix} 3 & 1 & 2 \\ 0 & -1 & 5 \\ 2 & 1 & 4 \end{bmatrix} \right) \\
&= 3 \det(A_{11}) - 1 \det(A_{12}) + 2 \det(A_{13}) \\
&= 3 \det \left( \begin{bmatrix} -1 & 5 \\ 1 & 4 \end{bmatrix} \right) - \det \left( \begin{bmatrix} 0 & 5 \\ 2 & 4 \end{bmatrix} \right) + 2 \det \left( \begin{bmatrix} 0 & -1 \\ 2 & 1 \end{bmatrix} \right) \\
&= 3((-1)(4) - (5)(1)) - ((0)(4) - (5)(2)) + 2((0)(1) - (-1)(2)) \\
&= 3(-9) - (-10) + 2(2) \\
&= -13
\end{aligned}$$

To evaluate the determinant of a  $4 \times 4$  matrix, we would have to evaluate the determinants of *four*  $3 \times 3$  matrices, each of which involves the finding the determinants of *three*  $2 \times 2$  matrices. As you can see, our method of evaluating determinants quickly gets out of hand and many of you may be reaching for the calculator. There is some mathematical machinery which can assist us in calculating determinants and we present that here. Before we state the theorem, we need some more terminology.

DEFINITION 1.13. Let  $A$  be an  $n \times n$  matrix and  $A_{ij}$  be defined as in Definition 1.11.

- The  **$ij$  minor** of  $A$ , denoted  $M_{ij}$  is defined by  $M_{ij} = \det(A_{ij})$ .
- The  **$ij$  cofactor** of  $A$ , denoted  $C_{ij}$  is defined by  $C_{ij} = (-1)^{i+j} M_{ij} = (-1)^{i+j} \det(A_{ij})$ .

We note that in Definition 1.12, the sum

$$a_{11} \det(A_{11}) - a_{12} \det(A_{12}) + \dots + (-1)^{1+n} a_{1n} \det(A_{1n})$$

can be rewritten as

$$a_{11}(-1)^{1+1} \det(A_{11}) + a_{12}(-1)^{1+2} \det(A_{12}) + \dots + a_{1n}(-1)^{1+n} \det(A_{1n})$$

which, in the language of cofactors is

$$a_{11} C_{11} + a_{12} C_{12} + \dots + a_{1n} C_{1n}$$

We are now ready to state our main theorem concerning determinants.

THEOREM 1.7. **Properties of the Determinant:** Let  $A = [a_{ij}]_{n \times n}$ .

- We may find the determinant by expanding along any row. That is, for any  $1 \leq k \leq n$ ,

$$\det(A) = a_{k1} C_{k1} + a_{k2} C_{k2} + \dots + a_{kn} C_{kn}$$

- If  $A'$  is the matrix obtained from  $A$  by:
  - interchanging any two rows, then  $\det(A') = -\det(A)$ .
  - replacing a row with a nonzero multiple (say  $c$ ) of itself, then  $\det(A') = c \det(A)$
  - replacing a row with itself plus a multiple of another row, then  $\det(A') = \det(A)$
- If  $A$  has two identical rows, or a row consisting of all 0's, then  $\det(A) = 0$ .
- If  $A$  is upper or lower triangular,<sup>a</sup> then  $\det(A)$  is the product of the entries on the main diagonal.<sup>b</sup>
- If  $B$  is an  $n \times n$  matrix, then  $\det(AB) = \det(A) \det(B)$ .
- $\det(A^n) = \det(A)^n$  for all natural numbers  $n$ .
- $A$  is invertible if and only if  $\det(A) \neq 0$ . In this case,  $\det(A^{-1}) = \frac{1}{\det(A)}$ .

<sup>a</sup>See Exercise 1.3.1 in 1.3.

<sup>b</sup>See page 34 in Section 1.3.

Unfortunately, while we can easily *demonstrate* the results in Theorem 1.7, the proofs of most of these properties are beyond the scope of this text. We could prove these properties for generic  $2 \times 2$  or even

$3 \times 3$  matrices by brute force computation, but this manner of proof belies the elegance and symmetry of the determinant. We will prove what few properties we can after we have developed some more tools such as the Principle of Mathematical Induction in Section ??.<sup>2</sup>

For the moment, let us demonstrate some of the properties listed in Theorem 1.7 on the matrix  $A$  below. (Others will be discussed in the Exercises.)

$$A = \begin{bmatrix} 3 & 1 & 2 \\ 0 & -1 & 5 \\ 2 & 1 & 4 \end{bmatrix}$$

We found  $\det(A) = -13$  by expanding along the first row. To take advantage of the 0 in the second row, we use Theorem 1.7 to find  $\det(A) = -13$  by expanding along that row.

$$\begin{aligned} \det \left( \begin{bmatrix} 3 & 1 & 2 \\ 0 & -1 & 5 \\ 2 & 1 & 4 \end{bmatrix} \right) &= 0C_{21} + (-1)C_{22} + 5C_{23} \\ &= (-1)(-1)^{2+2} \det(A_{22}) + 5(-1)^{2+3} \det(A_{23}) \\ &= -\det \left( \begin{bmatrix} 3 & 2 \\ 2 & 4 \end{bmatrix} \right) - 5 \det \left( \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix} \right) \\ &= -((3)(4) - (2)(2)) - 5((3)(1) - (2)(1)) \\ &= -8 - 5 \\ &= -13 \checkmark \end{aligned}$$

In general, the sign of  $(-1)^{i+j}$  in front of the minor in the expansion of the determinant follows an alternating pattern. Below is the pattern for  $2 \times 2$ ,  $3 \times 3$  and  $4 \times 4$  matrices, and it extends naturally to higher dimensions.

$$\begin{bmatrix} + & - \\ - & + \end{bmatrix} \quad \begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix} \quad \begin{bmatrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{bmatrix}$$

The reader is cautioned, however, against reading too much into these sign patterns. In the example above, we expanded the  $3 \times 3$  matrix  $A$  by its second row and the term which corresponds to the second entry ended up being negative even though the sign attached to the minor is (+). These signs represent only the signs of the  $(-1)^{i+j}$  in the formula; the sign of the corresponding entry as well as the minor itself determine the ultimate sign of the term in the expansion of the determinant.

To illustrate some of the other properties in Theorem 1.7, we use row operations to transform our  $3 \times 3$  matrix  $A$  into an upper triangular matrix, keeping track of the row operations, and labeling each successive matrix. In what follows we follow the Gauss Jordan algorithm without worrying about getting leading 1's.

<sup>2</sup>For a very elegant treatment, take a course in Linear Algebra. There, you will most likely see the treatment of determinants logically reversed than what is presented here. Specifically, the determinant is defined as a function which takes a square matrix to a real number and satisfies some of the properties in Theorem 1.7. From that function, a formula for the determinant is developed.

$$\begin{array}{ccc}
 \begin{bmatrix} 3 & 1 & 2 \\ 0 & -1 & 5 \\ 2 & 1 & 4 \end{bmatrix} & \xrightarrow[\text{with } -\frac{2}{3}R1 + R3]{\text{Replace } R3} & \begin{bmatrix} 3 & 1 & 2 \\ 0 & -1 & 5 \\ 0 & \frac{1}{3} & \frac{8}{3} \end{bmatrix} & \xrightarrow[\frac{1}{3}R2 + R3]{\text{Replace } R3 \text{ with}} & \begin{bmatrix} 3 & 1 & 2 \\ 0 & -1 & 5 \\ 0 & 0 & \frac{13}{3} \end{bmatrix} \\
 A & & B & & C
 \end{array}$$

Theorem 1.7 guarantees us that  $\det(A) = \det(B) = \det(C)$  since we are replacing a row with itself plus a multiple of another row moving from one matrix to the next. Furthermore, since  $C$  is upper triangular,  $\det(C)$  is the product of the entries on the main diagonal, in this case  $\det(C) = (3)(-1)\left(\frac{13}{3}\right) = -13$ . This demonstrates the utility of using row operations to assist in calculating determinants.

This also sheds some light on the connection between a determinant and invertibility. Recall from Section 1.4 that in order to find  $A^{-1}$ , we attempt to transform  $A$  to  $I_n$  using row operations

$$[A \mid I_n] \xrightarrow{\text{Gauss Jordan Elimination}} [I_n \mid A^{-1}]$$

As we apply our allowable row operations on  $A$  to put it into reduced row echelon form, the determinant of the intermediate matrices can vary from the determinant of  $A$  by at most a *nonzero* multiple. This means that if  $\det(A) \neq 0$ , then the determinant of  $A$ 's reduced row echelon form must also be nonzero. Per Definition 1.3, this means that all the main diagonal entries on  $A$ 's reduced row echelon form must be 1. Hence,  $A$ 's reduced row echelon form is  $I_n$  so  $A$  is invertible.

Conversely, if  $A$  is invertible, then  $A$  can be transformed into  $I_n$  using row operations. Since  $\det(I_n) = 1 \neq 0$ , our same logic implies  $\det(A) \neq 0$ . Basically, we have established that the determinant *determines* whether or not the matrix  $A$  is invertible.<sup>3</sup>

It is worth noting that when we first introduced the notion of a matrix inverse, it was in the context of solving a linear matrix equation. In effect, we were trying to 'divide' both sides of the matrix equation  $AX = B$  by the matrix  $A$ . Just like we cannot divide a real number by 0, Theorem 1.7 tells us we cannot 'divide' by a matrix whose *determinant* is 0. We also know that if the coefficient matrix of a system of linear equations is invertible, then system is consistent and independent. It follows, then, that if the determinant of said coefficient is not zero, the system is consistent and independent.

<sup>3</sup>In Section 1.5.2, we learn determinants (specifically cofactors) are deeply connected with the inverse of a matrix.



### 1.5.2 Cramer's Rule and Matrix Adjoints

In this section, we introduce a theorem which enables us to solve a system of linear equations by means of determinants only. As usual, the theorem is stated in full generality, using numbered unknowns  $x_1, x_2$ , etc., instead of the more familiar letters  $x, y, z$ , etc. As with many results in this chapter, the proof of the general case is best left to a course in Linear Algebra.

**THEOREM 1.8. Cramer's Rule:** Suppose  $AX = B$  is the matrix form of a system of  $n$  linear equations in  $n$  unknowns where  $A$  is the coefficient matrix,  $X$  is the unknowns matrix, and  $B$  is the constant matrix. If  $\det(A) \neq 0$ , then the corresponding system is consistent and independent and the solution for unknowns  $x_1, x_2, \dots, x_n$  is given by:

$$x_j = \frac{\det(A_j)}{\det(A)},$$

where  $A_j$  is the matrix  $A$  whose  $j$ th column has been replaced by the constants in  $B$ .

In words, Cramer's Rule tells us we can solve for each unknown, one at a time, by finding the ratio of the determinant of  $A_j$  to that of the determinant of the coefficient matrix. The matrix  $A_j$  is found by replacing the column in the coefficient matrix which holds the coefficients of  $x_j$  with the constants of the system. The following example fleshes out this method.

**EXAMPLE 1.5.1.** Use Cramer's Rule to solve for the indicated unknowns.

1. Solve  $\begin{cases} 2x_1 - 3x_2 = 4 \\ 5x_1 + x_2 = -2 \end{cases}$  for  $x_1$  and  $x_2$
2. Solve  $\begin{cases} 2x - 3y + z = -1 \\ x - y + z = 1 \\ 3x - 4z = 0 \end{cases}$  for  $z$ .

**Solution.**

1. Writing this system in matrix form, we find

$$A = \begin{bmatrix} 2 & -3 \\ 5 & 1 \end{bmatrix} \quad X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad B = \begin{bmatrix} 4 \\ -2 \end{bmatrix}$$

To find the matrix  $A_1$ , we remove the column of the coefficient matrix  $A$  which holds the coefficients of  $x_1$  and replace it with the corresponding entries in  $B$ . Likewise, we replace the column of  $A$  which corresponds to the coefficients of  $x_2$  with the constants to form the matrix  $A_2$ . This yields

$$A_1 = \begin{bmatrix} 4 & -3 \\ -2 & 1 \end{bmatrix} \quad A_2 = \begin{bmatrix} 2 & 4 \\ 5 & -2 \end{bmatrix}$$

Computing determinants, we get  $\det(A) = 17$ ,  $\det(A_1) = -2$  and  $\det(A_2) = -24$ , so that

$$x_1 = \frac{\det(A_1)}{\det(A)} = -\frac{2}{17} \quad x_2 = \frac{\det(A_2)}{\det(A)} = -\frac{24}{17}$$

The reader can check that the solution to the system is  $(-\frac{2}{17}, -\frac{24}{17})$ .

2. To use Cramer's Rule to find  $z$ , we identify  $x_3$  as  $z$ . We have

$$A = \begin{bmatrix} 2 & -3 & 1 \\ 1 & -1 & 1 \\ 3 & 0 & -4 \end{bmatrix} \quad X = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad B = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \quad A_3 = A_z = \begin{bmatrix} 2 & -3 & -1 \\ 1 & -1 & 1 \\ 3 & 0 & 0 \end{bmatrix}$$

Expanding both  $\det(A)$  and  $\det(A_z)$  along the third rows (to take advantage of the 0's) gives

$$z = \frac{\det(A_z)}{\det(A)} = \frac{-12}{-10} = \frac{6}{5}$$

The reader is encouraged to solve this system for  $x$  and  $y$  similarly and check the answer.  $\square$

It is worth noting that finding determinants is very 'computationally expensive' meaning they take quite a bit of time and effort to compute for humans and machines alike. For that reason, determinants, and, in particular, Cramer's Rule are used mostly in theoretical contexts as convenient ways to describe and discuss solutions to systems of linear equations.<sup>4</sup>

Our last application of determinants is to develop an alternative method for finding the inverse of a matrix. As with the discussion in Section 1.4 when we developed the first algorithm to find matrix inverses, we ask that you indulge us as we proceed to work to describe a *general* method here via *example*.

Let us consider the  $3 \times 3$  matrix  $A$  which we so extensively studied in Section 1.5.1

$$A = \begin{bmatrix} 3 & 1 & 2 \\ 0 & -1 & 5 \\ 2 & 1 & 4 \end{bmatrix}$$

We found through a variety of methods that  $\det(A) = -13$ . To our surprise and delight, its inverse below has a remarkable number of 13's in the denominators of its entries. This is no coincidence.

$$A^{-1} = \begin{bmatrix} \frac{9}{13} & \frac{2}{13} & -\frac{7}{13} \\ -\frac{10}{13} & -\frac{8}{13} & \frac{15}{13} \\ -\frac{2}{13} & \frac{1}{13} & \frac{3}{13} \end{bmatrix}$$

Recall that to find  $A^{-1}$ , we are essentially solving the matrix equation  $AX = I_3$ , where  $X = [x_{ij}]_{3 \times 3}$  is a  $3 \times 3$  matrix. Because of how matrix multiplication is defined, the first column of  $I_3$  is the product of  $A$  with the first column of  $X$ , the second column of  $I_3$  is the product of  $A$  with the second column of  $X$  and the third column of  $I_3$  is the product of  $A$  with the third column of  $X$ . In other words, we are solving three equations<sup>5</sup>

<sup>4</sup>A classic case of this is the [Wronskian](#) from Differential Equations.

<sup>5</sup>The reader is encouraged to stop and think this through.

$$A \begin{bmatrix} x_{11} \\ x_{21} \\ x_{31} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad A \begin{bmatrix} x_{12} \\ x_{22} \\ x_{32} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad A \begin{bmatrix} x_{13} \\ x_{23} \\ x_{33} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

We can solve each of these systems using Cramer's Rule. Focusing on the first system, we have

$$A_1 = \begin{bmatrix} 1 & 1 & 2 \\ 0 & -1 & 5 \\ 0 & 1 & 4 \end{bmatrix} \quad A_2 = \begin{bmatrix} 3 & 1 & 2 \\ 0 & 0 & 5 \\ 2 & 0 & 4 \end{bmatrix} \quad A_3 = \begin{bmatrix} 3 & 1 & 1 \\ 0 & -1 & 0 \\ 2 & 1 & 0 \end{bmatrix}$$

If we expand  $\det(A_1)$  along the first row, we get

$$\begin{aligned} \det(A_1) &= \det \left( \begin{bmatrix} -1 & 5 \\ 1 & 4 \end{bmatrix} \right) - \det \left( \begin{bmatrix} 0 & 5 \\ 0 & 4 \end{bmatrix} \right) + 2 \det \left( \begin{bmatrix} 0 & -1 \\ 0 & 1 \end{bmatrix} \right) \\ &= \det \left( \begin{bmatrix} -1 & 5 \\ 1 & 4 \end{bmatrix} \right) \end{aligned}$$

Amazingly, this is none other than the  $C_{11}$  cofactor of  $A$ . The reader is invited to check this, as well as the claims that  $\det(A_2) = C_{12}$  and  $\det(A_3) = C_{13}$ .<sup>6</sup> (To see this, though it seems unnatural to do so, expand along the first row.) Cramer's Rule tells us

$$x_{11} = \frac{\det(A_1)}{\det(A)} = \frac{C_{11}}{\det(A)}, \quad x_{21} = \frac{\det(A_2)}{\det(A)} = \frac{C_{12}}{\det(A)}, \quad x_{31} = \frac{\det(A_3)}{\det(A)} = \frac{C_{13}}{\det(A)}$$

So the first column of the inverse matrix  $X$  is:

$$\begin{bmatrix} x_{11} \\ x_{21} \\ x_{31} \end{bmatrix} = \begin{bmatrix} \frac{C_{11}}{\det(A)} \\ \frac{C_{12}}{\det(A)} \\ \frac{C_{13}}{\det(A)} \end{bmatrix} = \frac{1}{\det(A)} \begin{bmatrix} C_{11} \\ C_{12} \\ C_{13} \end{bmatrix}$$

Notice the reversal of the subscripts going from the unknown to the corresponding cofactor of  $A$ . This trend continues and we get

$$\begin{bmatrix} x_{12} \\ x_{22} \\ x_{32} \end{bmatrix} = \frac{1}{\det(A)} \begin{bmatrix} C_{21} \\ C_{22} \\ C_{23} \end{bmatrix} \quad \begin{bmatrix} x_{13} \\ x_{23} \\ x_{33} \end{bmatrix} = \frac{1}{\det(A)} \begin{bmatrix} C_{31} \\ C_{32} \\ C_{33} \end{bmatrix}$$

Putting all of these together, we have obtained a new and surprising formula for  $A^{-1}$ , namely

<sup>6</sup>In a solid Linear Algebra course you will learn that the properties in Theorem 1.7 hold equally well if the word 'row' is replaced by the word 'column'. We're not going to get into column operations in this text, but they do make some of what we're trying to say easier to follow.

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} C_{11} & C_{21} & C_{31} \\ C_{12} & C_{22} & C_{32} \\ C_{13} & C_{23} & C_{33} \end{bmatrix}$$

To see that this does indeed yield  $A^{-1}$ , we find all of the cofactors of  $A$

$$\begin{aligned} C_{11} &= -9, & C_{21} &= -2, & C_{31} &= 7 \\ C_{12} &= 10, & C_{22} &= 8, & C_{32} &= -15 \\ C_{13} &= 2, & C_{23} &= -1, & C_{33} &= -3 \end{aligned}$$

And, as promised,

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} C_{11} & C_{21} & C_{31} \\ C_{12} & C_{22} & C_{32} \\ C_{13} & C_{23} & C_{33} \end{bmatrix} = -\frac{1}{13} \begin{bmatrix} -9 & -2 & 7 \\ 10 & 8 & -15 \\ 2 & -1 & -3 \end{bmatrix} = \begin{bmatrix} \frac{9}{13} & \frac{2}{13} & -\frac{7}{13} \\ -\frac{10}{13} & -\frac{8}{13} & \frac{15}{13} \\ -\frac{2}{13} & \frac{1}{13} & \frac{3}{13} \end{bmatrix}$$

To generalize this to invertible  $n \times n$  matrices, we need another definition and a theorem. Our definition gives a special name to the cofactor matrix, and the theorem gives us a recipe for the matrix inverse.

**DEFINITION 1.14.** Let  $A$  be an  $n \times n$  matrix, and  $C_{ij}$  denote the  $ij$  cofactor of  $A$ . The **adjoint** of  $A$ , denoted  $\text{adj}(A)$  is the matrix whose  $ij$ -entry is the  $ji$  cofactor of  $A$ ,  $C_{ji}$ . That is

$$\text{adj}(A) = \begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix}$$

This new notation greatly shortens the statement of the formula for the inverse of a matrix.

**THEOREM 1.9.** Let  $A$  be an invertible  $n \times n$  matrix. Then

$$A^{-1} = \frac{1}{\det(A)} \text{adj}(A)$$

For  $2 \times 2$  matrices, Theorem 1.9 reduces to a fairly simple formula.

**EQUATION 1.2.** For an invertible  $2 \times 2$  matrix,

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

The proof of Theorem 1.9 is, like so many of the results in this section, best left to a course in Linear Algebra. In such a course, not only do you gain some more sophisticated proof techniques, you also gain a larger perspective. Within the scope of this text, we will prove a few results involving determinants in Section ?? using the Principle of Mathematical Induction. Until then, make sure you have a handle on the *mechanics* of matrices and the theory will come eventually.

## 1.5.3 Exercises

Exercise ideas: revisit fitting curves to three points - function condition means det not zero (?)

Determinant formula for line

Geometry of determinant - wait to vectors (?)

Follow up in inverse determinant section: What about  $ax^m + bx^n + cx^p$  for these three points?

In Exercises 1 - 8, compute the determinant of the given matrix. (Some of these matrices appeared in Exercises 1 - 8 in Section 1.4.)

$$1. B = \begin{bmatrix} 12 & -7 \\ -5 & 3 \end{bmatrix}$$

$$2. C = \begin{bmatrix} 6 & 15 \\ 14 & 35 \end{bmatrix}$$

$$3. Q = \begin{bmatrix} x & x^2 \\ 1 & 2x \end{bmatrix}$$

$$4. L = \begin{bmatrix} \frac{1}{x^3} & \frac{\ln(x)}{x^3} \\ -\frac{3}{x^4} & \frac{1 - 3\ln(x)}{x^4} \end{bmatrix}$$

$$5. F = \begin{bmatrix} 4 & 6 & -3 \\ 3 & 4 & -3 \\ 1 & 2 & 6 \end{bmatrix}$$

$$6. G = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 11 \\ 3 & 4 & 19 \end{bmatrix}$$

$$7. V = \begin{bmatrix} i & j & k \\ -1 & 0 & 5 \\ 9 & -4 & -2 \end{bmatrix}$$

$$8. H = \begin{bmatrix} 1 & 0 & -3 & 0 \\ 2 & -2 & 8 & 7 \\ -5 & 0 & 16 & 0 \\ 1 & 0 & 4 & 1 \end{bmatrix}$$

In Exercises 9 - 14, use Cramer's Rule to solve the system of linear equations.

$$9. \begin{cases} 3x + 7y = 26 \\ 5x + 12y = 39 \end{cases}$$

$$10. \begin{cases} 2x - 4y = 5 \\ 10x + 13y = -6 \end{cases}$$

$$11. \begin{cases} x + y = 8000 \\ 0.03x + 0.05y = 250 \end{cases}$$

$$12. \begin{cases} \frac{1}{2}x - \frac{1}{5}y = 1 \\ 6x + 7y = 3 \end{cases}$$

$$13. \begin{cases} x + y + z = 3 \\ 2x - y + z = 0 \\ -3x + 5y + 7z = 7 \end{cases}$$

$$14. \begin{cases} 3x + y - 2z = 10 \\ 4x - y + z = 5 \\ x - 3y - 4z = -1 \end{cases}$$

In Exercises 15 - 16, use Cramer's Rule to solve for  $x_4$ .

$$15. \begin{cases} x_1 - x_3 = -2 \\ 2x_2 - x_4 = 0 \\ x_1 - 2x_2 + x_3 = 0 \\ -x_3 + x_4 = 1 \end{cases}$$

$$16. \begin{cases} 4x_1 + x_2 = 4 \\ x_2 - 3x_3 = 1 \\ 10x_1 + x_3 + x_4 = 0 \\ -x_2 + x_3 = -3 \end{cases}$$

In Exercises 17 - 18, find the inverse of the given matrix using their determinants and adjoints.

17.  $B = \begin{bmatrix} 12 & -7 \\ -5 & 3 \end{bmatrix}$

18.  $F = \begin{bmatrix} 4 & 6 & -3 \\ 3 & 4 & -3 \\ 1 & 2 & 6 \end{bmatrix}$

19. Carl's Sasquatch Attack! Game Card Collection is a mixture of common and rare cards. Each common card is worth \$0.25 while each rare card is worth \$0.75. If his entire 117 card collection is worth \$48.75, how many of each kind of card does he own?
20. How much of a 5 gallon 40% salt solution should be replaced with pure water to obtain 5 gallons of a 15% solution?
21. How much of a 10 liter 30% acid solution must be replaced with pure acid to obtain 10 liters of a 50% solution?
22. Daniel's Exotic Animal Rescue houses snakes, tarantulas and scorpions. When asked how many animals of each kind he boards, Daniel answered: 'We board 49 total animals, and I am responsible for each of their 272 legs and 28 tails.' How many of each animal does the Rescue board? (Recall: tarantulas have 8 legs and no tails, scorpions have 8 legs and one tail, and snakes have no legs and one tail.)
23. This exercise is a continuation of Exercise 16 in Section 1.4. Just because a system is consistent independent doesn't mean it will admit a solution that makes sense in an applied setting. Using the nutrient values given for Ippizuti Fish, Misty Mushrooms, and Sun Berries, use Cramer's Rule to determine the number of servings of Ippizuti Fish needed to meet the needs of a daily diet which requires 2500 calories, 1000 grams of protein, and 400 milligrams of Vitamin X. Now use Cramer's Rule to find the number of servings of Misty Mushrooms required. Does a solution to this diet problem exist?

24. Let  $R = \begin{bmatrix} -7 & 3 \\ 11 & 2 \end{bmatrix}$ ,  $S = \begin{bmatrix} 1 & -5 \\ 6 & 9 \end{bmatrix}$ ,  $T = \begin{bmatrix} 11 & 2 \\ -7 & 3 \end{bmatrix}$ , and  $U = \begin{bmatrix} -3 & 15 \\ 6 & 9 \end{bmatrix}$

(a) Show that  $\det(RS) = \det(R)\det(S)$

(b) Show that  $\det(T) = -\det(R)$

(c) Show that  $\det(U) = -3\det(S)$

25. For  $M$ ,  $N$ , and  $P$  below, show that  $\det(M) = 0$ ,  $\det(N) = 0$  and  $\det(P) = 0$ .

$$M = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 7 & 8 & 9 \end{bmatrix}, \quad N = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \quad P = \begin{bmatrix} 1 & 2 & 3 \\ -2 & -4 & -6 \\ 7 & 8 & 9 \end{bmatrix}$$

26. Let  $A$  be an arbitrary invertible  $3 \times 3$  matrix.

- (a) Show that  $\det(I_3) = 1$ . (See footnote<sup>7</sup> below.)  
 (b) Using the facts that  $AA^{-1} = I_3$  and  $\det(AA^{-1}) = \det(A)\det(A^{-1})$ , show that

$$\det(A^{-1}) = \frac{1}{\det(A)}$$

The purpose of Exercises 27 - 30 is to introduce you to the eigenvalues and eigenvectors of a matrix.<sup>8</sup> We begin with an example using a  $2 \times 2$  matrix and then guide you through some exercises using a  $3 \times 3$  matrix. Consider the matrix

$$C = \begin{bmatrix} 6 & 15 \\ 14 & 35 \end{bmatrix}$$

from Exercise 2. We know that  $\det(C) = 0$  which means that  $CX = 0_{2 \times 2}$  does not have a unique solution. So there is a nonzero matrix  $Y$  with  $CY = 0_{2 \times 2}$ . In fact, every matrix of the form

$$Y = \begin{bmatrix} -\frac{5}{2}t \\ t \end{bmatrix}$$

is a solution to  $CX = 0_{2 \times 2}$ , so there are infinitely many matrices such that  $CX = 0_{2 \times 2}$ . But consider the matrix

$$X_{41} = \begin{bmatrix} 3 \\ 7 \end{bmatrix}$$

It is NOT a solution to  $CX = 0_{2 \times 2}$ , but rather,

$$CX_{41} = \begin{bmatrix} 6 & 15 \\ 14 & 35 \end{bmatrix} \begin{bmatrix} 3 \\ 7 \end{bmatrix} = \begin{bmatrix} 123 \\ 287 \end{bmatrix} = 41 \begin{bmatrix} 3 \\ 7 \end{bmatrix}$$

In fact, if  $Z$  is of the form

$$Z = \begin{bmatrix} \frac{3}{7}t \\ t \end{bmatrix}$$

then

$$CZ = \begin{bmatrix} 6 & 15 \\ 14 & 35 \end{bmatrix} \begin{bmatrix} \frac{3}{7}t \\ t \end{bmatrix} = \begin{bmatrix} \frac{123}{7}t \\ 41t \end{bmatrix} = 41 \begin{bmatrix} \frac{3}{7}t \\ t \end{bmatrix} = 41Z$$

for all  $t$ . The big question is "How did we know to use 41?"

We need a number  $\lambda$  such that  $CX = \lambda X$  has nonzero solutions. We have demonstrated that  $\lambda = 0$  and  $\lambda = 41$  both worked. Are there others? If we look at the matrix equation more closely, what we *really*

<sup>7</sup>If you think about it for just a moment, you'll see that  $\det(I_n) = 1$  for any natural number  $n$ . The formal proof of this fact requires the Principle of Mathematical Induction (Section ??) so we'll stick with  $n = 3$  for the time being.

<sup>8</sup>This material is usually given its own chapter in a Linear Algebra book so clearly we're not able to tell you everything you need to know about eigenvalues and eigenvectors. They are a nice application of determinants, though, so we're going to give you enough background so that you can start playing around with them.

wanted was a nonzero solution to  $(C - \lambda I_2)X = 0_{2 \times 2}$  which we know exists if and only if the determinant of  $C - \lambda I_2$  is zero.<sup>9</sup> So we computed

$$\det(C - \lambda I_2) = \det \left( \begin{bmatrix} 6 - \lambda & 15 \\ 14 & 35 - \lambda \end{bmatrix} \right) = (6 - \lambda)(35 - \lambda) - 14 \cdot 15 = \lambda^2 - 41\lambda$$

This is called the **characteristic polynomial** of the matrix  $C$  and it has two zeros:  $\lambda = 0$  and  $\lambda = 41$ . That's how we knew to use 41 in our work above. The fact that  $\lambda = 0$  showed up as one of the zeros of the characteristic polynomial just means that  $C$  itself had determinant zero which we already knew. Those two numbers are called the **eigenvalues** of  $C$ . The corresponding matrix solutions to  $CX = \lambda X$  are called the **eigenvectors** of  $C$  and the 'vector' portion of the name will make more sense after you've studied vectors. Now it's your turn. In the following exercises, you'll be using the matrix  $G$  from Exercise 6.

$$G = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 11 \\ 3 & 4 & 19 \end{bmatrix}$$

27. Show that the characteristic polynomial of  $G$  is  $p(\lambda) = -\lambda(\lambda - 1)(\lambda - 22)$ . That is, compute  $\det(G - \lambda I_3)$ .
28. Let  $G_0 = G$ . Find the parametric description of the solution to the system of linear equations given by  $GX = 0_{3 \times 3}$ .
29. Let  $G_1 = G - I_3$ . Find the parametric description of the solution to the system of linear equations given by  $G_1X = 0_{3 \times 3}$ . Show that any solution to  $G_1X = 0_{3 \times 3}$  also has the property that  $GX = 1X$ .
30. Let  $G_{22} = G - 22I_3$ . Find the parametric description of the solution to the system of linear equations given by  $G_{22}X = 0_{3 \times 3}$ . Show that any solution to  $G_{22}X = 0_{3 \times 3}$  also has the property that  $GX = 22X$ .

---

<sup>9</sup>Think about this.



**1.5.4 Answers**

1.  $\det(B) = 1$

2.  $\det(C) = 0$

3.  $\det(Q) = x^2$

4.  $\det(L) = \frac{1}{x^7}$

5.  $\det(F) = -12$

6.  $\det(G) = 0$

7.  $\det(V) = 20i + 43j + 4k$

8.  $\det(H) = -2$

9.  $x = 39, y = -13$

10.  $x = \frac{41}{66}, y = -\frac{31}{33}$

11.  $x = 7500, y = 500$

12.  $x = \frac{76}{47}, y = -\frac{45}{47}$

13.  $x = 1, y = 2, z = 0$

14.  $x = \frac{121}{60}, y = \frac{131}{60}, z = -\frac{53}{60}$

15.  $x_4 = 4$

16.  $x_4 = -1$

17.  $B^{-1} = \begin{bmatrix} 3 & 7 \\ 5 & 12 \end{bmatrix}$

18.  $F^{-1} = \begin{bmatrix} -\frac{5}{2} & \frac{7}{2} & \frac{1}{2} \\ \frac{7}{4} & -\frac{9}{4} & -\frac{1}{4} \\ -\frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{bmatrix}$

19. Carl owns 78 common cards and 39 rare cards.

20. 3.125 gallons.

21.  $\frac{20}{7} \approx 2.85$  liters.

22. The rescue houses 15 snakes, 21 tarantulas and 13 scorpions.

23. Using Cramer's Rule, we find we need 53 servings of Ippizuti Fish to satisfy the dietary requirements. The number of servings of Misty Mushrooms required, however, is  $-1120$ . Since it's impossible to have a negative number of servings, there is no solution to the applied problem, despite there being a solution to the mathematical problem. A cautionary tale about using Cramer's Rule: just because you are guaranteed a mathematical answer for each variable doesn't mean the solution will make sense in the 'real' world.

## 1.6 Partial Fraction Decomposition

This section uses systems of linear equations to rewrite rational functions in a form more palatable to Calculus students. In College Algebra, the function

$$f(x) = \frac{x^2 - x - 6}{x^4 + x^2} \quad (1)$$

is written in the best form possible to construct a sign diagram and to find zeros and asymptotes, but certain applications in Calculus require us to rewrite  $f(x)$  as

$$f(x) = \frac{x+7}{x^2+1} - \frac{1}{x} - \frac{6}{x^2} \quad (2)$$

If we are given the form of  $f(x)$  in (1), it is a matter of Intermediate Algebra to determine a common denominator to obtain the form of  $f(x)$  given in (1). The focus of this section is to develop a method by which we start with  $f(x)$  in the form of (1) and ‘resolve it into *partial fractions*’ to obtain the form in (2). Essentially, we need to reverse the least common denominator process. Starting with the form of  $f(x)$  in (1), we begin by factoring the denominator

$$\frac{x^2 - x - 6}{x^4 + x^2} = \frac{x^2 - x - 6}{x^2(x^2 + 1)}$$

We now think about which individual denominators could contribute to obtain  $x^2(x^2 + 1)$  as the least common denominator. Certainly  $x^2$  and  $x^2 + 1$ , but are there any other factors? Since  $x^2 + 1$  is an irreducible quadratic<sup>1</sup> there are no factors of it that have real coefficients which can contribute to the denominator.

The factor  $x^2$ , however, is not irreducible, since we can think of it as  $x^2 = xx = (x - 0)(x - 0)$ , a so-called ‘repeated’ linear factor.<sup>2</sup> This means it’s possible that a term with a denominator of just  $x$  contributed to the expression as well. What about something like  $x(x^2 + 1)$ ? This, too, could contribute, but we would then wish to break down that denominator into  $x$  and  $(x^2 + 1)$ , so we leave out a term of that form.

At this stage, we have guessed

$$\frac{x^2 - x - 6}{x^4 + x^2} = \frac{x^2 - x - 6}{x^2(x^2 + 1)} = \frac{?}{x} + \frac{?}{x^2} + \frac{?}{x^2 + 1}$$

Our next task is to determine what form the unknown numerators take. It stands to reason that since the expression  $\frac{x^2 - x - 6}{x^4 + x^2}$  is ‘proper’ in the sense that the degree of the numerator is less than the degree of the denominator, we are safe to make the [ansatz](#) that all of the partial fraction resolvents are also. This means that the numerator of the fraction with  $x$  as its denominator is just a constant and the numerators on the terms involving the denominators  $x^2$  and  $x^2 + 1$  are at most linear polynomials.

In other words, we guess that there are real numbers  $A$ ,  $B$ ,  $C$ ,  $D$  and  $E$  so that

$$\frac{x^2 - x - 6}{x^4 + x^2} = \frac{x^2 - x - 6}{x^2(x^2 + 1)} = \frac{A}{x} + \frac{Bx + C}{x^2} + \frac{Dx + E}{x^2 + 1}$$

<sup>1</sup>Recall this means it has no real zeros; see Section ??.

<sup>2</sup>Recall this means  $x = 0$  is a zero of multiplicity 2.

However, if we look more closely at the term  $\frac{Bx+C}{x^2}$ , we see that  $\frac{Bx+C}{x^2} = \frac{Bx}{x^2} + \frac{C}{x^2} = \frac{B}{x} + \frac{C}{x^2}$ . The term  $\frac{B}{x}$  has the same form as the term  $\frac{A}{x}$  which means it contributes nothing new to our expansion. Hence, we drop it and, after re-labeling, we find ourselves with our new guess:

$$\frac{x^2 - x - 6}{x^4 + x^2} = \frac{x^2 - x - 6}{x^2(x^2 + 1)} = \frac{A}{x} + \frac{B}{x^2} + \frac{Cx + D}{x^2 + 1}$$

Our next task is to determine the values of our unknowns. Clearing denominators gives

$$x^2 - x - 6 = Ax(x^2 + 1) + B(x^2 + 1) + (Cx + D)x^2$$

Gathering the like powers of  $x$  we have

$$x^2 - x - 6 = (A + C)x^3 + (B + D)x^2 + Ax + B$$

In order for this to hold for all values of  $x$  in the domain of  $f$ , we equate the coefficients of corresponding powers of  $x$  on each side of the equation<sup>3</sup> and obtain the system of linear equations

$$\begin{cases} (E1) & A + C & = & 0 & \text{From equating coefficients of } x^3 \\ (E2) & B + D & = & 1 & \text{From equating coefficients of } x^2 \\ (E3) & A & = & -1 & \text{From equating coefficients of } x \\ (E4) & B & = & -6 & \text{From equating the constant terms} \end{cases}$$

To solve this system of equations, we could use any of the methods presented in Sections 1.1 through 1.5, but none of these methods are as efficient as the good old-fashioned substitution from High School algebra. From  $E3$ , we have  $A = -1$  and we substitute this into  $E1$  to get  $C = 1$ . Similarly, since  $E4$  gives us  $B = -6$ , we have from  $E2$  that  $D = 7$ . We get

$$\frac{x^2 - x - 6}{x^4 + x^2} = \frac{x^2 - x - 6}{x^2(x^2 + 1)} = -\frac{1}{x} - \frac{6}{x^2} + \frac{x + 7}{x^2 + 1}$$

which matches the formula given in (2).

As we have seen in this opening example, resolving a rational function into partial fractions takes two steps: first, we need to determine the *form* of the decomposition, and then we need to determine the unknown coefficients which appear in said form.

Theorem ?? guarantees that any polynomial with real coefficients can be factored over the real numbers as a product of linear factors and irreducible quadratic factors. Once we have this factorization of the denominator of a rational function, the next theorem tells us the form the decomposition takes. The reader is encouraged to review the Factor Theorem (Theorem ??) and its connection to the role of multiplicity to fully appreciate the statement of the following theorem.

---

<sup>3</sup>We will justify this shortly.

THEOREM 1.10. Suppose  $R(x) = \frac{N(x)}{D(x)}$  is a rational function where the degree of  $N(x)$  less than the degree of  $D(x)$  and  $N(x)$  and  $D(x)$  have no common factors.<sup>a</sup>

- If  $\alpha$  is a real zero of  $D$  of multiplicity  $m$  which corresponds to the linear factor  $ax + b$ , the partial fraction decomposition includes

$$\frac{A_1}{ax + b} + \frac{A_2}{(ax + b)^2} + \dots + \frac{A_m}{(ax + b)^m}$$

for real numbers  $A_1, A_2, \dots, A_m$ .

- If  $\alpha$  is a non-real zero of  $D$  of multiplicity  $m$  which corresponds to the irreducible quadratic  $ax^2 + bx + c$ , the partial fraction decomposition includes

$$\frac{B_1x + C_1}{ax^2 + bx + c} + \frac{B_2x + C_2}{(ax^2 + bx + c)^2} + \dots + \frac{B_mx + C_m}{(ax^2 + bx + c)^m}$$

for real numbers  $B_1, B_2, \dots, B_m$  and  $C_1, C_2, \dots, C_m$ .

<sup>a</sup>In other words,  $R(x)$  is a proper rational function which has been fully reduced.

The proof of Theorem 1.10 is best left to a course in Abstract Algebra. Notice that the theorem provides for the general case, so we need to use subscripts,  $A_1, A_2$ , etc., to denote different unknown coefficients as opposed to the usual convention of  $A, B$ , etc.. The stress on multiplicities is to help us correctly group factors in the denominator. For example, consider the rational function

$$\frac{3x - 1}{(x^2 - 1)(2 - x - x^2)}$$

Factoring the denominator to find the zeros, we get  $(x + 1)(x - 1)(1 - x)(2 + x)$ . We find  $x = -1$  and  $x = -2$  are zeros of multiplicity one but that  $x = 1$  is a zero of multiplicity two due to the two different factors  $(x - 1)$  and  $(1 - x)$ . One way to handle this is to note that  $(1 - x) = -(x - 1)$  so

$$\frac{3x - 1}{(x + 1)(x - 1)(1 - x)(2 + x)} = \frac{3x - 1}{-(x - 1)^2(x + 1)(x + 2)} = \frac{1 - 3x}{(x - 1)^2(x + 1)(x + 2)}$$

from which we proceed with the partial fraction decomposition

$$\frac{1 - 3x}{(x - 1)^2(x + 1)(x + 2)} = \frac{A}{x - 1} + \frac{B}{(x - 1)^2} + \frac{C}{x + 1} + \frac{D}{x + 2}$$

Turning our attention to non-real zeros, we note that the tool of choice to determine the irreducibility of a quadratic  $ax^2 + bx + c$  is the discriminant,  $b^2 - 4ac$ . If  $b^2 - 4ac < 0$ , the quadratic admits a *pair* of non-real complex conjugate zeros. Even though *one* irreducible quadratic gives *two* distinct non-real zeros, we list the terms with denominators involving a given irreducible quadratic only once to avoid duplication in the form of the decomposition. The trick, of course, is factoring the denominator or otherwise finding the

zeros and their multiplicities in order to apply Theorem 1.10. We recommend that the reader review the techniques set forth in Sections ?? and ??.

Next, we state a theorem that if two polynomials are equal, the corresponding coefficients of the like powers of  $x$  are equal. This is the principal by which we shall determine the unknown coefficients in our partial fraction decomposition.

**THEOREM 1.11.** Suppose

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_2 x^2 + a_1 x + a_0 = b_m x^m + m_{m-1} x^{m-1} + \cdots + b_2 x^2 + b_1 x + b_0$$

for all  $x$  in an open interval  $I$ . Then  $n = m$  and  $a_i = b_i$  for all  $i = 1 \dots n$ .

Believe it or not, the proof of Theorem 1.11 is a consequence of Theorem ??. Define  $p(x)$  to be the difference of the left hand side of the equation in Theorem 1.11 and the right hand side. Then  $p(x) = 0$  for all  $x$  in the open interval  $I$ . If  $p(x)$  were a nonzero polynomial of degree  $k$ , then, by Theorem ??,  $p$  could have at most  $k$  zeros in  $I$ ,  $k$  being a *finite* number. Since  $p(x) = 0$  for *all* real numbers  $x$  in  $I$ ,  $p$  has infinitely many zeros, and hence,  $p$  is the zero polynomial. This means there can be no nonzero terms in  $p(x)$  and the theorem follows. Arguably, the best way to make sense of either of the two preceding theorems is to work some examples.

**EXAMPLE 1.6.1.** Resolve the following rational functions into partial fractions.

1.  $R(x) = \frac{x+5}{2x^2-x-1}$

2.  $f(t) = \frac{3}{t^3-2t^2+t}$

3.  $F(s) = \frac{3}{s^3-s^2+s}$

4.  $r(x) = \frac{4x^3}{x^2-2}$

5.  $G(t) = \frac{t^3+5t-1}{t^4+6t^2+9}$

6.  $H(s) = \frac{8s^2}{s^4+16}$

**Solution.**

1. We begin by factoring the denominator to find  $2x^2 - x - 1 = (2x + 1)(x - 1)$ . We get  $x = -\frac{1}{2}$  and  $x = 1$  are both zeros of multiplicity one and thus we know

$$\frac{x+5}{2x^2-x-1} = \frac{x+5}{(2x+1)(x-1)} = \frac{A}{2x+1} + \frac{B}{x-1}$$

Clearing denominators, we get  $x+5 = A(x-1) + B(2x+1)$  so that  $x+5 = (A+2B)x + B - A$ . Equating coefficients, we get the system

$$\begin{cases} A+2B = 1 \\ -A+B = 5 \end{cases}$$

This system is readily handled using the Addition Method from Section ??, and after adding both equations, we get  $3B = 6$  so  $B = 2$ . Using back substitution, we find  $A = -3$ . Our answer is easily checked by getting a common denominator and adding the fractions.

$$\frac{x+5}{2x^2-x-1} = \frac{2}{x-1} - \frac{3}{2x+1}$$

2. Factoring the denominator gives  $t^3 - 2t^2 + t = t(t^2 - 2t + 1) = t(t-1)^2$  which gives  $t = 0$  as a zero of multiplicity one and  $t = 1$  as a zero of multiplicity two. We have

$$\frac{3}{t^3 - 2t^2 + t} = \frac{3}{t(t-1)^2} = \frac{A}{t} + \frac{B}{t-1} + \frac{C}{(t-1)^2}$$

Clearing denominators, we get  $3 = A(t-1)^2 + Bt(t-1) + Ct$ , which, after gathering up the like terms becomes  $3 = (A+B)t^2 + (-2A-B+C)t + A$ . Our system is

$$\begin{cases} A+B &= 0 \\ -2A-B+C &= 0 \\ A &= 3 \end{cases}$$

Substituting  $A = 3$  into  $A+B = 0$  gives  $B = -3$ , and substituting both for  $A$  and  $B$  in  $-2A-B+C = 0$  gives  $C = 3$ . Our final answer is

$$\frac{3}{t^3 - 2t^2 + t} = \frac{3}{t} - \frac{3}{t-1} + \frac{3}{(t-1)^2}$$

3. The denominator factors as  $s(s^2 - s + 1)$ . We see immediately that  $s = 0$  is a zero of multiplicity one, but the zeros of  $s^2 - s + 1$  aren't as easy to discern. The quadratic doesn't factor easily, so we check the discriminant and find it to be  $(-1)^2 - 4(1)(1) = -3 < 0$ . We find its zeros are not real so it is an irreducible quadratic. The form of the partial fraction decomposition is then

$$\frac{3}{s^3 - s^2 + s} = \frac{3}{s(s^2 - s + 1)} = \frac{A}{s} + \frac{Bs + C}{s^2 - s + 1}$$

Clearing denominators gives  $3 = A(s^2 - s + 1) + (Bs + C)s$  or  $3 = (A+B)s^2 + (-A+C)s + A$ , hence

$$\begin{cases} A+B &= 0 \\ -A+C &= 0 \\ A &= 3 \end{cases}$$

From  $A = 3$  and  $A+B = 0$ , we get  $B = -3$ . From  $-A+C = 0$ , we get  $C = A = 3$ . We get

$$\frac{3}{s^3 - s^2 + s} = \frac{3}{s} + \frac{3-3s}{s^2 - s + 1}$$

4. Since  $\frac{4x^3}{x^2-2}$  isn't proper, we first use long division and obtain a quotient of  $4x$  with a remainder of  $8x$ . Rewriting,  $\frac{4x^3}{x^2-2} = 4x + \frac{8x}{x^2-2}$  so we focus on resolving  $\frac{8x}{x^2-2}$  into partial fractions. The quadratic  $x^2 - 2$ , though it doesn't factor nicely, is, nevertheless, reducible. Solving  $x^2 - 2 = 0$  gives us  $x = \pm\sqrt{2}$ , so using Theorem ??,<sup>4</sup> we have  $x^2 - 2 = (x - \sqrt{2})(x + \sqrt{2})$ . Hence,

<sup>4</sup>Alternatively, we can recognize  $x^2 - 2 = x^2 - (\sqrt{2})^2$  and use the Difference of Squares formula on page ??.

$$\frac{8x}{x^2 - 2} = \frac{8x}{(x - \sqrt{2})(x + \sqrt{2})} = \frac{A}{x - \sqrt{2}} + \frac{B}{x + \sqrt{2}}$$

Clearing fractions, we get  $8x = A(x + \sqrt{2}) + B(x - \sqrt{2})$  or  $8x = (A + B)x + (A - B)\sqrt{2}$  which gives

$$\begin{cases} A + B = 8 \\ (A - B)\sqrt{2} = 0 \end{cases}$$

From  $(A - B)\sqrt{2} = 0$ , we get  $A = B$ , which, when substituted into  $A + B = 8$  gives  $B = 4$ . Hence,  $A = B = 4$  and we get

$$\frac{4x^3}{x^2 - 2} = 4x + \frac{8x}{x^2 - 2} = 4x + \frac{4}{x + \sqrt{2}} + \frac{4}{x - \sqrt{2}}$$

5. At first glance, the denominator  $D(t) = t^4 + 6t^2 + 9$  appears irreducible. However,  $D(t)$  has three terms, and the exponent on the first term is exactly twice that of the second. Rewriting  $D(t) = (t^2)^2 + 6t^2 + 9$ , we see it is a quadratic in disguise and factor  $D(t) = (t^2 + 3)^2$ . Since  $t^2 + 3$  clearly has no real zeros, it is irreducible and the form of the decomposition is

$$\frac{t^3 + 5t - 1}{t^4 + 6t^2 + 9} = \frac{t^3 + 5t - 1}{(t^2 + 3)^2} = \frac{At + B}{t^2 + 3} + \frac{Ct + D}{(t^2 + 3)^2}$$

After the usual clearing of denominators, we have  $t^3 + 5t - 1 = (At + B)(t^2 + 3) + Ct + D$  which gives  $t^3 + 5t - 1 = At^3 + Bt^2 + (3A + C)t + 3B + D$ . Our system is

$$\begin{cases} A = 1 \\ B = 0 \\ 3A + C = 5 \\ 3B + D = -1 \end{cases}$$

We have  $A = 1$  and  $B = 0$  from which we get  $C = 2$  and  $D = -1$ . Our final answer is

$$\frac{t^3 + 5t - 1}{t^4 + 6t^2 + 9} = \frac{t}{t^2 + 3} + \frac{2t - 1}{(t^2 + 3)^2}$$

6. Once again, the difficulty in our last example is factoring the denominator. In an attempt to get a quadratic in disguise, we write

$$s^4 + 16 = (s^2)^2 + 4^2 = (s^2)^2 + 8s^2 + 4^2 - 8s^2 = (s^2 + 4)^2 - 8s^2$$

and obtain a difference of two squares:  $(s^2 + 4)^2$  and  $8s^2 = (2s\sqrt{2})^2$ . Hence,

$$s^4 + 16 = (s^2 + 4 - 2s\sqrt{2})(s^2 + 4 + 2s\sqrt{2}) = (s^2 - 2s\sqrt{2} + 4)(s^2 + 2s\sqrt{2} + 4)$$

The discriminant of both of these quadratics works out to be  $-8 < 0$ , which means they are irreducible. We leave it to the reader to verify that, despite having the same discriminant, these quadratics have different zeros. The partial fraction decomposition takes the form

$$\frac{8s^2}{s^4 + 16} = \frac{8s^2}{(s^2 - 2s\sqrt{2} + 4)(s^2 + 2s\sqrt{2} + 4)} = \frac{As + B}{s^2 - 2s\sqrt{2} + 4} + \frac{Cs + D}{s^2 + 2s\sqrt{2} + 4}$$

We get  $8s^2 = (As + B)(s^2 + 2s\sqrt{2} + 4) + (Cs + D)(s^2 - 2s\sqrt{2} + 4)$  or

$$8s^2 = (A + C)s^3 + (2A\sqrt{2} + B - 2C\sqrt{2} + D)s^2 + (4A + 2B\sqrt{2} + 4C - 2D\sqrt{2})s + 4B + 4D$$

which gives the system

$$\begin{cases} A + C = 0 \\ 2A\sqrt{2} + B - 2C\sqrt{2} + D = 8 \\ 4A + 2B\sqrt{2} + 4C - 2D\sqrt{2} = 0 \\ 4B + 4D = 0 \end{cases}$$

From  $A + C = 0$ , we get  $A = -C$ . Likewise, from  $4B + 4D = 0$ , we get  $B = -D$ . Substituting these into the remaining two equations gives

$$\begin{cases} -2C\sqrt{2} - D - 2C\sqrt{2} + D = 8 \\ -4C - 2D\sqrt{2} + 4C - 2D\sqrt{2} = 0 \end{cases}$$

or

$$\begin{cases} -4C\sqrt{2} = 8 \\ -4D\sqrt{2} = 0 \end{cases}$$

We get  $C = -\sqrt{2}$  so that  $A = -C = \sqrt{2}$  and  $D = 0$  which means  $B = -D = 0$ . We get

$$\frac{8s^2}{s^4 + 16} = \frac{s\sqrt{2}}{s^2 - 2s\sqrt{2} + 4} - \frac{s\sqrt{2}}{s^2 + 2s\sqrt{2} + 4}$$

□



## 1.6.1 Exercises

When we get to the exercises, have one where students split one over the complex numbers and refer back to the claim in 3.1 about rational functions being sums of shifted Laurent monomials.

In Exercises 1 - 6, find only the *form* needed to begin the process of partial fraction decomposition. Do not create the system of linear equations or attempt to find the actual decomposition.

1.  $\frac{7}{(x-3)(x+5)}$

2.  $\frac{5x+4}{x(x-2)(2-x)}$

3.  $\frac{m}{(7x-6)(x^2+9)}$

4.  $\frac{ax^2+bx+c}{x^3(5x+9)(3x^2+7x+9)}$

5.  $\frac{\text{A polynomial of degree } < 9}{(x+4)^5(x^2+1)^2}$

6.  $\frac{\text{A polynomial of degree } < 7}{x(4x-1)^2(x^2+5)(9x^2+16)}$

In Exercises 7 - 18, find the partial fraction decomposition of the following rational expressions.

7.  $\frac{2x}{x^2-1}$

8.  $\frac{-7x+43}{3x^2+19x-14}$

9.  $\frac{11x^2-5x-10}{5x^3-5x^2}$

10.  $\frac{-2x^2+20x-68}{x^3+4x^2+4x+16}$

11.  $\frac{-x^2+15}{4x^4+40x^2+36}$

12.  $\frac{-21x^2+x-16}{3x^3+4x^2-3x+2}$

13.  $\frac{5x^4-34x^3+70x^2-33x-19}{(x-3)^2}$

14.  $\frac{x^6+5x^5+16x^4+80x^3-2x^2+6x-43}{x^3+5x^2+16x+80}$

15.  $\frac{-7x^2-76x-208}{x^3+18x^2+108x+216}$

16.  $\frac{-10x^4+x^3-19x^2+x-10}{x^5+2x^3+x}$

17.  $\frac{4x^3-9x^2+12x+12}{x^4-4x^3+8x^2-16x+16}$

18.  $\frac{2x^2+3x+14}{(x^2+2x+9)(x^2+x+5)}$

19. As we stated at the beginning of this section, the technique of resolving a rational function into partial fractions is a skill needed for Calculus. However, we hope to have shown you that it is worth doing if, for no other reason, it reinforces a hefty amount of algebra. One of the common algebraic errors the authors find students make is something along the lines of

$$\frac{8}{x^2-9} \neq \frac{8}{x^2} - \frac{8}{9}$$

Think about why if the above were true, this section would have no need to exist.

**1.6.2 Answers**

1.  $\frac{A}{x-3} + \frac{B}{x+5}$

2.  $\frac{A}{x} + \frac{B}{x-2} + \frac{C}{(x-2)^2}$

3.  $\frac{A}{7x-6} + \frac{Bx+C}{x^2+9}$

4.  $\frac{A}{x} + \frac{B}{x^2} + \frac{C}{x^3} + \frac{D}{5x+9} + \frac{Ex+F}{3x^2+7x+9}$

5.  $\frac{A}{x+4} + \frac{B}{(x+4)^2} + \frac{C}{(x+4)^3} + \frac{D}{(x+4)^4} + \frac{E}{(x+4)^5} + \frac{Fx+G}{x^2+1} + \frac{Hx+I}{(x^2+1)^2}$

6.  $\frac{A}{x} + \frac{B}{4x-1} + \frac{C}{(4x-1)^2} + \frac{Dx+E}{x^2+5} + \frac{Fx+G}{9x^2+16}$

7.  $\frac{2x}{x^2-1} = \frac{1}{x+1} + \frac{1}{x-1}$

8.  $\frac{-7x+43}{3x^2+19x-14} = \frac{5}{3x-2} - \frac{4}{x+7}$

9.  $\frac{11x^2-5x-10}{5x^3-5x^2} = \frac{3}{x} + \frac{2}{x^2} - \frac{4}{5(x-1)}$

10.  $\frac{-2x^2+20x-68}{x^3+4x^2+4x+16} = -\frac{9}{x+4} + \frac{7x-8}{x^2+4}$

11.  $\frac{-x^2+15}{4x^4+40x^2+36} = \frac{1}{2(x^2+1)} - \frac{3}{4(x^2+9)}$

12.  $\frac{-21x^2+x-16}{3x^3+4x^2-3x+2} = -\frac{6}{x+2} - \frac{3x+5}{3x^2-2x+1}$

13.  $\frac{5x^4-34x^3+70x^2-33x-19}{(x-3)^2} = 5x^2-4x+1 + \frac{9}{x-3} - \frac{1}{(x-3)^2}$

14.  $\frac{x^6+5x^5+16x^4+80x^3-2x^2+6x-43}{x^3+5x^2+16x+80} = x^3 + \frac{x+1}{x^2+16} - \frac{3}{x+5}$

15.  $\frac{-7x^2-76x-208}{x^3+18x^2+108x+216} = -\frac{7}{x+6} + \frac{8}{(x+6)^2} - \frac{4}{(x+6)^3}$

16.  $\frac{-10x^4+x^3-19x^2+x-10}{x^5+2x^3+x} = -\frac{10}{x} + \frac{1}{x^2+1} + \frac{x}{(x^2+1)^2}$

17.  $\frac{4x^3-9x^2+12x+12}{x^4-4x^3+8x^2-16x+16} = \frac{1}{x-2} + \frac{4}{(x-2)^2} + \frac{3x+1}{x^2+4}$

18.  $\frac{2x^2+3x+14}{(x^2+2x+9)(x^2+x+5)} = \frac{1}{x^2+2x+9} + \frac{1}{x^2+x+5}$

## 1.7 Systems of Non-Linear Equations

In this section, we study systems of non-linear equations. In non-linear equations, we can have variables to powers other than 1, we can have different variables multiplied together, or variable can occur as arguments of exponential and logarithmic functions.

Unlike the systems of linear equations for which we have developed several algorithmic solution techniques, there is no general algorithm to solve systems of non-linear equations. Moreover, all of the usual hazards of non-linear equations like extraneous solutions and domain restrictions are once again present.

Along with the tried and true techniques of substitution and elimination, we shall often need equal parts tenacity and ingenuity to see a problem through to the end. You may find it necessary to review topics throughout the text which pertain to solving equations involving the various functions we have studied thus far. To get the section rolling we begin with a fairly routine example.

EXAMPLE 1.7.1. Solve the following systems of equations. Verify answers algebraically and graphically.

$$1. \begin{cases} x^2 + y^2 = 4 \\ 4x^2 + 9y^2 = 36 \end{cases}$$

$$2. \begin{cases} x^2 + y^2 = 4 \\ 4x^2 - 9y^2 = 36 \end{cases}$$

$$3. \begin{cases} x^2 + y^2 = 4 \\ y - 2x = 0 \end{cases}$$

$$4. \begin{cases} x^2 + y^2 = 4 \\ y - x^2 = 0 \end{cases}$$

SOLUTION:

1. Since both equations contain  $x^2$  and  $y^2$  only, we can use elimination as demonstrated in Section ??:

$$\begin{cases} (E1) & x^2 + y^2 = 4 \\ (E2) & 4x^2 + 9y^2 = 36 \end{cases} \xrightarrow[-4E1 + E2]{\text{Replace } E2 \text{ with}} \begin{cases} (E1) & x^2 + y^2 = 4 \\ (E2) & 5y^2 = 20 \end{cases}$$

From  $5y^2 = 20$ , we get  $y^2 = 4$  or  $y = \pm 2$ . To find the associated  $x$  values, we substitute each value of  $y$  into one of the equations to find the resulting value of  $x$ .

Choosing  $x^2 + y^2 = 4$ , we find that for both  $y = -2$  and  $y = 2$ , we get  $x = 0$ . Our solution is thus  $\{(0, 2), (0, -2)\}$ . To verify these answers algebraically, we would need to show that the pair  $(x, y) = (0, 2)$  and  $(x, y) = (0, -2)$  each satisfy *both* equations. We leave this to the reader.

To check our answer graphically, we sketch both equations and look for their points of intersection.

The graph of  $x^2 + y^2 = 4$  is a circle centered at  $(0, 0)$  with a radius of 2. To graph  $4x^2 + 9y^2 = 36$ , we convert to standard form  $\frac{x^2}{9} + \frac{y^2}{4} = 1$  and recognize it as an ellipse centered at  $(0, 0)$  with a major axis along the  $x$ -axis of length 6 and a minor axis along the  $y$ -axis of length 4.

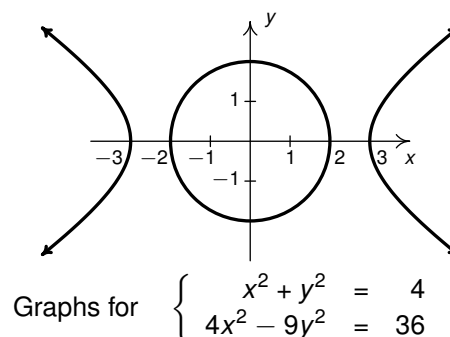
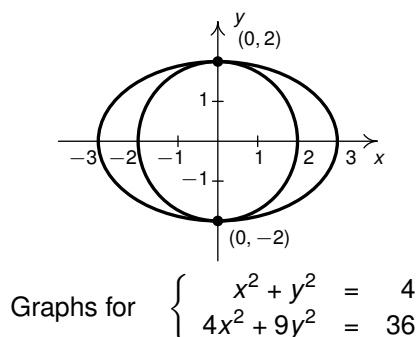
We see from the graph that the two curves intersect at their  $y$ -intercepts only,  $(0, \pm 2)$ .

2. We proceed as before to eliminate one of the variables

$$\begin{cases} (E1) & x^2 + y^2 = 4 \\ (E2) & 4x^2 - 9y^2 = 36 \end{cases} \xrightarrow[-4E1 + E2]{\text{Replace } E2 \text{ with}} \begin{cases} (E1) & x^2 + y^2 = 4 \\ (E2) & -13y^2 = 20 \end{cases}$$

Since the equation  $-13y^2 = 20$  admits no real solution, the system is inconsistent. To verify this graphically, we note that  $x^2 + y^2 = 4$  is the same circle as before, but when writing the second equation in standard form,  $\frac{x^2}{9} - \frac{y^2}{4} = 1$ , we find a hyperbola centered at  $(0, 0)$  opening to the left and right with a transverse axis of length 6 and a conjugate axis of length 4.

We see that the circle and the hyperbola have no points in common, hence, there are no solutions.



3. Since there are no like terms among the two equations, elimination won't work here. Instead, we proceed using substitution.

From the equation  $y - 2x = 0$ , we get  $y = 2x$ . Substituting this into  $x^2 + y^2 = 4$  gives  $x^2 + (2x)^2 = 4$ . Solving, we find  $5x^2 = 4$  or  $x = \pm \frac{2\sqrt{5}}{5}$ .

Returning to the equation we used for the substitution,  $y = 2x$ , we find  $y = \frac{4\sqrt{5}}{5}$  when  $x = \frac{2\sqrt{5}}{5}$ , so one solution is  $\left(\frac{2\sqrt{5}}{5}, \frac{4\sqrt{5}}{5}\right)$  and the other is  $\left(-\frac{2\sqrt{5}}{5}, -\frac{4\sqrt{5}}{5}\right)$ . Hence, our final answer is  $\left\{\left(\frac{2\sqrt{5}}{5}, \frac{4\sqrt{5}}{5}\right), \left(-\frac{2\sqrt{5}}{5}, -\frac{4\sqrt{5}}{5}\right)\right\}$ . As before, we leave the algebraic check to the reader.

The graph of  $x^2 + y^2 = 4$  is our circle from before and the graph of  $y - 2x = 0$ , or  $y = 2x$  is a line through the origin with slope 2. Even though we cannot easily verify the numerical values of the points of intersection from our sketch, we can be sure there are just two solutions: one in Quadrant I and one in Quadrant III. This observation, combined with our (your) algebraic check gives us confidence our solution is correct.<sup>1</sup>

4. While it may be tempting to solve  $y - x^2 = 0$  as  $y = x^2$  and substitute, we note that this system is set up for elimination.<sup>2</sup>

$$\begin{cases} (E1) & x^2 + y^2 = 4 \\ (E2) & y - x^2 = 0 \end{cases} \xrightarrow[\text{E1} + \text{E2}]{\text{Replace E2 with}} \begin{cases} (E1) & x^2 + y^2 = 4 \\ (E2) & y^2 + y = 4 \end{cases}$$

From  $y^2 + y = 4$  we get  $y^2 + y - 4 = 0$  which gives  $y = \frac{-1 \pm \sqrt{17}}{2}$ . Due to the complicated nature of these answers, it is worth our time to make a quick sketch of both equations first to head off any extraneous solutions we may encounter.

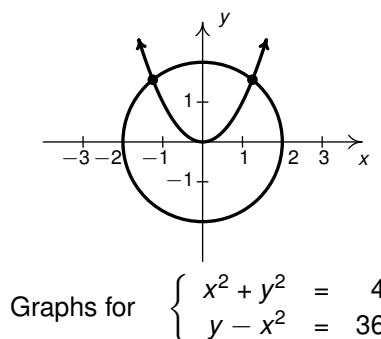
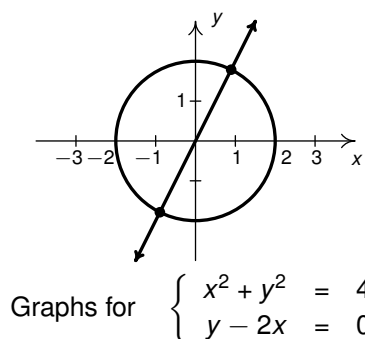
<sup>1</sup>Of course, we could check our answers more accurately using a graphing utility.

<sup>2</sup>We encourage the reader to solve the system using substitution to see that you get the same solution.

We see that the circle  $x^2 + y^2 = 4$  intersects the parabola  $y = x^2$  exactly *twice*, and both of these points have a positive  $y$  value.

Of the two solutions for  $y$ , only  $y = \frac{-1+\sqrt{17}}{2}$  is positive, so to get our solution, we substitute this into  $y - x^2 = 0$  and solve for  $x$ . We get  $x = \pm \sqrt{\frac{-1+\sqrt{17}}{2}} = \pm \frac{\sqrt{-2+2\sqrt{17}}}{2}$ .

Our final answer is  $\left\{ \left( \frac{\sqrt{-2+2\sqrt{17}}}{2}, \frac{-1+\sqrt{17}}{2} \right), \left( -\frac{\sqrt{-2+2\sqrt{17}}}{2}, \frac{-1+\sqrt{17}}{2} \right) \right\}$ . Checking these answers algebraically amounts to a true test of anyone's algebraic mettle and as such is left to the reader.



□

A couple of remarks about Example 1.7.1 are in order. First note that, unlike systems of linear equations, it is possible for a system of non-linear equations to have more than one solution without having infinitely many solutions. In fact, while we characterize systems of nonlinear equations as being 'consistent' or 'inconsistent,' we generally don't use the labels 'dependent' or 'independent'.

Secondly, as we saw with the last problem, sometimes making a quick sketch of the problem situation can save a lot of time and effort. While in general the curves in a system of non-linear equations may not be easily visualized, it pays to take advantage when they are. Our next example provides some considerable review of many of the topics introduced in this text.

EXAMPLE 1.7.2. Solve the following systems of equations. Verify answers algebraically and graphically.

$$1. \begin{cases} x^2 + 2xy - 16 = 0 \\ y^2 + 2xy - 16 = 0 \end{cases} \quad 2. \begin{cases} y + 4e^{2x} = 1 \\ y^2 + 2e^x = 1 \end{cases} \quad 3. \begin{cases} z(x-2) = x \\ yz = y \\ (x-2)^2 + y^2 = 1 \end{cases}$$

**Solution.**

- At first glance, it doesn't appear as though elimination will do us any good since it's clear that we cannot completely eliminate one of the variables. The alternative, however, namely solving one of the equations for one variable and substituting into the other, is very intimidating.

Returning to elimination, we note that it is possible to eliminate the troublesome  $xy$  term, and the constant term as well. Doing so we get a more tractable relationship between  $x$  and  $y$ :

$$\left\{ \begin{array}{l} (E1) \quad x^2 + 2xy - 16 = 0 \\ (E2) \quad y^2 + 2xy - 16 = 0 \end{array} \right. \xrightarrow[\text{Replace } E2 \text{ with}]{-E1 + E2} \left\{ \begin{array}{l} (E1) \quad x^2 + 2xy - 16 = 0 \\ (E2) \quad y^2 - x^2 = 0 \end{array} \right.$$

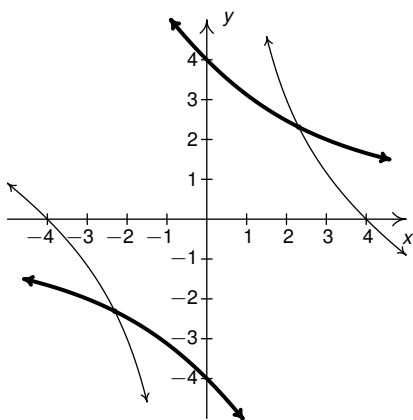
We get  $y^2 - x^2 = 0$  or  $y = \pm x$ . Substituting  $y = x$  into  $E1$  we get  $x^2 + 2x^2 - 16 = 0$  so that  $x^2 = \frac{16}{3}$  or  $x = \pm \frac{4\sqrt{3}}{3}$ . On the other hand, when we substitute  $y = -x$  into  $E1$ , we get  $x^2 - 2x^2 - 16 = 0$  or  $x^2 = -16$  which gives no real solutions.

Substituting each of  $x = \pm \frac{4\sqrt{3}}{3}$  into  $y = x$  yields the solution  $\left\{ \left( \frac{4\sqrt{3}}{3}, \frac{4\sqrt{3}}{3} \right), \left( -\frac{4\sqrt{3}}{3}, -\frac{4\sqrt{3}}{3} \right) \right\}$ . As usual, we leave it to the reader to verifying this solution algebraically.

To verify this equation graphically, we solve  $x^2 + 2xy - 16 = 0$  for  $y$  to obtain  $y = \frac{16-x^2}{2x}$ . We produce the graph of this equation using the techniques described in Section ??.

Solving the second equation,  $y^2 + 2xy - 16 = 0$ , for  $y$ , however, is more complicated. The quadratic formula gives  $y = -x \pm \sqrt{x^2 + 16}$  which requires Calculus or a graphing utility to graph.

As it happens, however, we don't need either because the equation  $y^2 + 2xy - 16 = 0$  can be obtained from the equation  $x^2 + 2xy - 16 = 0$  by interchanging 'y' and 'x'. Thinking back to Section ??, this means we can obtain the graph of  $y^2 + 2xy - 16 = 0$  by reflecting the graph of  $x^2 + 2xy - 16 = 0$  across the line  $y = x$ . Doing so confirms that the two graphs intersect twice: once in Quadrant I, and once in Quadrant III as required.<sup>3</sup>



The graphs of  $x^2 + 2xy - 16 = 0$  and  $y^2 + 2xy - 16 = 0$

2. Unlike the previous problem, there seems to be no avoiding substitution and a bit of algebraic unpleasantness. Solving  $y + 4e^{2x} = 1$  for  $y$ , we get  $y = 1 - 4e^{2x}$  which, when substituted into the second equation, yields  $(1 - 4e^{2x})^2 + 2e^x = 1$ .

After expanding and gathering like terms, we get  $16e^{4x} - 8e^{2x} + 2e^x = 0$ . Factoring gives us  $2e^x (8e^{3x} - 4e^x + 1) = 0$ , and since  $2e^x \neq 0$  for any real  $x$ , we are left with solving  $8e^{3x} - 4e^x + 1 = 0$ .

<sup>3</sup>An amazing coincidence. Of course, we could use a graphing utility to verify our solutions if this didn't happen to be the case.

We have three terms, and even though this is not a 'quadratic in disguise', we can benefit from the substitution  $u = e^x$ . The equation becomes  $8u^3 - 4u + 1 = 0$ . Using the techniques set forth in Section ??, we find  $u = \frac{1}{2}$  is a zero and factor  $8u^3 - 4u + 1$ . We find  $8u^3 - 4u + 1 = 0$  is equivalent to  $(u - \frac{1}{2})(8u^2 + 4u - 2) = 0$ . So in addition to  $u = \frac{1}{2}$ , we need to solve  $8u^2 + 4u - 2 = 0$ .

We use the quadratic formula to solve  $8u^2 + 4u - 2 = 0$  and find  $u = \frac{-1 \pm \sqrt{5}}{4}$ . Since  $u = e^x$ , we now must solve  $e^x = \frac{1}{2}$  and  $e^x = \frac{-1 \pm \sqrt{5}}{4}$ . From  $e^x = \frac{1}{2}$ , we get  $x = \ln\left(\frac{1}{2}\right) = -\ln(2)$ .

As for  $e^x = \frac{-1 \pm \sqrt{5}}{4}$ , we first note that  $\frac{-1 - \sqrt{5}}{4} < 0$ , so  $e^x = \frac{-1 - \sqrt{5}}{4}$  has no real solutions. We are left with  $e^x = \frac{-1 + \sqrt{5}}{4}$ , so that  $x = \ln\left(\frac{-1 + \sqrt{5}}{4}\right)$ .

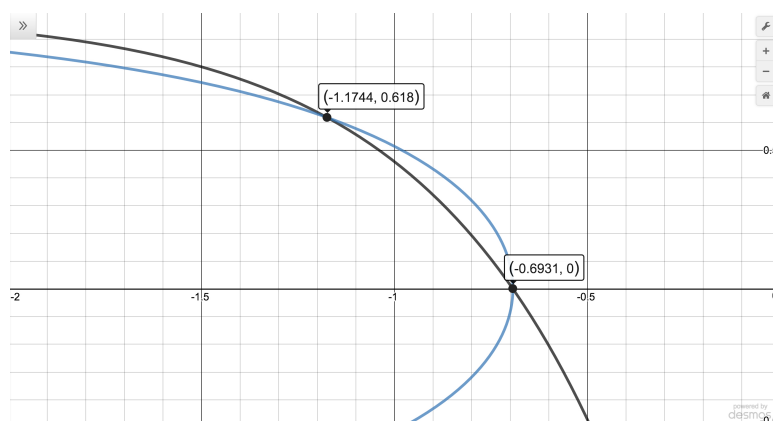
We now return to  $y = 1 - 4e^{2x}$  to find the accompanying  $y$  values for each of our solutions for  $x$ :

$$\begin{aligned} x &= -\ln(2) : \\ y &= 1 - 4e^{2x} \\ &= 1 - 4e^{-2\ln(2)} \\ &= 1 - 4e^{\ln(\frac{1}{4})} \\ &= 1 - 4\left(\frac{1}{4}\right) \\ &= 0 \end{aligned}$$

$$\begin{aligned} x &= \ln\left(\frac{-1 + \sqrt{5}}{4}\right) : \\ y &= 1 - 4e^{2x} \\ &= 1 - 4e^{2\ln\left(\frac{-1 + \sqrt{5}}{4}\right)} \\ &= 1 - 4e^{\ln\left(\frac{-1 + \sqrt{5}}{4}\right)^2} \\ &= 1 - 4\left(\frac{-1 + \sqrt{5}}{4}\right)^2 \\ &= 1 - 4\left(\frac{3 - \sqrt{5}}{8}\right) \\ &= \frac{-1 + \sqrt{5}}{2} \end{aligned}$$

We get two solutions,  $\left\{(-\ln(2), 0), \left(\ln\left(\frac{-1 + \sqrt{5}}{4}\right), \frac{-1 + \sqrt{5}}{2}\right)\right\}$ . It is a good review of the properties of logarithms to verify both solutions algebraically, so we leave that to the reader.

While we are able to sketch  $y = 1 - 4e^{2x}$  using the techniques in Section ??, the second equation is more difficult and we resort to using a graphing utility. We see the two graphs intersect at  $(-0.6931, 0) \approx (-\ln(2), 0)$  and  $(-1.1744, 0.618) \approx \left(\ln\left(\frac{-1 + \sqrt{5}}{4}\right), \frac{-1 + \sqrt{5}}{2}\right)$ .



Graphs for 
$$\begin{cases} y + 4e^{2x} = 1 \\ y^2 + 2e^x = 1 \end{cases}$$

3. Our last system involves three variables and provides an opportunity to gain some insight on how to keep such systems organized. Labeling the equations as before, we have

$$\begin{cases} E1 & z(x-2) = x \\ E2 & yz = y \\ E3 & (x-2)^2 + y^2 = 1 \end{cases}$$

The easiest equation to start with appears to be  $E2$ . While it may be tempting to divide both sides of  $E2$  by  $y$ , we caution against this practice because it presupposes  $y \neq 0$ . Instead, we take  $E2$  and rewrite it as  $yz - y = 0$  so  $y(z - 1) = 0$ . From this, we get two cases:  $y = 0$  or  $z = 1$ .

CASE 1:  $y = 0$ . Substituting  $y = 0$  into  $E1$  and  $E3$ , we get

$$\begin{cases} E1 & z(x-2) = x \\ E3 & (x-2)^2 = 1 \end{cases}$$

Solving  $E3$  for  $x$  gives  $x = 1$  or  $x = 3$ . Substituting these values into  $E1$  gives  $z = -1$  when  $x = 1$  and  $z = 3$  when  $x = 3$ . We obtain two solutions,  $(1, 0, -1)$  and  $(3, 0, 3)$ .

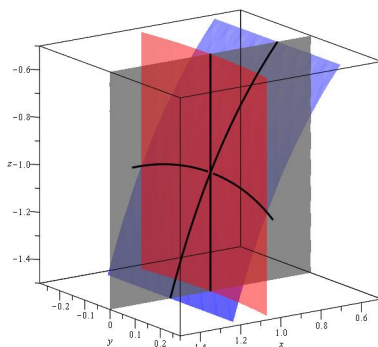
CASE 2:  $z = 1$ . Substituting  $z = 1$  into  $E1$  and  $E3$  gives us

$$\begin{cases} E1 & (1)(x-2) = x \\ E3 & (1-2)^2 + y^2 = 1 \end{cases}$$

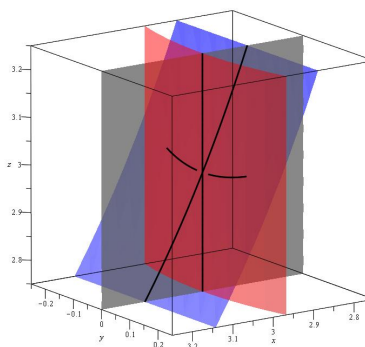
Equation  $E1$  gives us  $x - 2 = x$  or  $-2 = 0$ , which is a contradiction. This means we have no solution to the system in this case, even though  $E3$  is solvable and gives  $y = 0$ .

Hence, our final answer is  $\{(1, 0, -1), (3, 0, 3)\}$ . These points are easy enough to check algebraically in our three original equations, so that is left to the reader.

As for verifying these solutions graphically, they require plotting surfaces in three dimensions and looking for intersection points. While this is beyond the scope of this book, we provide a snapshot of the graphs of our three equations near one of the solution points,  $(1, 0, -1)$  and  $(3, 0, 3)$ .  $\square$



Near  $(-1, 0, 1)$



Near  $(3, 0, 3)$



Example 1.7.2 showcases some of the ingenuity and tenacity mentioned at the beginning of the section. Sometimes you just have to look at a system the right way to find the most efficient method to solve it. Sometimes you just have to try something.

Next we explore some common application problems which give rise to systems of nonlinear equations.

EXAMPLE 1.7.3. Carl decides to explore the Meander River, the location of several recent Sasquatch sightings. From camp, he canoes downstream five miles to check out a purported Sasquatch nest. Finding nothing, he immediately turns around, retraces his route (this time traveling upstream), and returns to camp 3 hours after he left. If Carl canoes at a rate of 6 miles per hour in still water, how fast was the Meander River flowing on that day?

**Solution.** We are given information about distances, rates (speeds) and times. The basic principle relating these quantities is:

$$\text{distance} = \text{rate} \cdot \text{time}$$

The first observation to make, however, is that the distance, rate and time given to us aren't 'compatible': the distance given is the distance for only *part* of the trip, the rate given is the speed Carl can canoe in still water, not in a flowing river, and the time given is the duration of the *entire* trip. Ultimately, we are after the speed of the river, so let's call that  $R$  measured in miles per hour to be consistent with the other rate given to us. To get started, let's divide the trip into its two parts: the initial trip downstream and the return trip upstream. For the downstream trip, all we know is that the distance traveled is 5 miles.

$$\begin{aligned} \text{distance downstream} &= \text{rate traveling downstream} \cdot \text{time traveling downstream} \\ 5 \text{ miles} &= \text{rate traveling downstream} \cdot \text{time traveling downstream} \end{aligned}$$

Since the return trip upstream followed the same route as the trip downstream, we know that the distance traveled upstream is also 5 miles.

$$\begin{aligned} \text{distance upstream} &= \text{rate traveling upstream} \cdot \text{time traveling upstream} \\ 5 \text{ miles} &= \text{rate traveling upstream} \cdot \text{time traveling upstream} \end{aligned}$$

We are told Carl can canoe at a rate of 6 miles per hour in still water. How does this figure into the rates traveling upstream and downstream? The speed the canoe travels in the river is a combination of the speed at which Carl can propel the canoe in still water, 6 miles per hour, and the speed of the river, which we're calling  $R$ . When traveling downstream, the river is helping Carl along, so we *add* these two speeds:

$$\begin{aligned} \text{rate traveling downstream} &= \text{rate Carl propels the canoe} + \text{speed of the river} \\ &= 6 \frac{\text{miles}}{\text{hour}} + R \frac{\text{miles}}{\text{hour}} \end{aligned}$$

So our downstream speed is  $(6 + R) \frac{\text{miles}}{\text{hour}}$ . Substituting this into our 'distance-rate-time' equation for the downstream part of the trip, we get:

$$\begin{aligned} 5 \text{ miles} &= \text{rate traveling downstream} \cdot \text{time traveling downstream} \\ 5 \text{ miles} &= (6 + R) \frac{\text{miles}}{\text{hour}} \cdot \text{time traveling downstream} \end{aligned}$$

When traveling upstream, Carl works *against* the current. Since Carl must move the canoe faster than the river's speed to move upstream, we *subtract* the river's speed *from* Carl's canoeing speed to get:

$$\begin{aligned}\text{rate traveling upstream} &= \text{rate Carl propels the canoe} - \text{river speed} \\ &= 6 \frac{\text{miles}}{\text{hour}} - R \frac{\text{miles}}{\text{hour}}\end{aligned}$$

Proceeding as before, we get

$$\begin{aligned}5 \text{ miles} &= \text{rate traveling upstream} \cdot \text{time traveling upstream} \\ 5 \text{ miles} &= (6 - R) \frac{\text{miles}}{\text{hour}} \cdot \text{time traveling upstream}\end{aligned}$$

The last piece of information given to us is that the total trip lasted 3 hours. If we let  $t_{\text{down}}$  denote the time of the downstream trip and  $t_{\text{up}}$  the time of the upstream trip, we have:  $t_{\text{down}} + t_{\text{up}} = 3$  hours. Substituting  $t_{\text{down}}$  and  $t_{\text{up}}$  into the 'distance-rate-time' equations, we get a system of three equations in three unknowns below. Note that since the variables in equations  $E1$  and  $E2$  are multiplied together, these two equations are nonlinear.

$$\begin{cases} E1 & (6 + R) t_{\text{down}} = 5 \\ E2 & (6 - R) t_{\text{up}} = 5 \\ E3 & t_{\text{down}} + t_{\text{up}} = 3 \end{cases}$$

Since we are ultimately after  $R$ , we need to use these three equations to get at least one equation involving *only*  $R$ . We start with equation  $E1$ . We know that both  $(6 + R) \neq 0$  and  $t_{\text{down}} \neq 0$  since the product of these two quantities is 5 and is nonzero.<sup>4</sup> Hence, we may solve  $E1$  for  $t_{\text{down}}$  by dividing both sides by the quantity  $(6 + R)$  to get  $t_{\text{down}} = \frac{5}{6+R}$ . Similarly, we use  $E2$  to get  $t_{\text{up}} = \frac{5}{6-R}$ . Substituting these into  $E3$ , we get:<sup>5</sup>

$$\frac{5}{6+R} + \frac{5}{6-R} = 3.$$

Clearing denominators, we get  $5(6 - R) + 5(6 + R) = 3(6 + R)(6 - R)$  which reduces to  $R^2 = 16$ . We find  $R = \pm 4$ , and since  $R$  represents the speed of the river, we choose  $R = 4$ . On the day in question, the Meander River is flowing at a rate of 4 miles per hour.  $\square$

One of the important lessons to learn from Example 1.7.3 is that speeds, and more generally, rates, are additive. As we see in our next example, the concept of rate and its associated principles can be applied to a wide variety of problems - not just 'distance-rate-time' scenarios.

**EXAMPLE 1.7.4.** Working alone, Taylor can weed the garden in 4 hours. If Carl helps, they can weed the garden in 3 hours. How long would it take for Carl to weed the garden on his own?

**Solution.** The key relationship between work and time which we use in this problem is:

$$\text{amount of work done} = \text{rate of work} \cdot \text{time spent working}$$

We are told that, working alone, Taylor can weed the garden in 4 hours. In Taylor's case then:

$$\begin{aligned}\text{amount of work Taylor does} &= \text{rate of Taylor working} \cdot \text{time Taylor spent working} \\ 1 \text{ garden} &= (\text{rate of Taylor working}) \cdot (4 \text{ hours})\end{aligned}$$

<sup>4</sup>This is a restatement of the Zero Product Property from Section ??.

<sup>5</sup>The reader is encouraged to verify that the units in this equation are consistent. For starters, the units on the '3' is 'hours.'

So we have that the rate Taylor works is  $\frac{1 \text{ garden}}{4 \text{ hours}} = \frac{1}{4} \frac{\text{garden}}{\text{hour}}$ . We are also told that when working together, Taylor and Carl can weed the garden in just 3 hours. We have:

$$\begin{aligned} \text{amount of work done together} &= \text{rate of working together} \cdot \text{time spent working together} \\ 1 \text{ garden} &= (\text{rate of working together}) \cdot (3 \text{ hours}) \end{aligned}$$

From this, we find that the rate of Taylor and Carl working together is  $\frac{1 \text{ garden}}{3 \text{ hours}} = \frac{1}{3} \frac{\text{garden}}{\text{hour}}$ . We are asked to find out how long it would take for Carl to weed the garden on his own. Let us call this unknown  $t$ , measured in hours to be consistent with the other times given to us in the problem. Then:

$$\begin{aligned} \text{amount of work Carl does} &= \text{rate of Carl working} \cdot \text{time Carl spent working} \\ 1 \text{ garden} &= (\text{rate of Carl working}) \cdot (t \text{ hours}) \end{aligned}$$

In order to find  $t$ , we need to find the rate of Carl working, so let's call this quantity  $R$ , with units  $\frac{\text{garden}}{\text{hour}}$ . Using the fact that rates are additive, we have:

$$\begin{aligned} \text{rate working together} &= \text{rate of Taylor working} + \text{rate of Carl working} \\ \frac{1}{3} \frac{\text{garden}}{\text{hour}} &= \frac{1}{4} \frac{\text{garden}}{\text{hour}} + R \frac{\text{garden}}{\text{hour}} \end{aligned}$$

so that  $R = \frac{1}{12} \frac{\text{garden}}{\text{hour}}$ . Substituting this into our 'work-rate-time' equation for Carl, we get:

$$\begin{aligned} 1 \text{ garden} &= (\text{rate of Carl working}) \cdot (t \text{ hours}) \\ 1 \text{ garden} &= \left( \frac{1}{12} \frac{\text{garden}}{\text{hour}} \right) \cdot (t \text{ hours}) \end{aligned}$$

Solving  $1 = \frac{1}{12}t$ , we get  $t = 12$ , so it takes Carl 12 hours to weed the garden on his own.<sup>6</sup> □

As is common with 'word problems' like Examples 1.7.3 and 1.7.4, there is no short-cut to the answer. Note that in Examples 1.7.3, we formalized the system of non-linear equations before solving whereas in Example 1.7.4, the system remained much in the background.<sup>7</sup> We encourage the reader to carefully think through and apply the basic principles of rate to each (potentially different!) situation. It is time well spent. We also encourage the tracking of units, especially in the early stages of the problem. Not only does this promote uniformity in the units, it also serves as a quick means to check if an equation makes sense.<sup>8</sup>

<sup>6</sup>Carl would much rather spend his time writing open-source Mathematics texts than gardening anyway.

<sup>7</sup>The reader is invited to formalize this system in Exercise ??.

<sup>8</sup>In other words, make sure you don't try to add apples to oranges!

## 1.7.1 Exercises

Exercise idea: follow up on last example and have students formalize the system there using given variables.

In Exercises 1 - 6, solve the given system of nonlinear equations. Sketch the graph of both equations on the same set of axes to verify the solution set.

$$\begin{array}{lll} 1. \begin{cases} x^2 - y = 4 \\ x^2 + y^2 = 4 \end{cases} & 2. \begin{cases} x^2 + y^2 = 4 \\ x^2 - y = 5 \end{cases} & 3. \begin{cases} x^2 + y^2 = 16 \\ 16x^2 + 4y^2 = 64 \end{cases} \\ 4. \begin{cases} x^2 + y^2 = 16 \\ 9x^2 - 16y^2 = 144 \end{cases} & 5. \begin{cases} x^2 + y^2 = 16 \\ \frac{1}{9}y^2 - \frac{1}{16}x^2 = 1 \end{cases} & 6. \begin{cases} x^2 + y^2 = 16 \\ x - y = 2 \end{cases} \end{array}$$

In Exercises 7 - 15, solve the given system of nonlinear equations. Use a graph to help you avoid any potential extraneous solutions.

$$\begin{array}{lll} 7. \begin{cases} x^2 - y^2 = 1 \\ x^2 + 4y^2 = 4 \end{cases} & 8. \begin{cases} \sqrt{x+1} - y = 0 \\ x^2 + 4y^2 = 4 \end{cases} & 9. \begin{cases} x + 2y^2 = 2 \\ x^2 + 4y^2 = 4 \end{cases} \\ 10. \begin{cases} (x-2)^2 + y^2 = 1 \\ x^2 + 4y^2 = 4 \end{cases} & 11. \begin{cases} x^2 + y^2 = 25 \\ y - x = 1 \end{cases} & 12. \begin{cases} x^2 + y^2 = 25 \\ x^2 + (y-3)^2 = 10 \end{cases} \\ 13. \begin{cases} y = x^3 + 8 \\ y = 10x - x^2 \end{cases} & 14. \begin{cases} x^2 - xy = 8 \\ y^2 - xy = 8 \end{cases} & 15. \begin{cases} x^2 + y^2 = 25 \\ 4x^2 - 9y = 0 \\ 3y^2 - 16x = 0 \end{cases} \end{array}$$

16. A certain bacteria culture follows the Law of Uninbited Growth, Equation ???. After 10 minutes, there are 10,000 bacteria. Five minutes later, there are 14,000 bacteria. How many bacteria were present initially? How long before there are 50,000 bacteria?

Consider the system of nonlinear equations below

$$\begin{cases} \frac{4}{x} + \frac{3}{y} = 1 \\ \frac{3}{x} + \frac{2}{y} = -1 \end{cases}$$

If we let  $u = \frac{1}{x}$  and  $v = \frac{1}{y}$  then the system becomes

$$\begin{cases} 4u + 3v = 1 \\ 3u + 2v = -1 \end{cases}$$

This associated system of linear equations can then be solved using any of the techniques presented earlier in the chapter to find that  $u = -5$  and  $v = 7$ . Thus  $x = \frac{1}{u} = -\frac{1}{5}$  and  $y = \frac{1}{v} = \frac{1}{7}$ .

We say that the original system is **linear in form** because its equations are not linear but a few substitutions reveal a structure that we can treat like a system of linear equations. Each system in Exercises 17 - 19 is linear in form. Make the appropriate substitutions and solve for  $x$  and  $y$ .

$$17. \begin{cases} 4x^3 + 3\sqrt{y} = 1 \\ 3x^3 + 2\sqrt{y} = -1 \end{cases} \quad 18. \begin{cases} 4e^x + 3e^{-y} = 1 \\ 3e^x + 2e^{-y} = -1 \end{cases} \quad 19. \begin{cases} 4\ln(x) + 3y^2 = 1 \\ 3\ln(x) + 2y^2 = -1 \end{cases}$$

20. Solve the following system

$$\begin{cases} x^2 + \sqrt{y} + \log_2(z) = 6 \\ 3x^2 - 2\sqrt{y} + 2\log_2(z) = 5 \\ -5x^2 + 3\sqrt{y} + 4\log_2(z) = 13 \end{cases}$$

21. Systems of nonlinear equations show up in third semester Calculus in the midst of some really cool problems. The system below came from a problem in which we were asked to find the dimensions of a rectangular box with a volume of 1000 cubic inches that has minimal surface area. The variables  $x$ ,  $y$  and  $z$  are the dimensions of the box and  $\lambda$  is called a Lagrange multiplier. With the help of your classmates, solve the system.<sup>9</sup>

$$\begin{cases} 2y + 2z = \lambda yz \\ 2x + 2z = \lambda xz \\ 2y + 2x = \lambda xy \\ xyz = 1000 \end{cases}$$

22. According to Theorem ?? in Section ??, the polynomial  $p(x) = x^4 + 4$  can be factored into the product linear and irreducible quadratic factors. In this exercise, we present a method for obtaining that factorization.

(a) Show that  $p$  has no real zeros.

(b) Because  $p$  has no real zeros, its factorization must be of the form  $(x^2 + ax + b)(x^2 + cx + d)$  where each factor is an irreducible quadratic. Expand this quantity and gather like terms together.

(c) Create and solve the system of nonlinear equations which results from equating the coefficients of the expansion found above with those of  $x^4 + 4$ . You should get four equations in the four unknowns  $a$ ,  $b$ ,  $c$  and  $d$ . Write  $p(x)$  in factored form.

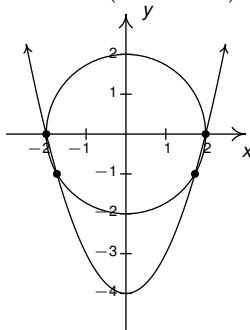
23. Factor  $q(x) = x^4 + 6x^2 - 5x + 6$ .

---

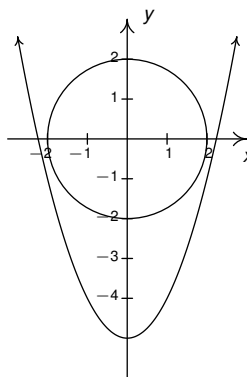
<sup>9</sup>If using  $\lambda$  bothers you, change it to  $w$  when you solve the system.

## 1.7.2 Answers

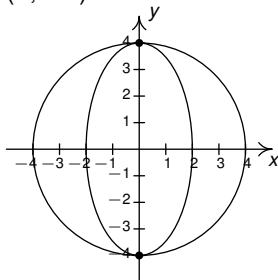
1.  $(\pm 2, 0), (\pm\sqrt{3}, -1)$



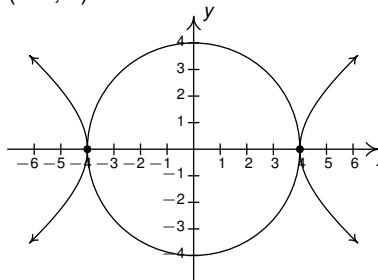
2. No solution



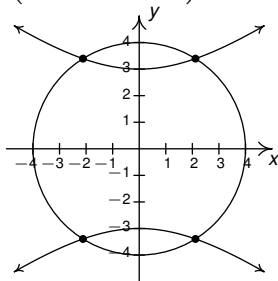
3.  $(0, \pm 4)$



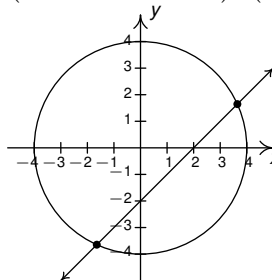
4.  $(\pm 4, 0)$



5.  $(\pm \frac{4\sqrt{7}}{5}, \pm \frac{12\sqrt{2}}{5})$



6.  $(1 + \sqrt{7}, -1 + \sqrt{7}), (1 - \sqrt{7}, -1 - \sqrt{7})$



7.  $(\pm \frac{2\sqrt{10}}{5}, \pm \frac{\sqrt{15}}{5})$

8.  $(0, 1)$

9.  $(0, \pm 1), (2, 0)$

10.  $(\frac{4}{3}, \pm \frac{\sqrt{5}}{3})$

11.  $(3, 4), (-4, -3)$

12.  $(\pm 3, 4)$

13.  $(-4, -56), (1, 9), (2, 16)$

14.  $(-2, 2), (2, -2)$

15.  $(3, 4)$

16. Initially, there are  $\frac{250000}{49} \approx 5102$  bacteria. It will take  $\frac{5 \ln(49/5)}{\ln(7/5)} \approx 33.92$  minutes for the colony to grow to 50,000 bacteria.

17.  $(-\sqrt[3]{5}, 49)$

18. No solution

19.  $(e^{-5}, \pm\sqrt{7})$

20.  $(1, 4, 8), (-1, 4, 8)$

21.  $x = 10, y = 10, z = 10, \lambda = \frac{2}{5}$

22. (c)  $x^4 + 4 = (x^2 - 2x + 2)(x^2 + 2x + 2)$

23.  $x^4 + 6x^2 - 5x + 6 = (x^2 - x + 1)(x^2 + x + 6)$

## 1.8 Inequalities and Regions in the Plane

With few exceptions, we have spent our time in this course graphing equations relating two variables. In this section, we explore graphing inequalities relating two variables which usually a two-dimensional *region* in the plane instead of a one dimensional line or curve.<sup>1</sup> In our first example, we restrict our attention to looking at regions in the  $xy$ -plane bounded by equations which describe  $y$  as a function of  $x$ .

EXAMPLE 1.8.1.

1. Sketch the following sets of points in the  $xy$ -plane.

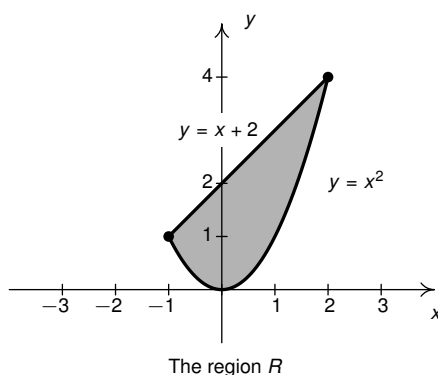
(a)  $R = \{(x, y) \mid y > |x|\}$

(b)  $S = \{(x, y) \mid y \leq 6 - x^2\}$

(c)  $T = \{(x, y) \mid |x| < y \leq 6 - x^2\}$

2. Find a set builder description for each of the following regions described below:

- (a) The region graphed below:



- (b) The region  $S$  between the graphs<sup>2</sup> of  $f(x) = x^3 - 3x$  and  $g(x) = x$ .

**Solution.**

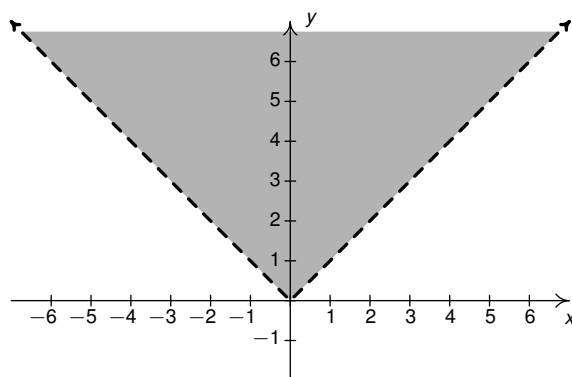
1. (a) The set  $R$  consists of all points  $(x, y)$  whose  $y$ -coordinate is greater than  $|x|$ . If we graph  $y = |x|$ , we get the familiar 'V' shape with vertex at  $(0, 0)$ . Hence, in order for an ordered pair  $(x, y)$  to satisfy  $y > |x|$ , the point  $(x, y)$  must be *above* the graph of  $y = |x|$ . Since the inequality here,  $y > |x|$  is strict, we use a dashed line with which to indicate the graph of  $y = |x|$  and shade the region above, or 'inside,' the 'V' as shown below on the left. Note that one way to check our answer is to choose points both inside and outside the shaded region to verify the inequality either holds or it doesn't, respectively.

<sup>1</sup>We have *some* experience describing simple regions in the plane using inequalities from our work in Section ??.

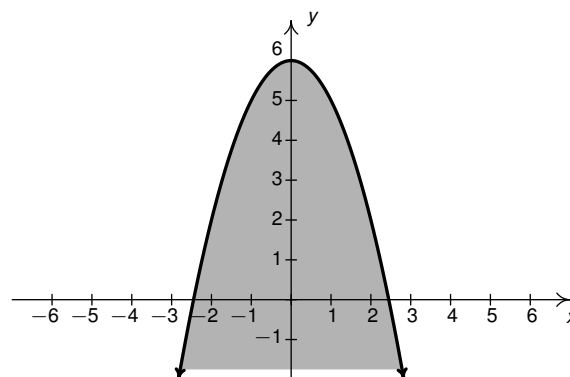
<sup>2</sup>sometimes written as 'the region bounded by the graphs' ...



- (b) Using the same reasoning as above we note a point  $(x, y)$  is in  $S$  if its  $y$ -coordinate is less than or equal to the  $y$ -coordinate on the parabola  $y = 6 - x^2$ . Hence, we graph the points *below* the parabola ( $y < 6 - x^2$ ) along with the points *on* the parabola ( $y = 6 - x^2$ ) below on the right.



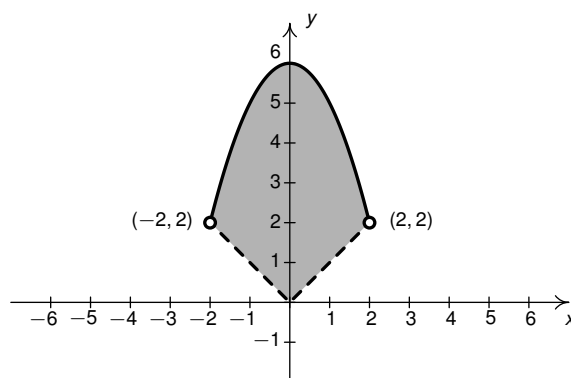
$$R = \{(x, y) : y > |x|\}$$



$$S = \{(x, y) : y \leq 6 - x^2\}$$

- (c) For a point  $(x, y)$  to be in  $T$ , the  $y$ -coordinate must satisfy  $|x| < y \leq 6 - x^2$  which means it must belong to both  $R$  and  $S$ .<sup>3</sup> Thus we shade the region *between*  $y = |x|$  and  $y = 6 - x^2$ , keeping those points on the parabola, but not the points on  $y = |x|$ .

To get an accurate graph, we need to find where these two graphs intersect, so we set  $|x| = 6 - x^2$ . Using the techniques discussed in Sections ?? and ??, we find  $x = -2, 2$ . To find the associated  $y$ -coordinates of the intersection points, we substitute  $x = \pm 2$  into either  $y = |x|$  or  $y = 6 - x^2$  (or both, to check!) In both cases, we find the associated  $y$ -coordinate to be 2, so the intersection points are  $(-2, 2)$  and  $(2, 2)$ . On our graph, however, these are the location of 'holes' owing to the strictness of the inequality  $y > |x|$ .



$$T = \{(x, y) : |x| < y \leq 6 - x^2\}$$

2. (a) Verbally, we may describe  $R$  as the region *between* the graphs of  $y = x^2$  and  $y = x + 2$ . More specifically,  $R$  is the set of points *above* (or on) the parabola  $y = x^2$  but *below* (or on) the line  $y = x + 2$ . That is, for the points  $(x, y)$  in  $R$ , we have  $x^2 \leq y \leq x + 2$ .

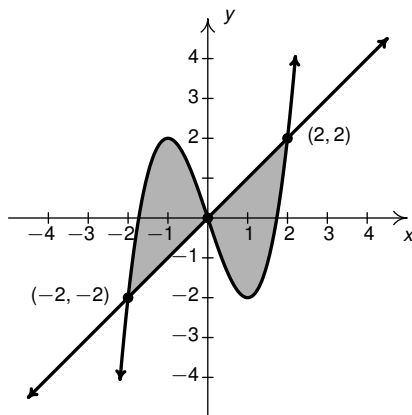
<sup>3</sup>Said differently,  $T$  is the set-theoretic intersection of  $R$  and  $S$ ,  $T = R \cap S$  as described in ??.

To find the restrictions on  $x$ , we find the intersection of the graphs. Solving  $x^2 = x + 2$  gives  $x = -1$  and  $x = 2$ . Hence, to be in  $R$ , points  $(x, y)$  need to satisfy  $-1 \leq x \leq 2$ . Hence, one way to describe  $R$  is  $R = \{(x, y) \mid -1 \leq x \leq 2, x^2 \leq y \leq x + 2\}$ .

While this answer is correct, the fact that the line and the parabola intersect only twice producing only *one* region where the inequality  $x^2 \leq y \leq x + 2$  is true, we may shorten our description of  $R$  to just  $R = \{(x, y) \mid x^2 \leq y \leq x + 2\}$ .

- (b) Since we are given a verbal description of  $S$ , we first sketch the graph to get a sense of the geometric situation. To graph  $f(x) = x^3 - 3x$ , we use what we learned in Section ??: end behavior, behavior near the  $x$ -intercepts, and symmetry. The graph of  $g(x) = x$  is a line with slope 1 through the origin.

Graphing both of these functions on the same pair of axes, we see three intersection points. Solving  $f(x) = g(x)$ , that is  $x^3 - 3x = x$  gives  $x = 0$  and  $x = \pm 2$ . These solutions correspond to the intersection points  $(-2, -2)$ ,  $(0, 0)$ , and  $(2, 2)$ .  $S$  is described as the region 'between' these graphs, so we shade between the graphs accordingly.



Graphing  $f(x) = x^3 - 3x$  and  $g(x) = x$

From the graph, we see that for  $-2 < x < 0$ , the graph of  $f$  is above the graph of  $g$ . Algebraically, this means that for all  $x$  with  $-2 \leq x \leq 0$ ,  $f(x) \geq g(x)$  or  $x^3 - 3x \geq x$ . At  $x = 0$ , the situation changes and we have  $f(x) \leq g(x)$  or  $x^3 - 3x \leq x$  for  $0 \leq x \leq 2$ . One to describe  $S$  is to describe each of these pieces and take their set theoretic union.

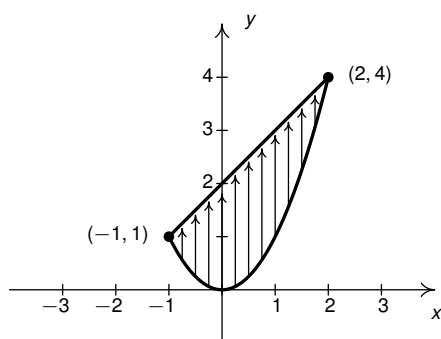
We first consider the piece for  $-2 \leq x \leq 0$ . Unlike the previous problem, we cannot completely describe this region in terms of the  $y$ -values since there are *two* regions where  $x \leq y \leq x^3 - 3x$  is true: one where  $-2 \leq x \leq 0$  (the one we want) and one here  $x \geq 2$  (which we don't want.) Hence, we describe this first piece as  $\{(x, y) \mid -2 \leq x \leq 0, x \leq y \leq x^3 - 3x\}$ . Following this methodology for the second piece, we obtain the complete description for  $S$  below:

$$S = \{(x, y) \mid -2 \leq x \leq 0, x \leq y \leq x^3 - 3x\} \cup \{(x, y) \mid 0 \leq x \leq 2, x^3 - 3x \leq y \leq x\}.$$

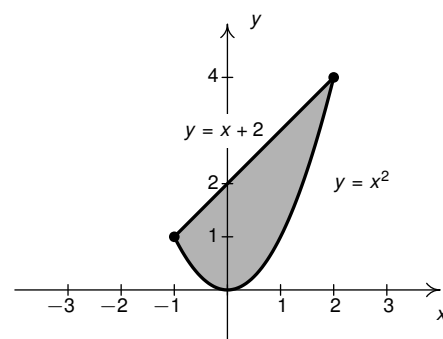
□

A few remarks about Example 1.8.1 are in order. First note that each of the regions presented here can be viewed as graphs of relations as described in Section ?? . Moreover, there are many ways to describe a region so our answers to number 2 above are by no means unique.

In particular, our solution  $R = \{(x, y) \mid -1 \leq x \leq 2, x^2 \leq y \leq x + 2\}$  to number 2a can be visualized as 'filling up' the region  $R$  from the bottom curve,  $y = x^2$  to the top curve,  $y = x + 2$  as  $x$  runs from the leftmost extent of the region at  $x = -1$  to the rightmost extent at  $x = 2$  as indicated below. The notion that each  $x$  determines where to start and stop filling the region is a consequence of us viewing the bounding curves as functions of  $x$ . That is, for each  $x$ , we can determine the lower boundary of the region,  $y = x^2$  and the upper boundary of the region,  $y = x + 2$ .

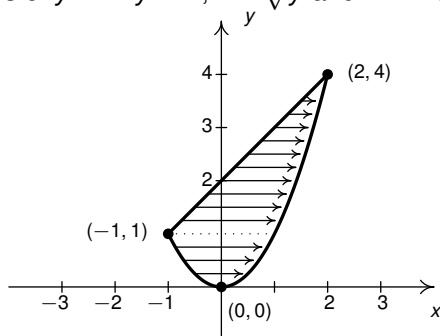


'Filling up'  $R$  from bottom to top.

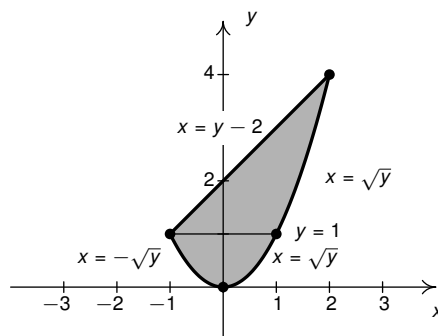


The region  $R$ .

There are times in Calculus where it may be convenient to describe the region  $R$  as filling up left-to-right as  $y$  varies from the bottom most extent,  $y = 0$  to the top most extent,  $y = 4$ . In this case, we need to describe the bounding curves as functions of  $y$ . To that end, we solve  $y = x + 2$  and  $y = x^2$  for  $x$  and get three functions of  $y$ :  $x = y - 2$ ,  $x = \sqrt{y}$  and  $x = -\sqrt{y}$  as labeled below.



'Filling up'  $R$  from left to right.



The region  $R$

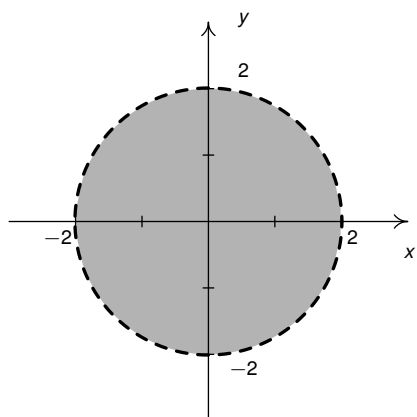
Based on the diagram, we see we need to describe  $R$  as *two* pieces. The first piece being bounded on the left by  $x = -\sqrt{y}$  and on the right  $x = \sqrt{y}$  from  $y = 0$  to  $y = 1$ , and the second piece bounded on the left by  $x = y - 2$  and on the right by  $x = \sqrt{y}$  from  $y = 1$  to  $y = 4$ . That is,

$$R = \{(x, y) \mid 0 \leq y \leq 1, -\sqrt{y} \leq x \leq \sqrt{y}\} \cup \{(x, y) \mid 1 \leq y \leq 4, y - 2 \leq x \leq \sqrt{y}\}.$$

Not all regions in the plane are best described using functions of  $x$  or  $y$ .<sup>4</sup> Suppose, for instance, we wish to sketch the region  $\{(x, y) \mid x^2 < 4 - y^2\}$ . Algebraically, we wish to plot all points  $(x, y)$  for which the inequality  $x^2 < 4 - y^2$  is true. One way to proceed is to mimic the ‘sign diagram’ routine we use for solving nonlinear inequalities in one variable: rewrite the inequality so as to obtain 0 on one side of the inequality, find the zeros of the non-zero side, choose test values determined by the zeros, and record our solution.

First, we gather all of the terms on one side and leave a 0 on the other:  $x^2 + y^2 - 4 < 0$ . Next, we find the zeros of the left hand side, that is, where is  $x^2 + y^2 - 4 = 0$ . Rewriting, we get  $x^2 + y^2 = 4$  which describes the circle of radius 2 centered at the origin. In other words, instead of obtaining a few *numbers* which divide the real number *line* into *intervals*, we get an equation of a *curve*, in this case, a circle, which divides the *plane* into two *regions* - the ‘inside’ and ‘outside’ of the circle.

Just like we used test *values* to determine whether or not an interval belongs to the solution of the inequality, we use test *points* in the each of the regions to see which of these belong to our solution set.<sup>5</sup> We choose  $(0, 0)$  to represent the region inside the circle and  $(0, 3)$  to represent the points outside of the circle. When we substitute  $(0, 0)$  into  $x^2 + y^2 - 4 < 0$ , we get  $-4 < 0$  which is true. This means  $(0, 0)$  and all the other points inside the circle are part of the solution. On the other hand, when we substitute  $(0, 3)$  into the same inequality, we get  $5 < 0$  which is false. This means  $(0, 3)$  along with all other points outside the circle are not part of the solution. What about points on the circle itself? Choosing a point on the circle, say  $(0, 2)$ , we get  $0 < 0$ , which means the circle itself does not satisfy the inequality.<sup>6</sup> As a result, we leave the circle dashed in the final diagram.



The solution to  $x^2 < 4 - y^2$

We put this technique to good use in the following example.

<sup>4</sup>See Section ?? for instance!

<sup>5</sup>The theory behind why all this works is, surprisingly, the same theory which guarantees that sign diagrams work the way they do - continuity and the Intermediate Value Theorem - but in this case, applied to functions of more than one variable.

<sup>6</sup>Another way to see this is that points on the circle satisfy  $x^2 + y^2 - 4 = 0$ , so they do not satisfy  $x^2 + y^2 - 4 < 0$ .

## EXAMPLE 1.8.2.

1. Sketch the following regions in the  $xy$ -plane:

(a)  $R = \{(x, y) \mid y^2 - 4 \leq x < y + 2\}$

(b) The solution to: 
$$\begin{cases} x^2 + y^2 \geq 4 \\ x^2 - 2x + y^2 - 2y \leq 0 \end{cases}$$

2. Find a set builder description for the region between the Unit Circle and the graph of  $x^2 + y^2 = 16$ .

**Solution.**

1. (a) The inequality  $y^2 - 4 \leq x < y + 2$  is a compound inequality which translates as  $y^2 - 4 \leq x$  and  $x < y + 2$ . Hence, our approach is to solve each inequality separately and take the set theoretic intersection to determine the region which satisfies both inequalities.

Starting with  $y^2 - 4 \leq x$ , we rewrite this as  $y^2 - x - 4 \leq 0$  and set about graphing  $y^2 - x - 4 = 0$ . Since only one variable is squared, we know this equation describes a parabola. Rewriting in standard form, we get  $y^2 = x + 4$ , so the vertex is  $(-4, 0)$  and the parabola opens to the right.

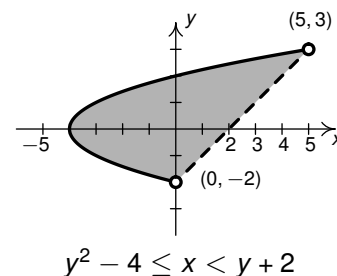
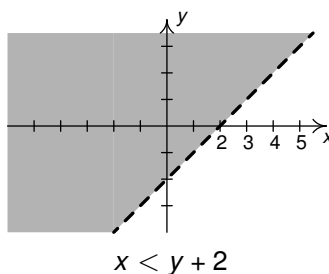
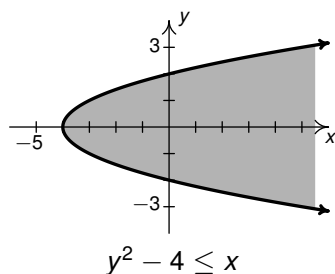
Using the test points  $(-5, 0)$  and  $(0, 0)$ , we find that the solution to the inequality includes the region to the *right* of, or 'inside', the parabola. The points on the parabola itself are also part of the solution, owing to the 'equality' in the 'inequality.' (We could also check a point on the parabola such as  $(-4, 0)$  satisfies the inequality.)

Turning our attention to  $x < y + 2$ , we first rewrite this as  $x - y - 2 < 0$  and focus our attention on  $x - y - 2 = 0$ . Rewriting, we have the line  $y = x - 2$  and using the test points  $(0, 0)$  and  $(0, -4)$ , we find points in the region *above* the line satisfy the inequality. (Owing to the strictness of the inequality, the points on the line itself do not.)<sup>7</sup>

We see that these two regions do overlap but to make the graph more precise, we seek the intersection of these two curves. That is, we need to solve the system of nonlinear equations

$$\begin{cases} (E1) & y^2 = x + 4 \\ (E2) & y = x - 2 \end{cases}$$

Solving  $E1$  for  $x$ , we get  $x = y^2 - 4$ . Substituting this into  $E2$  gives  $y = y^2 - 4 - 2$ , or  $y^2 - y - 6 = 0$ . We find  $y = -2$  and  $y = 3$  and since  $x = y^2 - 4$ , we get that the graphs intersect at  $(0, -2)$  and  $(5, 3)$ . Putting all of this together, we get our final answer below.



<sup>7</sup>We could also have rewritten  $x < y + 2$  as  $y > x - 2$  directly and sketched the region as explained in Example 1.8.1.

- (b) Like any system, our solution to this problem requires us to graph the points  $(x, y)$  which satisfy *both* inequalities. To do this, we solve each inequality separately and take the set theoretic intersection of the solution sets.

We begin with the inequality  $x^2 + y^2 \geq 4$  which we rewrite as  $x^2 + y^2 - 4 \geq 0$ . The points which satisfy  $x^2 + y^2 - 4 = 0$  form the circle  $x^2 + y^2 = 4$ . Using test points  $(0, 0)$ ,  $(0, 2)$ , and  $(0, 3)$  we find that our solution comprises the region *outside* the circle along with the circle itself.

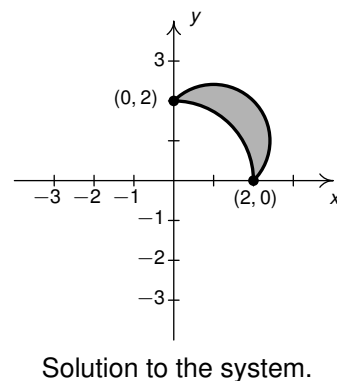
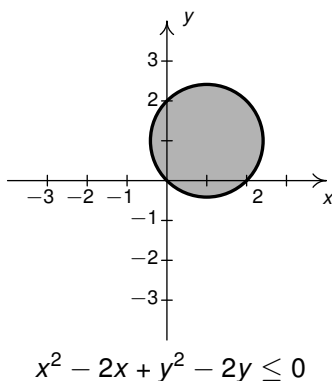
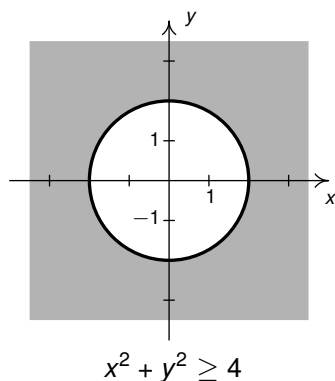
Moving to  $x^2 - 2x + y^2 - 2y \leq 0$ , we start with  $x^2 - 2x + y^2 - 2y = 0$ . Completing the squares, we obtain  $(x - 1)^2 + (y - 1)^2 = 2$ , which is a circle centered at  $(1, 1)$  with a radius of  $\sqrt{2}$ .

Choosing  $(1, 1)$  to represent the inside of the circle,  $(0, 0)$  as a point on the circle, and  $(1, 3)$  as a point outside of the circle, we find that the solution to the inequality is the inside of the circle, including the circle itself.

Our final answer, then, consists of the points on or outside of the circle  $x^2 + y^2 = 4$  which lie on or inside the circle  $(x - 1)^2 + (y - 1)^2 = 2$ . To produce the most accurate graph, we need to find where these circles intersect. To that end, we solve the system

$$\begin{cases} (E1) & x^2 + y^2 = 4 \\ (E2) & x^2 - 2x + y^2 - 2y = 0 \end{cases}$$

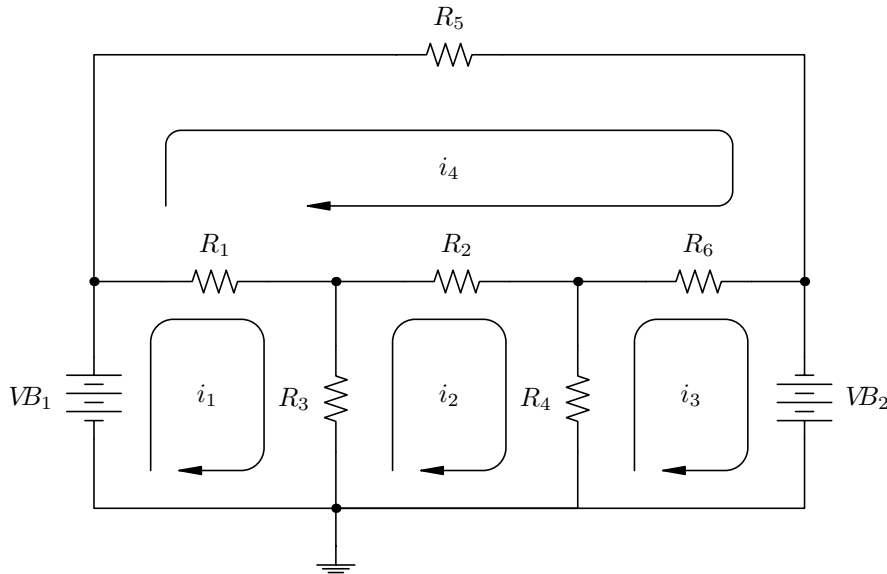
We can eliminate both the  $x^2$  and  $y^2$  by replacing  $E2$  with  $-E1 + E2$ . Doing so produces  $-2x - 2y = -4$ . Solving for  $y$  gives  $y = 2 - x$ . Substituting this into  $E1$  gives  $x^2 + (2 - x)^2 = 4$  which simplifies to  $x^2 + 4 - 4x + x^2 = 4$  or  $2x^2 - 4x = 0$ . Factoring yields  $2x(x - 2)$  which gives  $x = 0$  or  $x = 2$ . Substituting these values into  $y = 2 - x$  gives the points  $(0, 2)$  and  $(2, 0)$ . The intermediate graphs and final solution are below.



2. We first recall the Unit Circle is the circle centered at the origin with radius 1 and hence is described algebraically by the equation  $x^2 + y^2 = 1$ . The equation  $x^2 + y^2 = 16$  describes a circle centered at the origin with radius 4. Hence, the 'region between' these two circles indicates we want is *outside* the Unit Circle but *inside* the circle  $x^2 + y^2 = 16$ .

Based on our experience from the earlier problems, we know points outside the Unit Circle satisfy  $x^2 + y^2 > 1$  whereas the points inside the circle  $x^2 + y^2 = 16$  satisfy  $x^2 + y^2 < 16$ . Hence, the points we seek satisfy *both* inequalities, so our solution is  $\{(x, y) \mid 1 < x^2 + y^2 < 16\}$ .  $\square$

We close this section with a follow-up to Example 1.4.2 in Section 1.4. Recall in the circuit diagrammed below, we have two batteries with source voltages  $VB_1$  and  $VB_2$ , measured in volts  $V$ , along with six resistors with resistances  $R_1$  through  $R_6$ , measured in kilohms,  $k\Omega$ . Recall if we think of electrons flowing through the circuit, we can think of the voltage sources as providing the ‘push’ which makes the electrons move, the resistors as obstacles for the electrons to overcome, and the mesh current as a net rate of flow of electrons around the indicated loops.



Using [Ohm's Law](#) and [Kirchhoff's Voltage Law](#), we can relate the voltage supplied to the circuit by the two batteries to the voltage drops across the six resistors in order to find the four ‘mesh’ currents:  $i_1$ ,  $i_2$ ,  $i_3$  and  $i_4$ , measured in milliamps,  $mA$ . This gives rise to the following system of linear equations:

$$\begin{cases} (R_1 + R_3) i_1 - R_3 i_2 - R_1 i_4 = VB_1 \\ -R_3 i_1 + (R_2 + R_3 + R_4) i_2 - R_4 i_3 - R_2 i_4 = 0 \\ -R_4 i_2 + (R_4 + R_6) i_3 - R_6 i_4 = -VB_2 \\ -R_1 i_1 - R_2 i_2 - R_6 i_3 + (R_1 + R_2 + R_5 + R_6) i_4 = 0 \end{cases}$$

In Example 1.4.2, we found that under the assumptions  $VB_1 = 10V$ ,  $VB_2 = 5V$ , and all the resistances are all  $1k\Omega$ , the mesh currents worked out to be  $i_1 = 10.625 mA$ ,  $i_2 = 6.25 mA$ ,  $i_3 = 3.125 mA$ , and  $i_4 = 5 mA$ . In our final example, we assume  $VB_1 = 10V$  and  $VB_2 = 5V$  and work to find what combination of resistances would combine to produce these mesh currents.

EXAMPLE 1.8.3. For the circuit described above, if  $VB_1 = 10V$  and  $VB_2 = 5V$ , find the possible combinations of resistances which yield the currents  $i_1 = 10.625 mA$ ,  $i_2 = 6.25 mA$ ,  $i_3 = 3.125 mA$ , and  $i_4 = 5 mA$ .

**Solution.**

We begin by substituting the known values of the currents into the system of equations. We get:

$$\begin{cases} 5.625R_1 + 4.375R_3 = 10 \\ 1.25R_2 - 4.375R_3 + 3.125R_4 = 0 \\ -3.125R_4 - 1.875R_6 = -5 \\ -5.625R_1 - 1.25R_2 + 5R_5 + 1.875R_6 = 0 \end{cases}$$

The coefficient matrix for this system is  $4 \times 6$  (4 equations with 6 unknowns) and is therefore not invertible. We do know, however, this system is consistent, since setting all the resistance values equal to 1 corresponds to our situation Example 1.4.2. This means we have an underdetermined consistent system which is necessarily dependent. To solve this system, we encode it into an augmented matrix

$$\left[ \begin{array}{cccccc|c} 5.625 & 0 & 4.375 & 0 & 0 & 0 & 10 \\ 0 & 1.25 & -4.375 & 3.125 & 0 & 0 & 0 \\ 0 & 0 & 0 & -3.125 & 0 & -1.875 & -5 \\ -5.625 & -1.25 & 0 & 0 & 5 & 1.875 & 0 \end{array} \right]$$

A graphing utility gives the reduced-row echelon form of the matrix as:

$$\left[ \begin{array}{cccccc|c} 1 & 0 & 0.\bar{7} & 0 & 0 & 0 & 1.\bar{7} \\ 0 & 1 & -3.5 & 0 & 0 & -1.5 & -4 \\ 0 & 0 & 0 & 1 & 0 & 0.6 & 1.6 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 \end{array} \right]$$

from which we obtain the system:

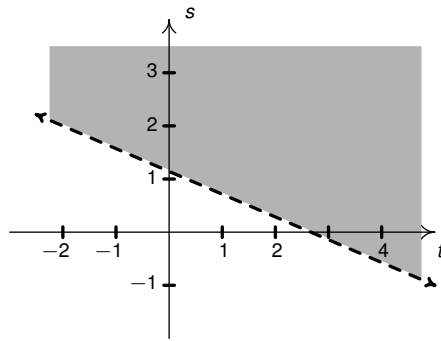
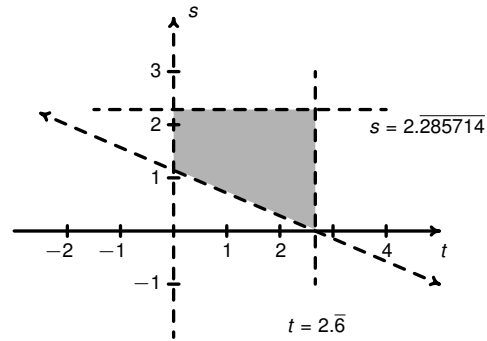
$$\begin{cases} R_1 + 0.\bar{7}R_3 = 1.\bar{7} \\ R_2 - 3.5R_3 - 1.5R_6 = -4 \\ R_4 + 0.6R_6 = 1.6 \\ R_5 = 1 \end{cases}$$

We can solve for  $R_1$ ,  $R_2$ ,  $R_4$  and  $R_5$  leaving  $R_3$  and  $R_6$  as free variables. Labeling  $R_3 = s$  and  $R_6 = t$ , we have  $R_1 = -0.\bar{7}s + 1.\bar{7}$ ,  $R_2 = 3.5s + 1.5t - 4$ ,  $R_4 = -0.6t + 1.6$  and  $R_5 = 1$ .

Since resistance values are always positive, we need to restrict our values of  $s$  and  $t$ . We know  $R_3 = s > 0$  and when we combine that with  $R_1 = -0.\bar{7}s + 1.\bar{7} > 0$ , we get  $0 < s < 2.\overline{285714}$ . Similarly,  $R_6 = t > 0$  and with  $R_4 = -0.6t + 1.6 > 0$ , we find  $0 < t < 2.\bar{6}$ .

In order visualize the inequality  $R_2 = 3.5s + 1.5t - 4 > 0$ , we graph the line  $3.5s + 1.5t - 4 = 0$  on the  $ts$ -plane below on the left and shade accordingly. Imposing the additional conditions  $0 < s < 2.\overline{285714}$  and  $0 < t < 2.\bar{6}$ , we find our values of  $s$  and  $t$  restricted to the region  $R$  graphed below on the right.



The region where  $3.5s + 1.5t - 4 > 0$ The region  $R$  for our parameters  $s$  and  $t$ .

Hence, our final answer is:

$$\begin{cases} R_1 = -0.\overline{7}s + 1.\overline{7} \\ R_2 = 3.5s + 1.5t - 4 \\ R_4 = -0.6t + 1.6 \\ R_5 = 1, \end{cases}$$

where  $s$  and  $t$  are pulled from the region  $R = \{(s, t) \mid 0 < s < 2.285714, 0 < t < 2.\overline{6}, 3.5s + 1.5t - 4 > 0\}$ . The reader is encouraged to check that the solution presented in Example 1.4.2, namely all resistance values equal to 1, corresponds to a pair  $(s, t)$  in this region.  $\square$

**1.8.1 Exercises**

In Exercises 1 - 6, sketch the solution to each system of nonlinear inequalities in the plane.

$$1. \begin{cases} x^2 - y^2 \leq 1 \\ x^2 + 4y^2 \geq 4 \end{cases}$$

$$2. \begin{cases} x^2 + y^2 < 25 \\ x^2 + (y - 3)^2 \geq 10 \end{cases}$$

$$3. \begin{cases} (x - 2)^2 + y^2 < 1 \\ x^2 + 4y^2 < 4 \end{cases}$$

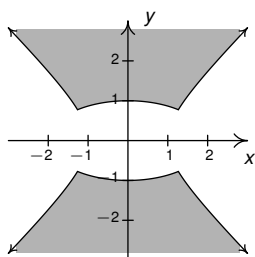
$$4. \begin{cases} y > 10x - x^2 \\ y < x^3 + 8 \end{cases}$$

$$5. \begin{cases} x + 2y^2 > 2 \\ x^2 + 4y^2 \leq 4 \end{cases}$$

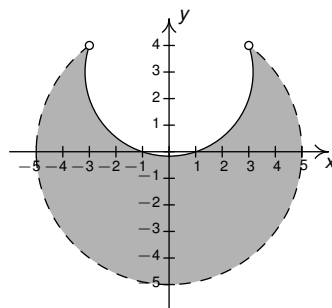
$$6. \begin{cases} x^2 + y^2 \geq 25 \\ y - x \leq 1 \end{cases}$$

## 1.8.2 Answers

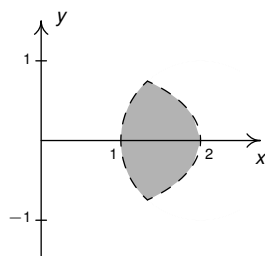
$$1. \begin{cases} x^2 - y^2 \leq 1 \\ x^2 + 4y^2 \geq 4 \end{cases}$$



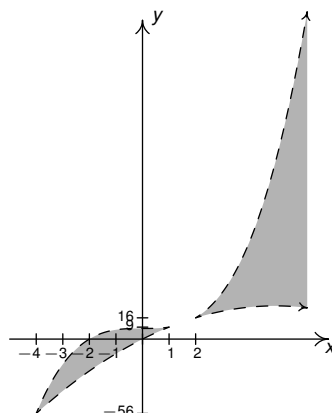
$$2. \begin{cases} x^2 + y^2 < 25 \\ x^2 + (y - 3)^2 \geq 10 \end{cases}$$



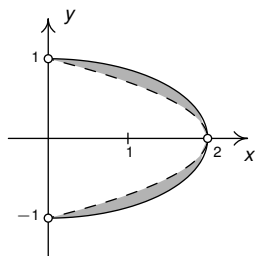
$$3. \begin{cases} (x - 2)^2 + y^2 < 1 \\ x^2 + 4y^2 < 4 \end{cases}$$



$$4. \begin{cases} y > 10x - x^2 \\ y < x^3 + 8 \end{cases}$$



$$5. \begin{cases} x + 2y^2 > 2 \\ x^2 + 4y^2 \leq 4 \end{cases}$$



$$6. \begin{cases} x^2 + y^2 \geq 25 \\ y - x \leq 1 \end{cases}$$

