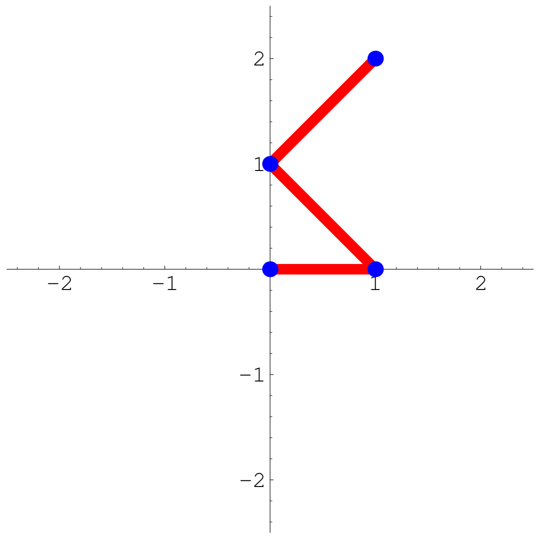
# Math 21b
# Linear Algebra and Differential Equations
# Fall 2004
# Janet Chen

These are the handouts that I used in class while teaching 21b. Most of them consist of examples or problems for the students to work out or discuss. This version of the handouts includes solutions.

# Linear Transformations in Geometry

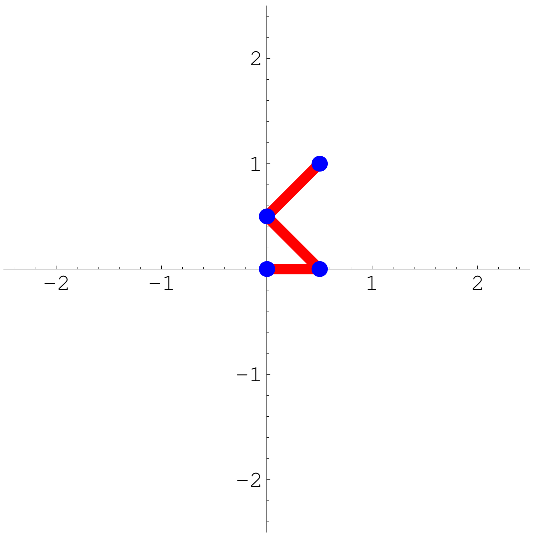
Consider the following figure, which has endpoints $\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.



Draw what happens to the figure after applying the linear transformation $T(\vec{x}) = A\vec{x}$ in each of the following cases. Try to describe the effect of the linear transformation in words. Is the linear transformation invertible?

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

**Solution.**




This linear transformation sends any vector

$$\begin{bmatrix} x \\ y \end{bmatrix}$$

to

$$\begin{bmatrix} x/2 \\ y/2 \end{bmatrix}.$$

Therefore, it shrinks the figure to half of its original size. We call this a "scaling by $\frac{1}{2}$." This transformation is invertible; the inverse matrix is

$$\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix},$$

or scaling by 2.

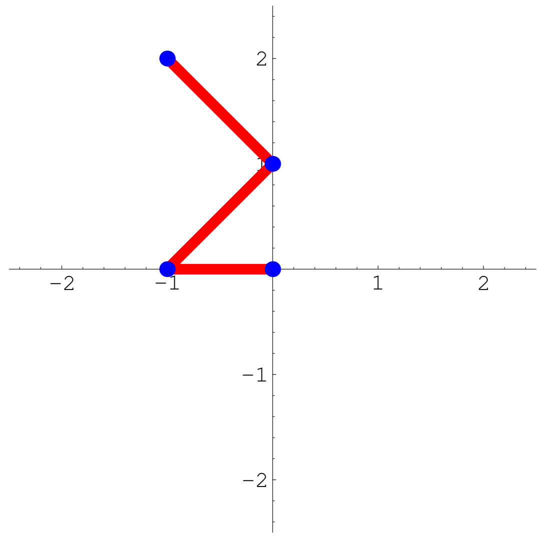
In general, if $k > 0$, the matrix

$$\begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix}$$

is called a "scaling by $k$."

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

**Solution.**




This linear transformation flips the figure over the $y$-axis (like a mirror). We call this transformation "reflection about the $y$-axis." It is invertible because we can simply reflect about the $y$-axis again to get the original figure back. That is, the inverse of

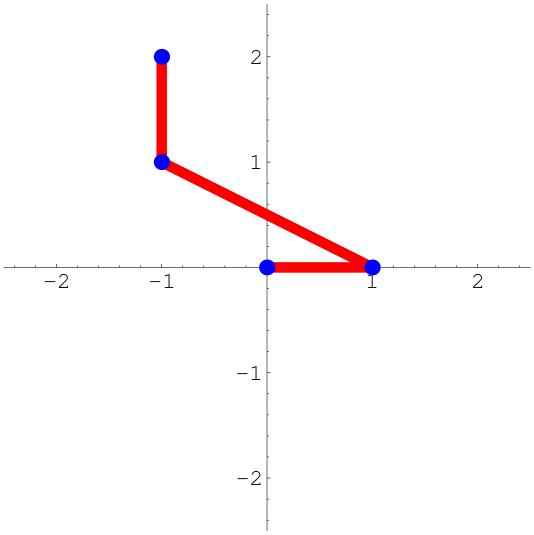
$$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

is just

$$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}.$$

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

**Solution.**




This linear transformation pushes things around horizontally but not vertically. It is called a "horizontal shear." In general, a horizontal shear is given by a matrix of the form

$$\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}.$$

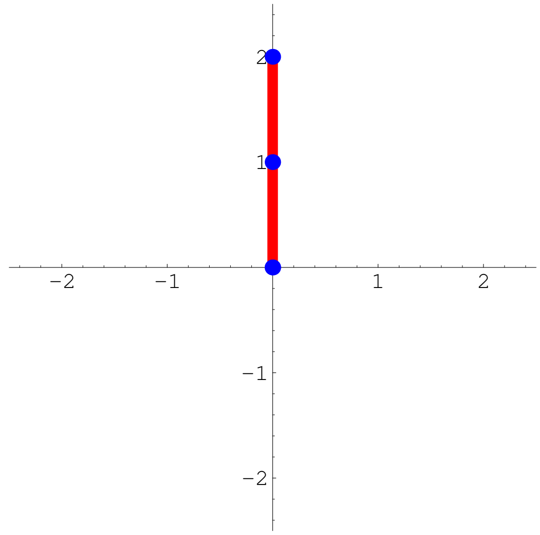
Any horizontal shear is invertible, and its inverse is just a horizontal shear in the opposite direction.

There are also vertical shears, which are given by matrices of the form

$$\begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}.$$

4. $A = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$

**Solution.**



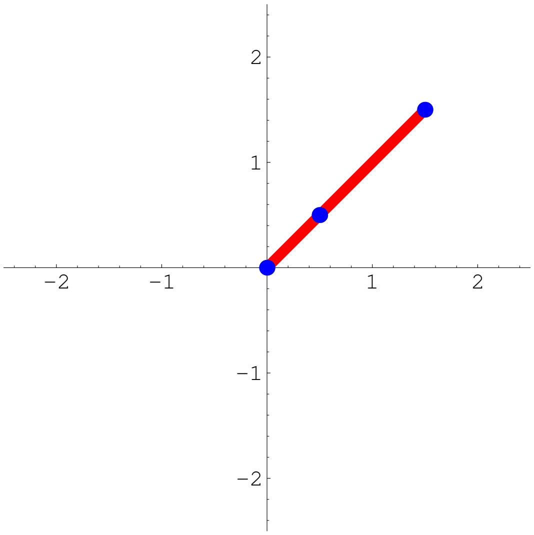

This linear transformation smashes the figure into the $y$-axis. It is called "projection onto the $y$-axis." This linear transformation is not invertible because multiple points may be smashed into the same point. For instance, we can see from the picture that $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} -1 \\ 0 \end{bmatrix}$ are both smashed into $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$. Another way of expressing this is to say that the system

$$A\vec{x} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

has more than one solution.

5. $A = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$

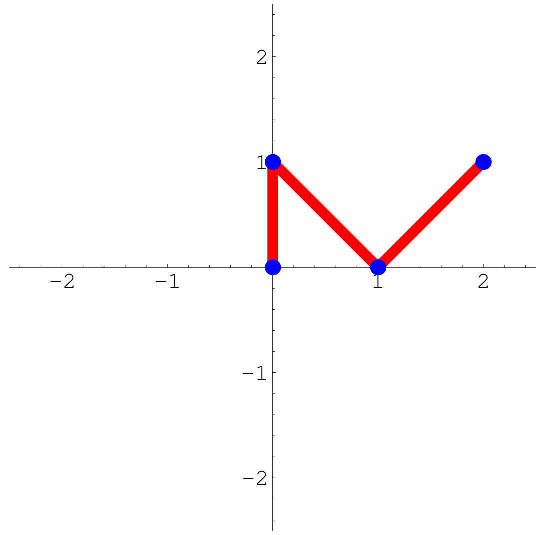
**Solution.**




This linear transformation smashes the figure onto the line $y = x$. We call it "projection onto the line $y = x$." This linear transformation is not invertible because multiple points may be smashed into the same point. For example, $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ are both smashed into $\begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$.

6. $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

**Solution.**




This linear transformation flips the figure over the line $y = x$. We call it "reflection about the line $y = x$." This linear transformation is invertible because, if we reflect over the line $y = x$ again, we get the original figure back. That is, the inverse of

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

is just

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

# Invertible Matrices

Let $A$ be an $n \times m$ matrix ($n$ rows and $m$ columns). Remember that $A$ is said to be invertible if the equation $A\vec{x} = \vec{y}$ has a unique solution for every $\vec{y} \in \mathbb{R}^n$.

1. *Suppose* $\text{rank}(A) < n$. *Is A invertible? That is, is it true that the system $A\vec{x} = \vec{y}$ has a unique solution for every $\vec{y} \in \mathbb{R}^n$? (Hint: what does* $\text{rref}(A)$ *look like?)*

   **Solution.** Saying that $\text{rank}(A) < n$ is the same as saying that the number of leading 1's in $\text{rref}(A)$ is less than the number of rows of $\text{rref}(A)$. That is, there is at least one row in $\text{rref}(A)$ which does not have a leading 1. That row must consist entirely of 0's.

   Therefore, when we row reduce an augmented matrix $\left[\begin{array}{c:c} A & \vec{b} \end{array}\right]$, we will end up with a row that looks like $\left[\begin{array}{ccc:c} 0 & \cdots & 0 & c \end{array}\right]$ for some number $c$. If $c \neq 0$, then the system is inconsistent. Therefore, $A$ is not invertible.

2. *Suppose* $\text{rank}(A) < m$. *Is A invertible?*

   **Solution.** If $\text{rank}(A) < m$, then there is at least one column in $\text{rref}(A)$ which does not contain a leading 1. In other words, any linear system $A\vec{x} = \vec{y}$ has at least one free variable. Such a system has either infinitely many solutions or no solutions. So, $A$ is not invertible.

3. *Finally, suppose* $\text{rank}(A) = n$ *and* $\text{rank}(A) = m$ *(in particular, $n = m$). Is A invertible?*

   **Solution.** In this case, each column has exactly one leading 1, and each row has exactly one leading 1. Since leading 1's must be to the left of any leading 1's in lower rows, $\text{rref}(A)$ must have leading 1's on the diagonal and 0's everywhere else. That is, $\text{rref}(A) = I_n$.

   Therefore, if we are solving a system $A\vec{x} = \vec{y}$, we can reduce the augmented matrix to end up with something that looks like $\left[\begin{array}{c:c} I_n & \vec{b} \end{array}\right]$. Then, $\vec{x} = \vec{b}$ is the unique solution to the system.

   Therefore, every system $A\vec{x} = \vec{y}$ has a unique solution, so $A$ is invertible.

# An Example of Finding Inverses

Remember that, to find the inverse of a matrix $A$, we want to solve the system $A\vec{x} = \vec{y}$ to write $\vec{x}$ in terms of $\vec{y}$. As an example, consider the matrix

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

If we write

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \text{ and } \vec{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix},$$

then the equation $A\vec{x} = \vec{y}$ can be rewritten as

$$\left| \begin{array}{ccccccc} 3x_1 & + & 2x_2 & = & y_1 & & \\ 7x_1 & + & 5x_2 & = & & & y_2 \end{array} \right|.$$

To solve this, we just use the row operations we know (and love!):

$$\left| \begin{array}{ccccccc} 3x_1 & + & 2x_2 & = & y_1 & & \\ 7x_1 & + & 5x_2 & = & & & y_2 \end{array} \right| \begin{array}{l} \div 3 \\ \phantom{x} \end{array} \to \left| \begin{array}{ccccccc} x_1 & + & \frac{2}{3}x_2 & = & \frac{1}{3}y_1 & & \\ 7x_1 & + & 5x_2 & = & & & y_2 \end{array} \right| \begin{array}{l} \phantom{x} \\ -7(\text{I}) \end{array}$$

$$\to \left| \begin{array}{ccccccc} x_1 & + & \frac{2}{3}x_2 & = & \frac{1}{3}y_1 & & \\ & & \frac{1}{3}x_2 & = & -\frac{7}{3}y_1 & + & y_2 \end{array} \right| \begin{array}{l} \phantom{x} \\ \times 3 \end{array}$$

$$\to \left| \begin{array}{ccccccc} x_1 & + & \frac{2}{3}x_2 & = & \frac{1}{3}y_1 & & \\ & & x_2 & = & -7y_1 & + & 3y_2 \end{array} \right| \begin{array}{l} -\frac{2}{3}(\text{II}) \\ \phantom{x} \end{array}$$

$$\to \left| \begin{array}{ccccccc} x_1 & & & = & 5y_1 & - & 2y_2 \\ & & x_2 & = & -7y_1 & + & 3y_2 \end{array} \right|$$

Therefore,

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 5 & -2 \\ -7 & 3 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix},$$

so

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

# Matrix Multiplication Practice

1. *Let* $A = \begin{bmatrix} -2 & 3 \\ 1 & 2 \end{bmatrix}$ *and* $B = \begin{bmatrix} 3 & -1 \\ 3 & -1 \end{bmatrix}$. *Find* $AB$ *and* $BA$ *(if they make sense).*

   **Solution.** $AB = \begin{bmatrix} 3 & -1 \\ 9 & -3 \end{bmatrix}$ and $BA = \begin{bmatrix} -7 & 7 \\ -7 & 7 \end{bmatrix}$. Notice that it is not true that $AB = BA$.

   Also, observe that $A$ has rank 2, $B$ has rank 1, and $AB$ and $BA$ both have rank 1. When we talk about images, kernels, and dimension next week, you should be able to explain why this is true.

2. *Let* $A = \begin{bmatrix} 1 & 2 \end{bmatrix}$ *and* $B = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$. *Find* $AB$ *and* $BA$ *(if they make sense).*

   **Solution.** $AB = [11]$ and $BA = \begin{bmatrix} 3 & 6 \\ 4 & 8 \end{bmatrix}$.

3. *Let* $A = \begin{bmatrix} 1 & 4 \\ -3 & 1 \\ 0 & 0 \end{bmatrix}$ *and* $B = \begin{bmatrix} 0 & 2 & 1 \end{bmatrix}$. *Find* $AB$ *and* $BA$ *(if they make sense).*

   **Solution.** $BA = \begin{bmatrix} -6 & 2 \end{bmatrix}$, but $AB$ is not defined.

4. *Let* $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ *and* $B = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$. *Find* $AB$, $(AB)^{-1}$, *and* $A^{-1}B^{-1}$.

   **Solution.** $AB = \begin{bmatrix} -1 & -1 \\ -1 & 0 \end{bmatrix}$. To find $(AB)^{-1}$, we want to row-reduce the matrix

$$\left[\begin{array}{cc:cc} -1 & -1 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{array}\right].$$

Here are the steps:

$$\left[\begin{array}{cc:cc} -1 & -1 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{array}\right] \times(-1) \quad \to \quad \left[\begin{array}{cc:cc} 1 & 1 & -1 & 0 \\ -1 & 0 & 0 & 1 \end{array}\right] +(\mathrm{I})$$
$$\to \left[\begin{array}{cc:cc} 1 & 1 & -1 & 0 \\ 0 & 1 & -1 & 1 \end{array}\right] -(\mathrm{II})$$
$$\to \left[\begin{array}{cc:cc} 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 1 \end{array}\right]$$

Thus, $(AB)^{-1} = \begin{bmatrix} 0 & -1 \\ -1 & 1 \end{bmatrix}$. To find $A^{-1}B^{-1}$, we first need to find $A^{-1}$ and $B^{-1}$. We could do this using the row reduction method above, or we could observe that $A$ is a horizontal shear and $B$ is a reflection. Therefore, $A^{-1}$ is a horizontal shear in the opposite direction and $B^{-1} = B$. So,

$$A^{-1}B^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 0 \end{bmatrix}.$$

Notice that $(AB)^{-1} \neq A^{-1}B^{-1}$! Instead, $(AB)^{-1} = B^{-1}A^{-1}$; intuitively, if $\vec{x}$ is a vector, then $AB\vec{x}$ is the vector obtained by first reflecting $\vec{x}$ over a line, then applying a horizontal shear. To get $\vec{x}$ back, we should first undo the shear and then undo the reflection.

In fact, it is always true that $(AB)^{-1} = B^{-1}A^{-1}$. Can you see why?

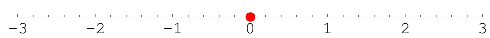
# Solutions of Systems

1. Define a linear transformation $T$ from $\mathbb{R}$ to $\mathbb{R}^2$ by

$$T(x) = \begin{bmatrix} 3x \\ 2x \end{bmatrix}.$$

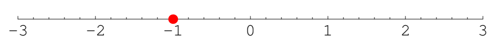
(a) *Draw the set of $x \in \mathbb{R}$ such that $T(x) = \vec{0}$. (This set is called the kernel of $T$.)*

**Solution.** The only solution of $T(x) = \vec{0}$ is $x = 0$.




(b) *Draw the set of $x \in \mathbb{R}$ such that $T(x) = \begin{bmatrix} -3 \\ -2 \end{bmatrix}$.*

**Solution.** The only solution is $x = -1$. Notice that, like in (1a), the solution set is a point.

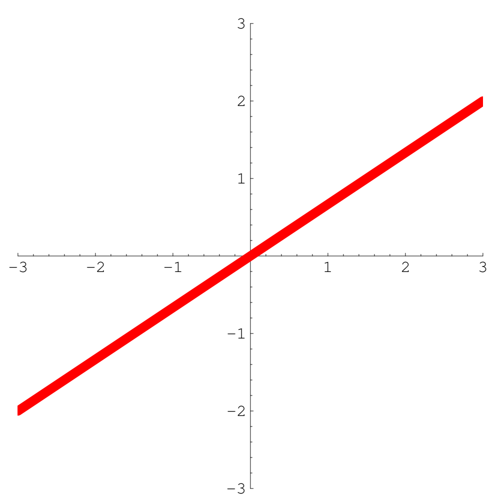



(c) *Draw the set of $x \in \mathbb{R}$ such that $T(x) = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$.*

**Solution.** There are no $x \in \mathbb{R}$ such that $T(x) = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$.

(d) *Draw the set of $\vec{y} \in \mathbb{R}^2$ such that $T(x) = \vec{y}$ has at least one solution. (This set is called the image of $T$.)*

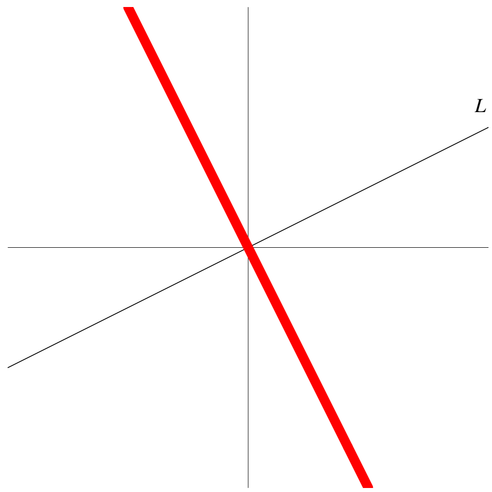
**Solution.** If $T(x) = \vec{y}$ has a solution $x = t$, then that means $\vec{y} = \begin{bmatrix} 3t \\ 2t \end{bmatrix}$. The set of all vectors of the form $\begin{bmatrix} 3t \\ 2t \end{bmatrix}$ is the line $y = \frac{2}{3}x$.



2. Let $L$ be a line through the origin. Remember that we defined a linear transformation $\text{proj}_L$ from $\mathbb{R}^2$ to $\mathbb{R}^2$.

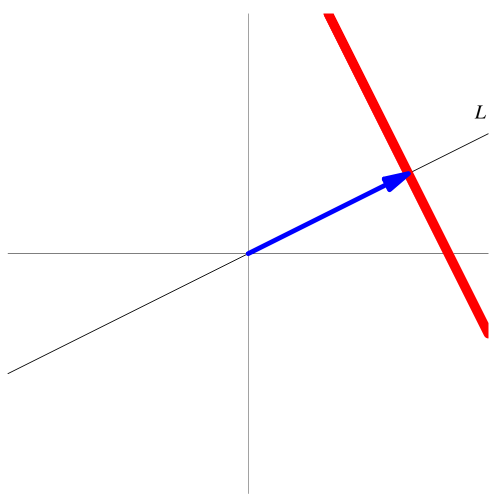
(a) *Draw the set of* $\vec{x} \in \mathbb{R}^2$ *such that* $\text{proj}_L(\vec{x}) = \vec{0}$. *(This is the kernel of* $\text{proj}_L$*.)*

**Solution.** Remember that $\text{proj}_L(\vec{x})$ is defined to be the unique vector such that $\vec{x} - \text{proj}_L(\vec{x})$ is perpendicular to $L$. Thus, $\text{proj}_L(\vec{x}) = 0$ if and only if $\vec{x}$ is perpendicular to $L$. So, the solutions of $\text{proj}_L(\vec{x}) = \vec{0}$ form a line perpendicular to $L$ going through the origin.




(b) *In the picture below, a vector* $\vec{y}$ *is drawn. Draw the set of* $\vec{x} \in \mathbb{R}^2$ *such that* $\text{proj}_L(\vec{x}) = \vec{y}$.

**Solution.** Any vector $\vec{x}$ can be written uniquely as $\vec{x} = \text{proj}_L(\vec{x}) + \vec{x}^{\perp}$ where $\vec{x}^{\perp}$ is a vector perpendicular to $L$. Therefore, saying that $\text{proj}_L(\vec{x}) = \vec{y}$ is the same as saying that $\vec{x} = \vec{y} + \vec{x}^{\perp}$ for some vector $\vec{x}^{\perp}$ which is perpendicular to $L$. The set of such $\vec{x}$ form a line perpendicular to $L$ which includes the tip of $\vec{y}$.

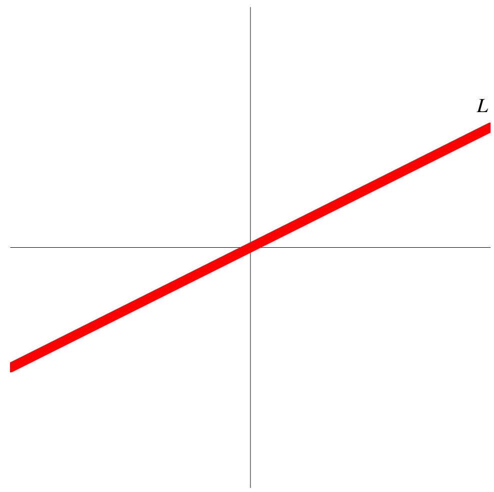



(c) *In the picture below, a vector* $\vec{y}$ *is drawn. Draw the set of* $\vec{x} \in \mathbb{R}^2$ *such that* $\text{proj}_L(\vec{x}) = \vec{y}$.

**Solution.** By definition, $\text{proj}_L(\vec{x})$ must be parallel to the line $L$. Since $\vec{y}$ is not parallel to $L$, there are no $\vec{x} \in \mathbb{R}^2$ satisfying $\text{proj}_L(\vec{x}) = \vec{y}$.

(d) *Draw the set of* $\vec{y} \in \mathbb{R}^2$ *such that* $\text{proj}_L(\vec{x}) = \vec{y}$ *has at least one solution. (This is the image of* $\text{proj}_L$*.)*

**Solution.** As we saw in (2b), if $\vec{y}$ is parallel to $L$, then $\text{proj}_L(\vec{x}) = \vec{y}$ has at least one solution (in fact, the solutions form a line perpendicular to $L$). On the other hand, we saw in (2c) that, if $\vec{y}$ is not parallel to $L$, then $\text{proj}_L(\vec{x}) = \vec{y}$ has no solutions. So, the set of $\vec{y} \in \mathbb{R}^2$ such that $\text{proj}_L(\vec{x}) = \vec{y}$ has at least one solution is exactly the set of $\vec{y}$ parallel to $L$.




3. *If* $T$ *is any linear transformation from* $\mathbb{R}^m$ *to* $\mathbb{R}^n$ *and* $\vec{y}$ *is any vector in* $\mathbb{R}^n$*, how does the solution set of* $T(\vec{x}) = \vec{y}$ *relate to the solution set of* $T(\vec{x}) = \vec{0}$*?*

**Solution.** First, here is some terminology. The set of $\vec{x} \in \mathbb{R}^m$ such that $T(\vec{x}) = \vec{0}$ is called the kernel of $T$ and is written $\ker T$. The set of $\vec{y}$ in $\mathbb{R}^n$ such that $T(\vec{x}) = \vec{y}$ has at least one solution is called the image of $T$ and is written $\text{im}\, T$.

As we can guess from the first two problems, there are two possibilities: either $T(\vec{x}) = \vec{y}$ has no solutions (that is, $\vec{y} \notin \text{im}\, T$), or the solution set of $T(\vec{x}) = \vec{y}$ looks like $\ker T$ (for instance, in problem 1, the solution sets were points; in problem 2, they were lines perpendicular to $L$). In fact, this is exactly what one of your previous homework problems showed! In §1.3, # 48, you showed the following: if $T(\vec{x}) = \vec{b}$ has a solution $\vec{x}_1$ (in particular, $\vec{b} \in \text{im}\, T$), then the solutions of $T(\vec{x}) = \vec{b}$ are exactly the vectors $\vec{x}_1 + \vec{x}_h$ where $\vec{x}_h \in \ker T$. That is, the solution set of $T(\vec{x}) = \vec{b}$ is obtained by translating $\ker T$ by the vector $\vec{x}_1$.

# Subspace and Basis Examples

1. *Let* $A = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \end{bmatrix}, \vec{b} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$*. Is the set of solutions of* $A\vec{x} = \vec{b}$ *a subspace of* $\mathbb{R}^3$*?*

   **Solution.** The set of solutions of $A\vec{x} = \vec{b}$ is not a subspace of $\mathbb{R}^3$. The easiest way to see this is to remember that any subspace of $\mathbb{R}^3$ must contain the zero vector

   $$\vec{0} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

   Since $\vec{0}$ is not a solution of $A\vec{x} = \vec{b}$, the set of solutions cannot be a subspace.

   Of course, we can also find the set of solutions of $A\vec{x} = \vec{b}$ explicitly. The augmented matrix of the system is just

   $$\left[\begin{array}{ccc:c} 1 & 0 & 3 & 2 \\ 0 & 1 & 0 & 4 \end{array}\right].$$

   This is already in reduced row-echelon form. The corresponding equations are $x_1 + 3x_3 = 2$ and $x_2 = 4$. In particular, $x_3$ is a free variable, so it can be any real number $t$. Then, $x_1 = 2 - 3t$, so the solutions are

   $$\vec{x} = \begin{bmatrix} 2-3t \\ 4 \\ t \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 0 \end{bmatrix} + t \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}.$$

   Geometrically, the solutions form a line which does not pass through the origin. Since any subspace must contain the origin, this line is not a subspace.

2. *Let* $\vec{v}_1 = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} 2 \\ 7 \\ 4 \end{bmatrix}$*. Are any of these vectors redundant?*

   **Solution.** $\vec{v}_1$ is non-zero, so it is not redundant. $\vec{v}_2$ is not a multiple of $\vec{v}_1$, so it does not lie in the span of $\vec{v}_1$. Therefore, $\vec{v}_2$ is not redundant either.

   To check whether $\vec{v}_3$ is redundant, we want to know if $\vec{v}_3$ lies in the span of $\vec{v}_1$ and $\vec{v}_2$; that is, we want to know if we can write $\vec{v}_3 = x_1\vec{v}_1 + x_2\vec{v}_2$ for some constants $x_1, x_2$. That is, we are wondering whether the system

   $$\left| \begin{array}{ccccc} x_1 & & & = & 2 \\ 3x_1 & + & x_2 & = & 7 \\ x_1 & + & 2x_2 & = & 4 \end{array} \right|$$

   has a solution. We can rewrite this system as

   $$\begin{bmatrix} 1 & 0 \\ 3 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 7 \\ 4 \end{bmatrix}.$$

   To solve the system, we can use Gauss-Jordan elimination.

   $$\left[\begin{array}{cc:c} 1 & 0 & 2 \\ 3 & 1 & 7 \\ 1 & 2 & 4 \end{array}\right] \begin{array}{c} \\ -3(\mathrm{I}) \\ -(\mathrm{I}) \end{array} \rightarrow \left[\begin{array}{cc:c} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 2 & 2 \end{array}\right] \begin{array}{c} \\ \\ -2(\mathrm{II}) \end{array} \rightarrow \left[\begin{array}{cc:c} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{array}\right].$$

   This system has at least one solution, so $\vec{v}_3$ must be redundant. In fact, the solution of the system is

   $$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix},$$

   so we can write $\vec{v}_3 = 2\vec{v}_1 + \vec{v}_2$.

3. *Let* $A = \begin{bmatrix} 1 & 0 & 2 \\ 3 & 1 & 7 \\ 1 & 2 & 4 \end{bmatrix}$*. Find a basis of* $\operatorname{im} A$.

**Solution.** We know that $\operatorname{im} A$ is spanned by the columns of $A$. That is, if we let

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} 2 \\ 7 \\ 4 \end{bmatrix},$$

then $\vec{v}_1, \vec{v}_2, \vec{v}_3$ span $\operatorname{im} A$. We saw in the previous problem that $\vec{v}_3$ is redundant; that is, it adds nothing to the span. So, $\vec{v}_1$ and $\vec{v}_2$ span $\operatorname{im} A$. Since we showed in the previous problem that $\vec{v}_1$ and $\vec{v}_2$ are linearly independent, this means that $\vec{v}_1$ and $\vec{v}_2$ form a basis of $\operatorname{im} A$.

4. *Let* $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_6$ *be vectors in* $\mathbb{R}^8$ *such that* $3\vec{v}_1 - 2\vec{v}_2 + 4\vec{v}_4 + 5\vec{v}_5 = \vec{0}$*. Explain why* $\vec{v}_1, \ldots, \vec{v}_6$ *must be linearly dependent.*

**Solution.** Since $3\vec{v}_1 - 2\vec{v}_2 + 4\vec{v}_4 + 5\vec{v}_5 = \vec{0}$, we can write $5\vec{v}_5 = -3\vec{v}_1 + 2\vec{v}_2 - 4\vec{v}_4$, so $\vec{v}_5 = -\frac{3}{5}\vec{v}_1 + \frac{2}{5}\vec{v}_2 - \frac{4}{5}\vec{v}_4$. In other words, $\vec{v}_5$ lies in the span of $\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4$, so $\vec{v}_5$ is redundant. Therefore, the vectors $\vec{v}_1, \ldots, \vec{v}_6$ must be linearly dependent.

5. *Let* $V$ *be a subspace of* $\mathbb{R}^n$*, and suppose we know that* $\vec{v}_1, \ldots, \vec{v}_5$ *form a basis of* $V$*. Is it possible that* $\vec{v}_1 + 2\vec{v}_2 - \vec{v}_4 = \vec{v}_3 + 2\vec{v}_4 + \vec{v}_5$*?*

**Solution.** No, it is not possible. After all, if $\vec{v}_1 + 2\vec{v}_2 - \vec{v}_4 = \vec{v}_3 + 2\vec{v}_4 + \vec{v}_5$, then we could solve for $\vec{v}_5$ to get $\vec{v}_5 = \vec{v}_1 + 2\vec{v}_2 - \vec{v}_3 - 3\vec{v}_4$. But this would mean that $\vec{v}_5$ was redundant, which is impossible since we said $\vec{v}_1, \ldots, \vec{v}_5$ form a basis of $V$.

6. *Let* $A = \begin{bmatrix} 1 & 0 & 2 \\ 3 & 1 & 7 \\ 1 & 2 & 4 \end{bmatrix}$*, like in Problem 3. Find* $\ker A$*. How can you use your answer to determine whether the columns of* $A$ *are linearly independent?*

**Solution.** To find $\ker A$, we want to solve the system $A\vec{x} = \vec{0}$. We can do this using Gauss-Jordan elimination:

$$\left[\begin{array}{ccc|c} 1 & 0 & 2 & 0 \\ 3 & 1 & 7 & 0 \\ 1 & 2 & 4 & 0 \end{array}\right] \begin{array}{l} \\ -3(\text{I}) \\ -(\text{I}) \end{array} \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 2 & 2 & 0 \end{array}\right] \begin{array}{l} \\ \\ -2(\text{II}) \end{array}$$

$$\rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

(Notice that the far right column always consists entirely of 0's when we're finding the kernel, so we don't really need to write this column.)

We end up with the equations

$$\begin{aligned} x_1 + 2x_3 &= 0 \\ x_2 + x_3 &= 0 \end{aligned}$$

The leading variables are $x_1$ and $x_2$, while $x_3$ is a free variable. Thus, $x_3$ can have any value $t$. We can solve for $x_1$ and $x_2$ in terms of the free variable, and we end up with $x_1 = -2t$, $x_2 = -t$. Thus, the solutions of the system $A\vec{x} = \vec{0}$ are all vectors of the form

$$\begin{bmatrix} -2t \\ -t \\ t \end{bmatrix} = t \begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix}.$$

Now, how does this relate to the question of whether the columns of $A$ are linearly independent? Let's write the columns of $A$ as $\vec{v}_1$, $\vec{v}_2$, and $\vec{v}_3$. Since

$$\begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix}$$

is in the kernel of $A$, we know that

$$[\vec{v}_1 \ \vec{v}_2 \ \vec{v}_3] \begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix} = 0.$$

This is exactly the same as saying that $-2\vec{v}_1 - \vec{v}_2 + \vec{v}_3 = \vec{0}$ (if you aren't convinced, multiply it out to see why!). But, as we saw in Problem 4, having an equation of this form means that $\vec{v}_1$, $\vec{v}_2$, and $\vec{v}_3$ must be linearly dependent; after all, we can solve for $\vec{v}_3$ to write $\vec{v}_3 = 2\vec{v}_1 + \vec{v}_2$, which shows that $\vec{v}_3$ is redundant.

To summarize, if $\vec{v}_1, \ldots, \vec{v}_r$ are vectors in $\mathbb{R}^n$, we can make a matrix $A = [\vec{v}_1 \ \cdots \ \vec{v}_r]$. If $\ker A$ contains a nonzero vector, then $\vec{v}_1, \ldots, \vec{v}_r$ must be linearly dependent.

# The Rank - Nullity Theorem

Let $A = \begin{bmatrix} 1 & 1 & 1 & 0 & -1 \\ 2 & 1 & 4 & -1 & 2 \\ -1 & 2 & -7 & 0 & 4 \\ 3 & 0 & 9 & -1 & -1 \end{bmatrix}$. You (should) know by now how to find $\operatorname{rref}(A)$, so I'll just tell you what it is: $\operatorname{rref}(A) = \begin{bmatrix} 1 & 0 & 3 & 0 & -2 \\ 0 & 1 & -2 & 0 & 1 \\ 0 & 0 & 0 & 1 & -5 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$. Let $\vec{v}_1, \ldots, \vec{v}_5$ be the columns of $A$.

1. *Find a basis of* $\ker A$. *What is the dimension of* $\ker A$*? How does the dimension of* $\ker A$ *relate to the rank of* $A$*? (Note: the dimension of* $\ker A$ *is called the nullity of* $A$*.)*

   **Solution.** Remember that $\ker A$ is just the set of solutions of $A\vec{x} = \vec{0}$. To solve this system, we row-reduce the augmented matrix $\left[\begin{array}{c|c} A & 0 \end{array}\right]$. No matter what row operations we do, the last column will always consist entirely of 0's. So, when we row-reduce the augmented matrix, we will just end up with $\left[\begin{array}{c|c} \operatorname{rref}(A) & 0 \end{array}\right]$, or

$$\left[\begin{array}{ccccc|c} 1 & 0 & 3 & 0 & -2 & 0 \\ 0 & 1 & -2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & -5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

   If we call our variables $x_1, \ldots, x_5$, the corresponding equations are

$$\begin{aligned} x_1 + 3x_3 - 2x_5 &= 0 \\ x_2 - 2x_3 + x_5 &= 0 \\ x_4 - 5x_5 &= 0 \end{aligned}$$

   The leading variables are $x_1$, $x_2$, and $x_4$, so the free variables are $x_3$ and $x_5$. If we let $x_3 = s$ and $x_5 = t$, then $x_1 = -3s + 2t$, $x_2 = 2s - t$, and $x_4 = 5t$. Therefore, $\ker A$ consists of all vectors of the form

$$\begin{bmatrix} -3s + 2t \\ 2s - t \\ s \\ 5t \\ t \end{bmatrix} = s \begin{bmatrix} -3 \\ 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 2 \\ -1 \\ 0 \\ 5 \\ 1 \end{bmatrix}.$$

   Thus, the vectors

$$\vec{u}_1 = \begin{bmatrix} -3 \\ 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \vec{u}_2 = \begin{bmatrix} 2 \\ -1 \\ 0 \\ 5 \\ 1 \end{bmatrix}$$

   span $\ker A$. To check that they are linearly independent, we use Summary 3.29, part iv on page 120 of the book. Suppose $s\vec{u}_1 + t\vec{u}_2 = \vec{0}$. Then,

$$\begin{bmatrix} -3s + 2t \\ 2s - t \\ s \\ 5t \\ t \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

   so $s = t = 0$. Thus, $\vec{u}_1$ and $\vec{u}_2$ form a basis of $\ker A$ (they span $\ker A$ and are linearly independent).

Notice that we get one basis vector of $\ker A$ for each free variable in the system. The number of free variables is the number of columns minus the number of leading variables, or $5 - \operatorname{rank}(A)$. Thus, the dimension of $\ker A$ is $5 - \operatorname{rank}(A)$.

2. For each basis vector $\vec{c} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \end{bmatrix}$ that you found in problem #1:

   (a) *Find $c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3 + c_4\vec{v}_4 + c_5\vec{v}_5$. (Why does your answer make sense?)*

   **Solution.** Let's first look at $\vec{c} = \vec{u}_1$, the first basis vector we found in part 1. Then, $c_1\vec{v}_1 + c_2\vec{v_2} + c_3\vec{v}_3 + c_4\vec{v}_4 + c_5\vec{v_5} = -3\vec{v}_1 + 2\vec{v}_2 + \vec{v}_3 = \vec{0}$. This makes sense because $c_1\vec{v}_1 + c_2\vec{v_2} + c_3\vec{v}_3 + c_4\vec{v}_4 + c_5\vec{v_5}$ is exactly the product $A\vec{c}$; since $\vec{c} \in \ker A$, $A\vec{c}$ must be $\vec{0}$.

   By the same reasoning, if we use $\vec{c} = \vec{u}_2$, we see that $2\vec{v}_1 - \vec{v}_2 + 5\vec{v}_4 + \vec{v}_5 = A\vec{u}_2 = \vec{0}$.

   (b) *Use your answer from #2b to find a redundant vector among $\vec{v}_1, \ldots, \vec{v}_5$.*

   **Solution.** We saw in part (a) that $-3\vec{v}_1 + 2\vec{v}_2 + \vec{v}_3 = \vec{0}$ and $2\vec{v}_1 - \vec{v}_2 + 5\vec{v}_4 + \vec{v}_5 = 0$. We can rewrite the equation $-3\vec{v}_1 + 2\vec{v}_2 + \vec{v}_3 = \vec{0}$ as $\vec{v}_3 = 3\vec{v}_1 - 2\vec{v}_2$, which shows that $\vec{v}_3$ is redundant. Similarly, we can rewrite the equation $2\vec{v}_1 - \vec{v}_2 + 5\vec{v}_4 + \vec{v}_5 = 0$ as $\vec{v}_5 = -2\vec{v}_1 + \vec{v}_2 - 5\vec{v}_4$, which shows that $\vec{v}_5$ is redundant.

3. *After throwing away the redundant vectors you found in #2, are the remaining $\vec{v}_i$ linearly independent? Why or why not?*

   **Solution.** The remaining $\vec{v}_i$ are $\vec{v}_1$, $\vec{v}_2$, and $\vec{v}_4$ (notice that these correspond to the columns of $\operatorname{rref}(A)$ which contain leading ones). Let $B = [\vec{v}_1\ \vec{v}_2\ \vec{v}_4]$. By Summary 3.29, part v on page 120 of the book, $\vec{v}_1$, $\vec{v}_2$, and $\vec{v}_4$ are linearly independent if and only if $\ker B = \{0\}$. To find $\ker B$, we want to solve the system $B\vec{x} = \vec{0}$. We do this by row-reducing the augmented matrix $\left[\begin{array}{c:c} B & \vec{0} \end{array}\right]$. As we explained in #1, we will end up with $\left[\begin{array}{c:c} \operatorname{rref}(B) & \vec{0} \end{array}\right]$, so we just need to find $\operatorname{rref}(B)$.

   We could row-reduce $B$ to find $\operatorname{rref}(B)$, but there's an easier way. To get from $A$ to $\operatorname{rref}(A)$, we do a series of elementary row operations. If we apply this same series of elementary row operations to $B$, we will end up with the 1st, 2nd, and 4th columns of $\operatorname{rref}(A)$ (if you're not convinced, try it!). So, $\operatorname{rref}(B)$ just consists of the 1st, 2nd, and 4th columns of $\operatorname{rref}(A)$:

$$\operatorname{rref}(B) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

   In particular, the only solution of the system with augmented matrix $\left[\begin{array}{c:c} \operatorname{rref}(B) & \vec{0} \end{array}\right]$ is $\vec{0}$, which means $\ker B = \{0\}$. So, $\vec{v}_1$, $\vec{v_2}$, and $\vec{v}_4$ are linearly independent.

4. *Find a basis of* $\operatorname{im} A$. *What is the dimension of* $\operatorname{im} A$? *How does the dimension of* $\operatorname{im} A$ *relate to the rank of A?*

   **Solution.** We know that the columns of $A$ span $\operatorname{im} A$. Since $\vec{v}_3$ and $\vec{v}_5$ (the columns corresponding to columns with free variables) are redundant and $\vec{v}_1$, $\vec{v}_2$, and $\vec{v}_4$ (the columns corresponding to columns with leading variables) are linearly independent, a basis of $\operatorname{im} A$ is given by $\vec{v}_1$, $\vec{v}_2$, and $\vec{v}_4$. These correspond to the columns of $\operatorname{rref}(A)$ with leading ones. In particular, the dimension of $\operatorname{im} A$ is the number of columns of $\operatorname{rref}(A)$ with leading ones, or $\operatorname{rank}(A)$ (which is 3 in this case).

# Coordinates

1. Let $\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\vec{v}_2 = \begin{bmatrix} 6 \\ -3 \end{bmatrix}$. Then, $\mathfrak{B} = (\vec{v}_1, \vec{v}_2)$ is a basis of $\mathbb{R}^2$. Let $L$ be the line $y = 2x$.

   (a) *Find* $[\vec{e}_1]_{\mathfrak{B}}$.

   **Solution.** We want to write $\vec{e}_1$ as $c_1\vec{v}_1 + c_2\vec{v}_2$. That is, we want to solve the system

   $$\begin{bmatrix} 1 & 6 \\ 2 & -3 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

   for $c_1, c_2$. Using Gauss-Jordan elimination, we have

   $$\left[\begin{array}{cc:c} 1 & 6 & 1 \\ 2 & -3 & 0 \end{array}\right] \begin{array}{c} \\ -2(\mathrm{I}) \end{array} \rightarrow \left[\begin{array}{cc:c} 1 & 6 & 1 \\ 0 & -15 & -2 \end{array}\right] \begin{array}{c} \\ \div(-15) \end{array}$$

   $$\rightarrow \left[\begin{array}{cc:c} 1 & 6 & 1 \\ 0 & 1 & 2/15 \end{array}\right] \begin{array}{c} -6(\mathrm{II}) \\ \end{array}$$

   $$\rightarrow \left[\begin{array}{cc:c} 1 & 0 & 1/5 \\ 0 & 1 & 2/15 \end{array}\right]$$

   Thus, $c_1 = \frac{1}{5}$ and $c_2 = \frac{2}{15}$, so $\vec{e}_1 = \frac{1}{5}\vec{v}_1 + \frac{2}{15}\vec{v}_2$, and $[\vec{e}_1]_{\mathfrak{B}} = \begin{bmatrix} 1/5 \\ 2/15 \end{bmatrix}$.

   (b) *Find a matrix $S$ such that $\vec{x} = S[\vec{x}]_{\mathfrak{B}}$ for all $\vec{x} \in \mathbb{R}^2$.*

   **Solution.** Observe that

   $$\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, [\vec{v}_1]_{\mathfrak{B}} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} 6 \\ -3 \end{bmatrix}, [\vec{v}_2]_{\mathfrak{B}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

   Therefore, if we let $S$ be the matrix

   $$S = \begin{bmatrix} 1 & 6 \\ 2 & -3 \end{bmatrix},$$

   then $\vec{v}_1 = S[\vec{v}_1]_{\mathfrak{B}}$ and $\vec{v}_2 = S[\vec{v}_2]_{\mathfrak{B}}$. Since $\vec{v}_1$ and $\vec{v}_2$ form a basis of $\mathbb{R}^2$, it follows that $\vec{x} = S[\vec{x}]_{\mathfrak{B}}$ for any $\vec{x} \in \mathbb{R}^2$.

   (c) Find the $\mathfrak{B}$-matrix of $\mathrm{proj}_L$.

   **Solution.** Since $\vec{v}_1$ lies on the line $L$, $\mathrm{proj}_L(\vec{v}_1) = \vec{v}_1 = 1 \cdot \vec{v}_1 + 0 \cdot \vec{v}_2$. Since $\vec{v}_2$ is perpendicular to $L$, $\mathrm{proj}_L(\vec{v}_2) = \vec{0} = 0 \cdot \vec{v}_1 + 0 \cdot \vec{v}_2$. In $\mathfrak{B}$-coordinates, we write

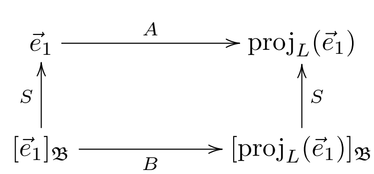
   $$[\mathrm{proj}_L(\vec{v}_1)]_{\mathfrak{B}} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ and } [\mathrm{proj}_L(\vec{v}_2)]_{\mathfrak{B}} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

   Therefore, the $\mathfrak{B}$-matrix of $\mathrm{proj}_L$ is

   $$B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

   (d) Use the $\mathfrak{B}$-matrix of $\mathrm{proj}_L$ to find the standard matrix $A$ of $\mathrm{proj}_L$.

   **Solution.** To write the standard matrix $A$ of $\mathrm{proj}_L$, we want to find $\mathrm{proj}_L(\vec{e}_1)$ and $\mathrm{proj}_L(\vec{e}_2)$. We can organize information into a commutative diagram:



(Here, $B$ is the matrix we found in part (b).) The meaning of the vertical arrows is this: we showed in part (b) that $\vec{x} = S[\vec{x}]_{\mathfrak{B}}$ for any $\vec{x} \in \mathbb{R}^2$. In particular, $\vec{e}_1 = S[\vec{e}_1]_{\mathfrak{B}}$ and $\text{proj}_L(\vec{e}_1) = S[\text{proj}_L(\vec{e}_1]_{\mathfrak{B}}$. The top horizontal arrow indicates that $\text{proj}_L(\vec{e}_1) = A\vec{e}_1$ (since $A$ is the standard matrix of $\text{proj}_L$). The bottom horizontal arrow means that $[\text{proj}_L(\vec{e}_1)]_{\mathfrak{B}} = B[\vec{e}_1]_{\mathfrak{B}}$; this is true since $B$ is the $\mathfrak{B}$-matrix of $\text{proj}_L$.

By part (a), we know that

$$[\vec{e}_1]_{\mathfrak{B}} = \begin{bmatrix} 1/5 \\ 2/15 \end{bmatrix}.$$

In other words, $\vec{e}_1 = \frac{1}{5}\vec{v}_1 + \frac{2}{15}\vec{v}_2$. Using our diagram, we see that

$$[\text{proj}_L(\vec{e}_1)]_{\mathfrak{B}} = B[\vec{e}_1]_{\mathfrak{B}} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/5 \\ 2/15 \end{bmatrix} = \begin{bmatrix} 1/5 \\ 0 \end{bmatrix}.$$

That is, $\text{proj}_L(\vec{e}_1) = \frac{1}{5}\vec{v}_1$. Finally, we want to write $\text{proj}_L(\vec{e}_1)$ in standard coordinates. We have

$$\text{proj}_L(\vec{e}_1) = S[\text{proj}_L(\vec{e}_1)_{\mathfrak{B}}] = \begin{bmatrix} 1 & 6 \\ 2 & -3 \end{bmatrix} \begin{bmatrix} 1/5 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/5 \\ 2/5 \end{bmatrix}.$$

Thus, we have found the first column of the standard matrix of $\text{proj}_L$.

Using the exact same argument for $\vec{e}_2$, we find that $[\vec{e}_2]_{\mathfrak{B}} = \begin{bmatrix} 2/5 \\ -1/15 \end{bmatrix}$, $[\text{proj}_L(\vec{e}_2)]_{\mathfrak{B}} = \begin{bmatrix} 2/5 \\ 0 \end{bmatrix}$, and $\text{proj}_L(\vec{e}_2) = \begin{bmatrix} 2/5 \\ 4/5 \end{bmatrix}$. Therefore, the standard matrix of $\text{proj}_L$ is

$$\begin{bmatrix} 1/5 & 2/5 \\ 2/5 & 4/5 \end{bmatrix}.$$

2. *Let* $\vec{v}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ *and* $\vec{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. *Then,* $\mathfrak{B} = (\vec{v}_1, \vec{v}_2)$ *is a basis of* $\mathbb{R}^2$. *Let* $A = \begin{bmatrix} 3 & 4 \\ 1 & 0 \end{bmatrix}$ *and* $T(\vec{x}) = A\vec{x}$. *Find the matrix of* $T$ *with respect to the basis* $\mathfrak{B}$.

**Solution.** There are different ways we could approach this.

- *Method 1: Construct the* $\mathfrak{B}$*-matrix column by column.*

  Remember that the columns of the $\mathfrak{B}$-matrix of $T$ are $[T(\vec{v}_1)]_{\mathfrak{B}}$ and $[T(\vec{v}_2)]_{\mathfrak{B}}$. We have

  $$T(\vec{v}_1) = A\vec{v}_1 = \begin{bmatrix} 3 & 4 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 10 \\ 2 \end{bmatrix}$$

  $$T(\vec{v}_2) = A\vec{v}_2 = \begin{bmatrix} 3 & 4 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \end{bmatrix}$$

  We can see by inspection that $T(\vec{v}_1) = 5\vec{v}_1 - 3\vec{v}_2$ and $T(\vec{v}_2) = 2\vec{v}_1 - 2\vec{v}_2$. So,

  $$[T(\vec{v}_1)]_{\mathfrak{B}} = \begin{bmatrix} 5 \\ -3 \end{bmatrix} \text{ and } [T(\vec{v}_2)]_{\mathfrak{B}} = \begin{bmatrix} 2 \\ -2 \end{bmatrix}.$$

  Therefore, the $\mathfrak{B}$-matrix of $T$ is

  $$\begin{bmatrix} 5 & 2 \\ -3 & -2 \end{bmatrix}.$$

- *Method 2: Use the formula $B = S^{-1}AS$.*

  We defined a matrix $S$ whose columns were the vectors in $\mathfrak{B}$. In this case,

  $$S = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}.$$

  Then, we showed that the $\mathfrak{B}$-matrix of $T(\vec{x}) = A\vec{x}$ is just $S^{-1}AS$. In this case (after some computation), we find that

  $$S^{-1} = \begin{bmatrix} 1/2 & 0 \\ -1/2 & 1 \end{bmatrix},$$

  so

  $$S^{-1}AS = \begin{bmatrix} 5 & 2 \\ -3 & -2 \end{bmatrix}.$$

3. Let $P$ be the plane $x_1 + 2x_2 - 3x_3 = 0$. Any vector $\vec{x} \in \mathbb{R}^3$ can be written uniquely as a sum $\vec{x}^P + \vec{x}^\perp$ where $\vec{x}^P$ is a vector in the plane $P$ and $\vec{x}^\perp$ is a vector perpendicular to $P$. We define a linear transformation $\text{proj}_P$ from $\mathbb{R}^3$ to $\mathbb{R}^3$ by $\text{proj}_P(\vec{x}) = \vec{x}^P$. (This is very much like the projection transformations we talked about before, except that we are now projecting onto a plane rather than a line.)

   (a) *Find a convenient basis $\mathfrak{B}$ of $\mathbb{R}^3$ and write the $\mathfrak{B}$-matrix of* $\text{proj}_P$.

   **Solution.** If $\vec{v}$ lies in the plane $P$, then $\text{proj}_P(\vec{v})$ is simply $\vec{v}$. On the other hand, if $\vec{v}$ is perpendicular to the plane $P$, then $\text{proj}_P(\vec{v}) = \vec{0}$. Thus, we would like to find a basis $\mathfrak{B} = (\vec{v}_1, \vec{v}_2, \vec{v}_3)$ of $\mathbb{R}^3$ consisting of vectors which either lie in $P$ or are perpendicular to $P$.

   We start by trying to find vectors lying in $P$. Notice that the equation $x_1 + 2x_3 - 3x_3 = 0$ can be written in matrix form as

   $$\begin{bmatrix} 1 & 2 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0 \qquad (*)$$

   In other words, the plane $P$ is exactly the kernel of the matrix $\begin{bmatrix} 1 & 2 & -3 \end{bmatrix}$. To find the kernel of this matrix, we write the augmented matrix $\left[\begin{array}{ccc:c} 1 & 2 & -3 & 0 \end{array}\right]$. This is already in reduced row-echelon form, so we can just read the solutions off. The free variables are $x_2$ and $x_3$, so the solutions look like

   $$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -2s + 3t \\ s \\ t \end{bmatrix} = s \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}.$$

   Thus, a basis of $P$ is

   $$\vec{v}_1 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}.$$

   On the other hand, by (*), the vector

   $$\vec{v}_3 = \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix}$$

   is perpendicular to $P$. Thus, $\mathfrak{B} = (\vec{v}_1, \vec{v}_2, \vec{v}_3)$ is a basis of $\mathbb{R}^3$. Moreover, since $\vec{v}_1$ and $\vec{v}_2$ lie in $P$, $\text{proj}_P(\vec{v}_1) = \vec{v}_1$ and $\text{proj}_P(\vec{v}_2)$. Since $\vec{v}_3$ is perpendicular to $P$, $\text{proj}_P(\vec{v}_3) = \vec{0}$. So, the $\mathfrak{B}$-matrix of $\text{proj}_P$ is

   $$B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

(b) *Use your answer from part (a) to find the standard matrix of* $\text{proj}_P$.

**Solution.** If

$$S = \begin{bmatrix} -2 & 3 & 1 \\ 1 & 0 & 2 \\ 0 & 1 & -3 \end{bmatrix},$$

the standard matrix of $\text{proj}_P$ is $SBS^{-1}$. Since

$$S^{-1} = \begin{bmatrix} -1/7 & 5/7 & 3/7 \\ 3/14 & 3/7 & 5/14 \\ 1/14 & 1/7 & -3/14 \end{bmatrix},$$

we have

$$SBS^{-1} = \begin{bmatrix} 13/14 & -1/7 & 3/14 \\ -1/7 & 5/7 & 3/7 \\ 3/14 & 3/7 & 5/14 \end{bmatrix}.$$

# Linear Spaces

1. In each problem, decide if the set is a linear space. If it is, find the neutral element.

   (a) *The set of* $2 \times 2$ *matrices with positive entries.*

   **Solution.** This is not a linear space because it is not closed under scalar multiplication: if we multiply a matrix with positive entries by $-1$, the result does not have positive entries.

   (b) $\mathbb{R}^{n\times m}$*, the set of* $n \times m$ *matrices.*

   **Solution.** If we add two $n \times m$ matrices, we get another $n \times m$ matrix. If we multiply a $n \times m$ matrix by a scalar, we also get an $n \times m$ matrix. Therefore, $\mathbb{R}^{n\times m}$ is a linear space. The neutral element is the zero matrix (the $n \times m$ matrix whose entries are all zero).

   (c) *The set of polynomials of the form* $x^2 + bx + c$ *where* $b, c \in \mathbb{R}$.

   **Solution.** This set is not a linear space because it is not closed under addition; for example, $x^2$ and $x^2 + 1$ are both polynomials of the specified form, but their sum is not.

   (d) $F(\mathbb{R}, \mathbb{R})$*, the set of all functions from* $\mathbb{R}$ *to* $\mathbb{R}$.

   **Solution.** This is a linear space since we can add functions and multiply a function by a scalar. The neutral element of $F(\mathbb{R}, \mathbb{R})$ is the zero function, which is defined by $f(x) = 0$.

2. Which of the following are subspaces of $\mathbb{R}^{2\times 2}$?

   (a) *The* $2 \times 2$ *matrices of rank* $1$.

   **Solution.** This is not a subspace of $\mathbb{R}^{2\times 2}$ because it does not contain the neutral element of $\mathbb{R}^{2\times 2}$ (remember that the neutral element of $\mathbb{R}^{2\times 2}$ is the zero matrix).

   (b) *The* $2 \times 2$ *matrices of rank less than or equal to* $1$.

   **Solution.** This is not a subspace of $\mathbb{R}^{2\times 2}$ because it is not closed under addition. For example,

   $$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \text{ and } B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

   both have rank 1, but their sum has rank 2.

   (c) *The* $2 \times 2$ *matrices of rank* $0$.

   **Solution.** The only $2 \times 2$ matrix of rank 0 is the zero matrix. If we add the zero matrix to itself or multiply it by a scalar, we still get the zero matrix. Thus, the set of $2 \times 2$ matrices of rank 0 is a subspace of $\mathbb{R}^{2\times 2}$.

3. Which of the following are subspaces of $F(\mathbb{R}, \mathbb{R})$?

   (a) *The set of all functions* $f$ *from* $\mathbb{R}$ *to* $\mathbb{R}$ *such that* $f(0) = 0$.

   **Solution.** This set contains the neutral element of $F(\mathbb{R}, \mathbb{R})$. If $f$ and $g$ are functions from $\mathbb{R}$ to $\mathbb{R}$ such that $f(0) = g(0) = 0$, then $(f + g)(0) = f(0) + g(0) = 0$. If $c$ is a constant, then $(cf)(0) = cf(0) = 0$. So, this is a subspace of $F(\mathbb{R}, \mathbb{R})$.

(b) *The set of all functions $f$ from $\mathbb{R}$ to $\mathbb{R}$ such that $f(0) = 0$ and $f(1) \geq 0$.*

**Solution.** Let's call this set $S$. $S$ is not a subspace of $F(\mathbb{R}, \mathbb{R})$ because it is not closed under scalar multiplication. For example, $f(x) = x^2$ satisfies $f(0) = 0$ and $f(1) \geq 0$, so $f \in S$. On the other hand, $(-f)(1) = -f(1) = -1$, so $-f \notin S$.

4. (a) *Find a basis $\mathfrak{B}$ of $\mathbb{R}^{2\times 2}$, and write the matrix $M = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ in $\mathfrak{B}$-coordinates.*

**Solution.** There are many different possible choices of basis. We'll use $\mathfrak{B} = (M_1, M_2, M_3, M_4)$ where

$$M_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, M_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, M_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \text{ and } M_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Then, the given matrix $M$ is $M_1 + 2M_2 + 3M_3 + 4M_4$, so

$$[M]_{\mathfrak{B}} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}.$$

(b) *Find a basis $\mathfrak{B}$ of $P_4$, and write the polynomial $x^4 - x^2 + 3$ in $\mathfrak{B}$-coordinates.*

**Solution.** Again, there are many different ways to choose a basis of $\mathfrak{B}$. One basis is $\mathfrak{B} = (x^4, x^3, x^2, x, 1)$. Then,

$$[x^4 - x^2 + 3]_{\mathfrak{B}} = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \\ 3 \end{bmatrix}.$$

5. *Find the dimension of $P_n = \{a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0 : a_i \in \mathbb{R}\}$, the set of polynomials of degree at most $n$. What can you say about the dimension of $P$, the set of all polynomials?*

**Solution.** Observe that $(x^n, x^{n-1}, \ldots, x, 1)$ is a basis of $P_n$. This basis consists of $n+1$ elements, so $\dim P_n = n + 1$.

Since $P$ contains $P_n$ for every $n$, $P$ must be infinite dimensional.

# The Gram-Schmidt Process

Let $M = \begin{bmatrix} 1 & 3 & 1 \\ 1 & -1 & 3 \\ 1 & -1 & 1 \\ 1 & 3 & -1 \end{bmatrix}$ and $V = \operatorname{im} M$. Then, the vectors

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} 3 \\ -1 \\ -1 \\ 3 \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} 1 \\ 3 \\ 1 \\ -1 \end{bmatrix}$$

form a basis of $V$ (do you see why?).

1. *Find an orthonormal basis of $V$.*

   **Solution.** We start by letting $\vec{u}_1 = \frac{1}{\|\vec{v}_1\|}\vec{v}_1$. Since $\vec{v}_1 \cdot \vec{v}_1 = 4$, $\|\vec{v}_1\| = 2$. So,

$$\vec{u}_1 = \begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix}.$$

   Since $\vec{u}_1 \cdot \vec{v}_2 = 2$,

$$\vec{v}_2^{\perp} = \vec{v}_2 - (\vec{u}_1 \cdot \vec{v}_2)\vec{u}_1 = \begin{bmatrix} 3 \\ -1 \\ -1 \\ 3 \end{bmatrix} - 2\begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \\ -2 \\ 2 \end{bmatrix}.$$

   Now, $\|\vec{v}_2^{\perp}\| = 4$, so we take

$$\vec{u}_2 = \frac{1}{4}\vec{v}_2^{\perp} = \begin{bmatrix} 1/2 \\ -1/2 \\ -1/2 \\ 1/2 \end{bmatrix}.$$

   Finally, $\vec{v}_3^{\perp} = \vec{v}_3 - (\vec{u}_1 \cdot \vec{v}_3)\vec{u}_1 - (\vec{u}_2 \cdot \vec{v}_3)\vec{u}_2$. Since $\vec{u}_1 \cdot \vec{v}_3 = 2$ and $\vec{u}_2 \cdot \vec{v}_3 = -2$,

$$\vec{v}_3^{\perp} = \vec{v}_3 - 2\vec{u}_1 + 2\vec{u}_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}.$$

   Now, $\|\vec{v}_3^{\perp}\| = 2$, so

$$\vec{u}_3 = \frac{1}{2}\vec{v}_3^{\perp} = \begin{bmatrix} 1/2 \\ 1/2 \\ -1/2 \\ -1/2 \end{bmatrix}.$$

   Thus, an orthonormal basis of $V$ is

$$\vec{u}_1 = \begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix}, \vec{u}_2 = \begin{bmatrix} 1/2 \\ -1/2 \\ -1/2 \\ 1/2 \end{bmatrix}, \vec{u}_3 = \begin{bmatrix} 1/2 \\ 1/2 \\ -1/2 \\ -1/2 \end{bmatrix}.$$

   Of course, it is easy to check that these vectors really are orthonormal.

2. *Find the QR-factorization of M.*

**Solution.** The matrix $Q$ simply consists of the new orthonormal basis vectors; that is, $Q = [\vec{u}_1\ \vec{u}_2\ \vec{u}_3]$. So,

$$Q = \begin{bmatrix} 1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & 1/2 \\ 1/2 & -1/2 & -1/2 \\ 1/2 & 1/2 & -1/2 \end{bmatrix}.$$

Let $\mathfrak{A} = (\vec{u}_1, \vec{u}_2, \vec{u}_3)$. Then, the matrix $R$ has columns $[\vec{v}_1]_{\mathfrak{A}}$, $[\vec{v}_2]_{\mathfrak{A}}$, and $[\vec{v}_3]_{\mathfrak{A}}$.

Since $\mathfrak{A}$ is an orthonormal basis of $V$, recall that any $\vec{x} \in V$ can be written as $\vec{x} = (\vec{u}_1 \cdot \vec{x})\vec{u}_1 + (\vec{u}_2 \cdot \vec{x})\vec{u}_2 + (\vec{u}_3 \cdot \vec{x})\vec{u}_3$. That is,

$$[\vec{x}]_{\mathfrak{B}} = \begin{bmatrix} \vec{u}_1 \cdot \vec{x} \\ \vec{u}_2 \cdot \vec{x} \\ \vec{u}_3 \cdot \vec{x} \end{bmatrix}.$$

In particular,

$$[\vec{v}_1]_{\mathfrak{B}} = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}, [\vec{v}_2]_{\mathfrak{B}} = \begin{bmatrix} 2 \\ 4 \\ 0 \end{bmatrix}, [\vec{v}_3]_{\mathfrak{B}} = \begin{bmatrix} 2 \\ -2 \\ 2 \end{bmatrix}.$$

Therefore,

$$R = \begin{bmatrix} 2 & 2 & 2 \\ 0 & 4 & -2 \\ 0 & 0 & 2 \end{bmatrix}.$$

We can easily check that this is correct: the columns of $Q$ are orthonormal, $R$ is upper triangular with positive entries on the diagonal, and $M = QR$, which means that we really do have the correct $QR$-factorization.

3. *Find the QR-factorization of* $M = \begin{bmatrix} 1 & 4 & 7 \\ 0 & -2 & 3 \\ 0 & 0 & -1 \end{bmatrix}$.

**Solution.** Rather than performing the Gram-Schmidt process on the columns of $M$, we can simply observe that

$$M = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 4 & 7 \\ 0 & 2 & -3 \\ 0 & 0 & 1 \end{bmatrix}.$$

The first matrix has orthogonal columns while the second one is upper triangular with positive entries, so this must be the $QR$-factorization of $M$ (because the $QR$-factorization is unique).

# Least Squares and Data Fitting

1. *Suppose we want to find a quadratic polynomial $f(x) = ax^2 + bx + c$ such that $f(-1) = 1$, $f(0) = 0$, $f(1) = 2$, and $f(2) = 5$. Express this problem as a linear system.*

   **Solution.** We want

$$\begin{aligned}1 &= f(-1) = a(-1)^2 + b(-1) + c = a - b + c\\ 0 &= f(0) = a(0)^2 + b(0) + c = c\\ 2 &= f(1) = a(1)^2 + b(1) + c = a + b + c\\ 5 &= f(2) = a(2)^2 + b(2) + c = 4a + 2b + c\end{aligned}$$

   That is, we are trying to solve the linear system

$$\begin{array}{rcrcrcr} a & - & b & + & c & = & 1\\ & & & & c & = & 0\\ a & + & b & + & c & = & 2\\ 4a & + & 2b & + & c & = & 5\end{array},$$

   which can be written as a matrix equation

$$\begin{bmatrix} 1 & -1 & 1\\ 0 & 0 & 1\\ 1 & 1 & 1\\ 4 & 2 & 1\end{bmatrix}\begin{bmatrix} a\\ b\\ c\end{bmatrix} = \begin{bmatrix} 1\\ 0\\ 2\\ 5\end{bmatrix}.$$

2. We want to fit a linear function of the form $f(x) = mx + c$ to the data points $(-1, 3)$, $(0, 1)$, and $(1, 1)$.

   (a) *Do you expect $m$ to be positive, negative, or zero?*

   **Solution.** Looking at the data points, we see that the best-fit line should have negative slope, so $m$ should be negative.

   (b) *Find the linear function $f(x) = mx + c$.*

   **Solution.** We can rewrite the problem as a system

$$\begin{array}{rcrcr} 3 & = & -m & + & c\\ 1 & = & & & c\\ 1 & = & m & + & c\end{array}$$

   We can rewrite this again as

$$\begin{bmatrix} -1 & 1\\ 0 & 1\\ 1 & 1\end{bmatrix}\begin{bmatrix} m\\ c\end{bmatrix} = \begin{bmatrix} 3\\ 1\\ 1\end{bmatrix}.$$

   Thus, if

$$A = \begin{bmatrix} -1 & 1\\ 0 & 1\\ 1 & 1\end{bmatrix} \text{ and } \vec{b} = \begin{bmatrix} 3\\ 1\\ 1\end{bmatrix},$$

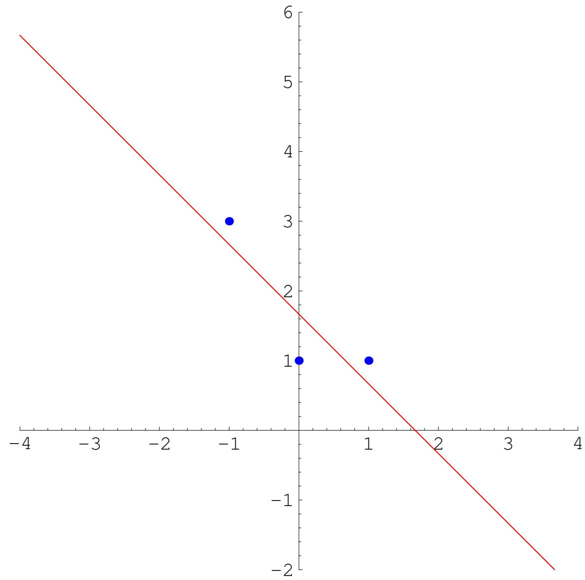
   we are looking for solutions of the system $A\vec{x} = \vec{b}$. This system is inconsistent, but we know that the least-squares solutions are the solutions of the system $A^TA\vec{x} = A^T\vec{b}$, or

$$\begin{bmatrix} 2 & 0\\ 0 & 3\end{bmatrix}\begin{bmatrix} m\\ c\end{bmatrix} = \begin{bmatrix} -2\\ 5\end{bmatrix}.$$

This system has a unique solution, which is

$$\begin{bmatrix} m \\ c \end{bmatrix} = \begin{bmatrix} -1 \\ 5/3 \end{bmatrix}.$$

Thus, the line that best fits the data points is $y = -x + \frac{5}{3}$.




3. The following table describes the percent of classes that Harvard students attend:

| Year ($y$) | Percent of Classes Attended ($p$) |
|---|---|
| 1 (Freshman) | 100 |
| 2 (Sophomore) | 90 |
| 3 (Junior) | 60 |
| 4 (Senior) | 10 |

We suspect that $p(y)$ looks like $ky^n$ for some constants $k$ and $n$, and we would like to find $k$ and $n$.

(a) *Do you expect $k$ and $n$ to be positive or negative? What number should $k$ be close to?*

**Solution.** Since $p(1)$ will be approximated by $k \cdot 1^n = k$, we expect $k$ to be close to 100 (in particular, $k$ should be positive). We expect $n$ to be negative since $p(y)$ decreases as $y$ increases.

(b) *Express this problem as a linear system.*

**Solution.** We have the equations

$$\begin{aligned} k \cdot 1^n &= 100 \\ k \cdot 2^n &= 90 \\ k \cdot 3^n &= 60 \\ k \cdot 4^n &= 10 \end{aligned}$$

The problem is that these are not linear equations. However, if we take the log of all of these equations and write $c = \log k$, then the equations become

$$\begin{aligned} c + n\log 1 &= \log 100 \\ c + n\log 2 &= \log 90 \\ c + n\log 3 &= \log 60 \\ c + n\log 4 &= \log 10 \end{aligned}$$

We can rewrite this system as a matrix equation

$$\begin{bmatrix} 1 & \log 1 \\ 1 & \log 2 \\ 1 & \log 3 \\ 1 & \log 4 \end{bmatrix} \begin{bmatrix} c \\ n \end{bmatrix} = \begin{bmatrix} \log 100 \\ \log 90 \\ \log 60 \\ \log 10 \end{bmatrix}.$$

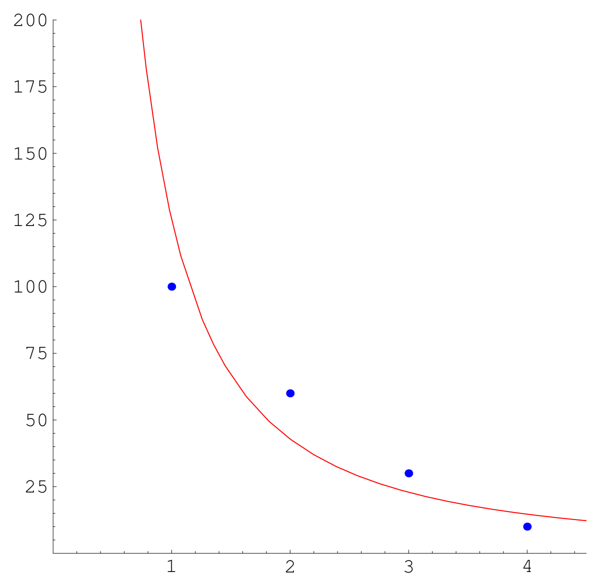
Thus, if

$$A = \begin{bmatrix} 1 & \log 1 \\ 1 & \log 2 \\ 1 & \log 3 \\ 1 & \log 4 \end{bmatrix} \text{ and } \vec{b} = \begin{bmatrix} \log 100 \\ \log 90 \\ \log 60 \\ \log 10 \end{bmatrix},$$

then the least-squares solutions are the solutions of the linear system $A^T A\vec{x} = A^T\vec{b}$. Using Mathematica, we see that this system has a unique solution,

$$\begin{bmatrix} c \\ n \end{bmatrix} \sim \begin{bmatrix} 4.83 \\ -1.55 \end{bmatrix}.$$

Since $k = e^c$, $k \sim 125$. So, $p(y) \sim 125y^{-1.55}$. The graph of $p$ looks like this:



## Odds and Ends

1. *If $A$ is an $n \times m$ matrix, what is the relationship between* $\text{rank}(A^T)$ *and* $\text{rank}(A)$*?*

   **Solution.** Since $A$ is an $n \times m$ matrix, $A^T$ is an $m \times n$ matrix. Thus,

$$\begin{aligned} \text{rank}(A^T) &= n - \dim(\ker A^T) \text{ by the rank-nullity theorem} \\ &= n - \dim(\text{im}\, A)^{\perp} \text{ since } \ker A^T = (\text{im}\, A)^{\perp} \end{aligned}$$

   Now, $\text{im}\, A$ is a subspace of $\mathbb{R}^n$, so $\dim(\text{im}\, A) + \dim(\text{im}\, A)^{\perp} = n$. (Recall: if $V$ is any subspace of $\mathbb{R}^n$, then $\text{proj}_V$ is a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^n$, $\text{im}\,\text{proj}_V = V$, and $\ker \text{proj}_V = V^{\perp}$, so the rank-nullity theorem says that $\dim V + \dim V^{\perp} = n$.) Therefore, $\text{rank}(A^T) = \dim(\text{im}\, A) = \text{rank}(A)$.

2. In each problem, decide if $S$ is a subspace of $\mathbb{R}^{3\times 3}$, the linear space of $3 \times 3$ matrices. If it is, find the dimension of $S$.

   (a) *$S$ is the set of symmetric $3 \times 3$ matrices.*

   **Solution.** $S$ is a subspace:

   - If $A$ and $B$ are symmetric, then $A = A^T$ and $B = B^T$, so $(A+B)^T = A^T + B^T = A + B$, which means that $A + B$ is symmetric. Thus, $S$ is closed under addition.
   - If $A$ is symmetric and $c$ is a scalar, then $(cA)^T = cA^T = cA$, so $cA$ is symmetric. That is, $S$ is closed under scalar multiplication.

   The symmetric $3 \times 3$ matrices look like

$$\begin{aligned} \begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix} &= a \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + b \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \\ &+ d \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} + e \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} + f \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{aligned}$$

   Thus, a basis of $\mathbb{R}^{3\times 3}$ is given by the following elements:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

   In particular, $\dim S = 6$.

   (b) *$S$ is the set of skew-symmetric $3 \times 3$ matrices.*

   **Solution.** $S$ is a subspace:

   - If $A$ and $B$ are skew-symmetric, then $A = -A^T$ and $B = -B^T$, so $A + B = -A^T - B^T = -(A+B)^T$. That is, $A + B$ is also skew-symmetric. So, $S$ is closed under addition.
   - If $A$ is skew-symmetric, then $cA$ is also skew-symmetric for any scalar $c$. So, $S$ is closed under scalar multiplication.

   Any skew-symmetric $3 \times 3$ matrix looks like

$$\begin{bmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{bmatrix} = a \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + b \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}.$$

So, the following elements form a basis of $S$:

$$\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}.$$

Thus, $\dim S = 3$.

(c) *$S$ is the set of orthogonal $3 \times 3$ matrices.*

**Solution.** $S$ is not a subspace. For example, the matrix $I_3$ is orthogonal, but $2I_3$ is not; therefore, $S$ is not closed under scalar multiplication.

3. *Explain why the matrix of a reflection must be symmetric.*

**Solution.** If $A$ is the matrix of a reflection, then we know $A = A^{-1}$ (the inverse of a reflection is just the reflection itself). On the other hand, any reflection is an orthogonal transformation, so $A^{-1} = A^T$. Thus, $A = A^T$, so $A$ is symmetric.

# More on Determinants

Let $A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 2 & 3 & 4 & 0 \\ 5 & 8 & 6 & 7 \\ 9 & 12 & 10 & 11 \end{bmatrix}$.

1. *Write an expression for* $\det A$ *in terms of determinants of* $2 \times 2$ *matrices.*

   **Solution.** We start by expanding along the first row, which gives

   $$\det A = -\begin{vmatrix} 2 & 4 & 0 \\ 5 & 6 & 7 \\ 9 & 10 & 11 \end{vmatrix}.$$

   Expanding along the first row of the new matrix, we have

   $$\det A = -\left(2\begin{vmatrix} 6 & 7 \\ 10 & 11 \end{vmatrix} - 4\begin{vmatrix} 5 & 7 \\ 9 & 11 \end{vmatrix}\right).$$

2. *Let* $B$ *be* $A$ *with the third and fourth rows swapped. That is,*

   $$B = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 2 & 3 & 4 & 0 \\ 9 & 12 & 10 & 11 \\ 5 & 8 & 6 & 7 \end{bmatrix}.$$

   *Write an expression for* $\det B$ *in terms of determinants of* $2 \times 2$ *matrices. How does this relate to your expression for* $\det A$*?*

   **Solution.** Everything is the same, except that the rows of the $2 \times 2$ minors are swapped. That is,

   $$\det B = -\left(2\begin{vmatrix} 10 & 11 \\ 6 & 7 \end{vmatrix} - 4\begin{vmatrix} 9 & 11 \\ 5 & 7 \end{vmatrix}\right).$$

   In particular, $\det B = -\det A$.

3. *Let* $B$ *be* $A$ *with a multiple of the fourth row added to the third row. That is,*

   $$B = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 2 & 3 & 4 & 0 \\ 9k+5 & 12k+8 & 10k+6 & 11k+7 \\ 9 & 12 & 10 & 11 \end{bmatrix}.$$

   *Write an expression for* $\det B$ *in terms of determinants of* $2 \times 2$ *matrices. How does this relate to your expression for* $\det A$*?*

   **Solution.** Again, only the $2 \times 2$ minors are affected:

   $$\det B = -\left(2\begin{vmatrix} 10k+6 & 11k+7 \\ 10 & 11 \end{vmatrix} - 4\begin{vmatrix} 9k+5 & 11k+7 \\ 9 & 11 \end{vmatrix}\right).$$

   In particular, $\det B = \det A$.

4. *Let* $A = \begin{bmatrix} 0 & 0 & 3 \\ 1 & 0 & 3 \\ 2 & -2 & -2 \end{bmatrix}$. *Use row reduction to find* $\det A$.

**Solution.** We row-reduce $A$, keeping track of how the determinant changes.

$$\begin{array}{rcll}
A & = & \begin{bmatrix} 0 & 0 & 3 \\ 1 & 0 & 3 \\ 2 & -2 & -2 \end{bmatrix} \begin{array}{l} \text{swap with (II)} \\ \\ \\ \end{array} & \\
\to B_1 & = & \begin{bmatrix} 1 & 0 & 3 \\ 0 & 0 & 3 \\ 2 & -2 & -2 \end{bmatrix} \begin{array}{l} \\ \\ -2\text{(I)} \end{array} & [\det B_1 = -\det A] \\
\to B_2 & = & \begin{bmatrix} 1 & 0 & 3 \\ 0 & 0 & 3 \\ 0 & -2 & -8 \end{bmatrix} \begin{array}{l} \\ \text{swap with (III)} \\ \\ \end{array} & [\det B_2 = \det B_1] \\
\to B_3 & = & \begin{bmatrix} 1 & 0 & 3 \\ 0 & -2 & -8 \\ 0 & 0 & 3 \end{bmatrix} \begin{array}{l} \\ \div(-2) \\ \\ \end{array} & [\det B_3 = -\det B_2] \\
\to B_4 & = & \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 3 \end{bmatrix} \begin{array}{l} \\ \\ \div 3 \end{array} & [\det B_4 = \frac{1}{-2}\det B_3] \\
\to B_5 & = & \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix} \begin{array}{l} -3\text{(III)} \\ -4\text{(III)} \\ \\ \end{array} & [\det B_5 = \frac{1}{3}\det B_4] \\
\to B_6 & = & \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} & [\det B_6 = \det B_5]
\end{array}$$

Now, we know that $\det B_6 = 1$, and we have

$$\begin{aligned}
\det B_6 &= \det B_5 \\
&= \left(\frac{1}{3}\right)\det B_4 \\
&= \left(\frac{1}{3}\right)\left(\frac{1}{-2}\right)\det B_3 \\
&= (-1)\left(\frac{1}{3}\right)\left(\frac{1}{-2}\right)\det B_2 \\
&= (-1)\left(\frac{1}{3}\right)\left(\frac{1}{-2}\right)\det B_1 \\
&= (-1)^2\left(\frac{1}{3}\right)\left(\frac{1}{-2}\right)\det A
\end{aligned}$$

So, $\det A = (-1)^2(3)(-2)\det B_6 = -6$. Actually, we could have stopped once we got to $B_3$ since $B_3$ is upper triangular, so we can easily find its determinnt.

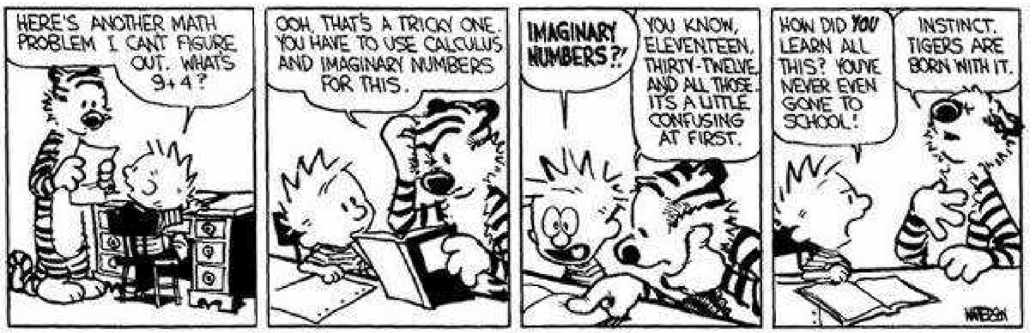
5. *If $A$ is an orthogonal $n \times n$ matrix, what can you say about* $\det A$*?*

**Solution.** Since $A$ is orthogonal, $A^T A = I_n$. Therefore, $(\det A^T)(\det A) = \det I_n = 1$. However, we know that $\det A^T = \det A$, so $(\det A)^2 = 1$, which means $\det A = \pm 1$.

6. *If $A$ is any $n \times m$ matrix with linearly independent columns, $A$ has a QR-factorization $A = QR$. Express* $\det(A^T A)$ *in terms of* $\det R$. *(Although $A$ is not necesarily square, $A^T A$ is always square, so* $\det(A^T A)$ *makes sense.)*

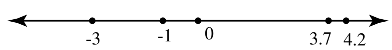
**Solution.** Since $A = QR$, $A^T A = (QR)^T(QR) = R^T Q^T QR$. Since $Q$ is an $n \times m$ matrix with orthonormal columns, $Q^T Q = I_m$, so $A^T A = R^T I_m R = R^T R$. Therefore, $\det(A^T A) = \det(R^T R)$. Now, $R$ is always a square matrix, so $\det(A^T A) = \det(R^T R) = \det(R^T)\det(R) = (\det R)^2$.

# Complex Numbers



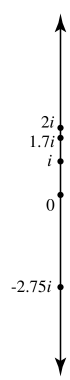

## Basic Definitions

You're used to real numbers and how they're graphed on the real number line.



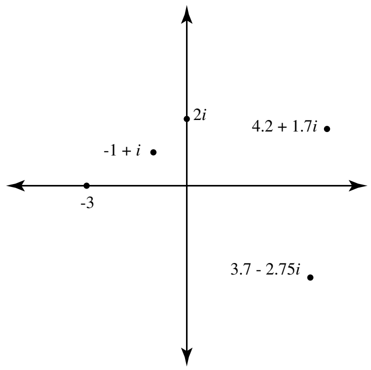

Real numbers are very useful, but there is one major problem we encounter when working with them: not all polynomials have real roots. For example, although the polynomials $x^2 - 3$ and $x^2 + 1$ look almost identical, the former has two real roots ($\pm\sqrt{3}$) while the latter has none. As we saw in class, this means that some matrices don't have any real eigenvalues. To fix this problem, mathematicians simply define a new number, called $i$, to be the square root of $-1$. Of course, $i$ is not a real number, since there is no real number whose square is $-1$; instead, we call $i$ an imaginary number. Now, the polynomial $x^2 + 1$ has two roots, $i$ and $-i$.

More generally, an imaginary number is any real number multiple of $i$, like 0, $2i$, $-2.75i$, or $1.7i$. We can graph the imaginary numbers on a number line, but we use a vertical line instead of a horizontal line.




We know how to add and multiply real numbers, and we would like to do the same with imaginary numbers. However, if two imaginary numbers are multiplied, the answer is a real number; for instance, $(2i)(3i) = 6i^2 = -6$. This suggests that we shouldn't look at real and imaginary numbers separately; instead, we study complex numbers, which are sums of real and imaginary numbers. That is, a complex number is just a number of the form $x + iy$ where $x$ and $y$ are both real numbers. So, $1 + \sqrt{3}i$, $4.2 + 1.7i$, $-3$, $3.7 - 2.75i$, and $2i$ are all complex numbers. We write $\mathbb{C}$ for the set of complex numbers.

Since we view the set of real numbers as a horizontal line and the set of imaginary numbers as a vertical line, it's natural to view the set of complex numbers as a plane in which the real numbers lie on the horizontal axis and the imaginary numbers lie on the vertical axis. This plane is called the complex plane.




## Arithmetic of Complex Numbers

As we said already, a complex number is just a number of the form $z = x + iy$ where $x, y \in \mathbb{R}$. We call $x$ the real part of $z$ and $y$ the imaginary part of $z$. To add complex numbers, we just add the real and imaginary parts separately. That is, $(x_1 + iy_1) + (x_2 + iy_2) = (x_1 + x_2) + i(y_1 + y_2)$.

**Example 1.** $(4.2 + 1.7i) + (3.7 - 2.75i) = (4.2 + 3.7) + (1.7 - 2.75)i = 7.9 - 1.05i$. ❖

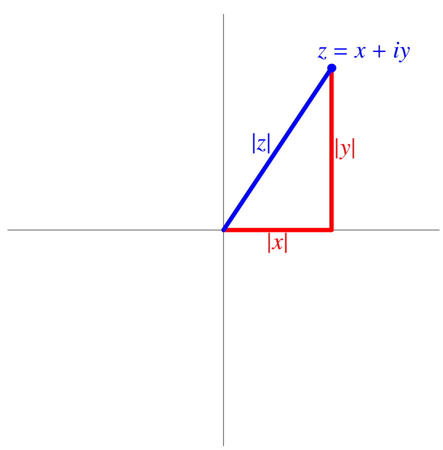
To multiply complex numbers, we use the distributive property. That is,

$$\begin{aligned}(x_1 + iy_1)(x_2 + iy_2) &= x_1(x_2 + iy_2) + iy_1(x_2 + iy_2) \\ &= x_1x_2 + ix_1y_2 + iy_1x_2 + i^2y_1y_2 \\ &= x_1x_2 + ix_1y_2 + iy_1x_2 - y_1y_2 \text{ since } i^2 = -1 \\ &= (x_1x_2 - y_1y_2) + i(x_1y_2 + y_1x_2)\end{aligned}$$

**Example 2.** $(1 + i)(5 + 3i) = 2 + 8i$. ❖

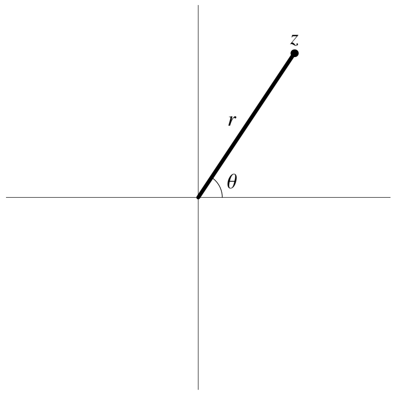
## The Absolute Value of a Complex Number

When $x$ is a real number, the absolute value of $x$ measures the distance from $x$ to 0 on the number line. For instance, 5 and $-5$ are both 5 units away from 0, so $|5| = |-5| = 5$. Similarly, the absolute value of a complex number $z = x + iy$ is defined to be the distance from $z$ to 0. By the Pythagorean theorem, $|z| = \sqrt{x^2 + y^2}$.



## Polar Coordinates

When we write a complex number $z$ as $x + iy$ with $x, y \in \mathbb{R}$, we say that we are writing $z$ in Cartesian coordinates. There is another useful way to write a complex number: any complex number $z$ is determined by its distance from the origin ($r$) and an angle ($\theta$):

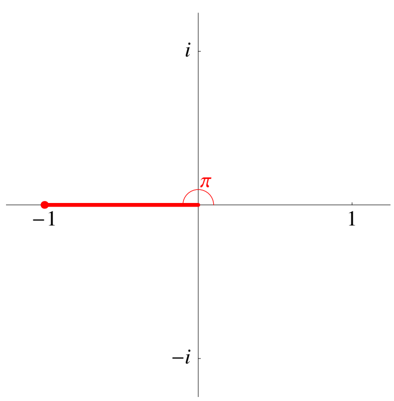



We call $r$ and $\theta$ the polar coordinates of $z$. We can see from the diagram that $z = r\cos\theta + i(r\sin\theta)$. We write this more simply as $z = re^{i\theta}$. (If you want to know why $e^{i\theta} = \cos\theta + i\sin\theta$, try writing out the Taylor series of $e^x$, $\sin x$, and $\cos x$.)

To summarize, we can write a complex number $z$ in the form $z = x + iy$ (Cartesian coordinates) or the form $z = re^{i\theta}$ (polar coordinates). These two forms are related as follows.

- $x = r\cos\theta$ and $y = r\sin\theta$.
- $r = |z| = \sqrt{x^2 + y^2}$ and $\tan\theta = \frac{y}{x}$.

**Example 3.** $e^{\pi i}$ is the complex number with $r = 1$ and $\theta = \pi$. By drawing a diagram, we see that $e^{\pi i} = -1$.

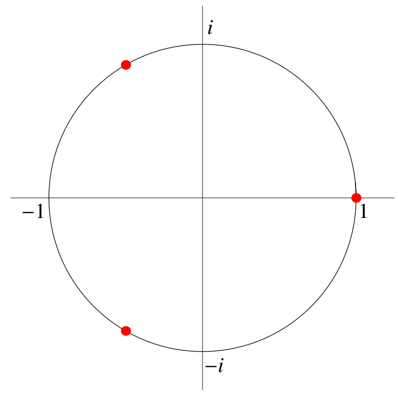



Similarly, $e^{2\pi i} = 1$. Notice that, if $n$ is any integer, then $e^{2\pi i n} = 1$. ❖

If we are adding complex numbers, it is useful to write them in Cartesian coordinates; on the other hand, if we are multiplying complex numbers, it is often easier to use polar coordinates. After all, $\left(r_1e^{i\theta_1}\right)\left(r_2e^{i\theta_2}\right)$ is just $r_1r_2e^{i\theta_1}e^{i\theta_2} = r_1r_2e^{i(\theta_1+\theta_2)}$. Here are two examples of how polar coordinates can be useful.

**Example 4.** Let $z = \frac{1}{2} + \frac{\sqrt{3}}{2}i$. Suppose we want to find $z^{100}$. We could just start computing powers of $z$, but that would get boring really fast. Instead, let's write $z$ in polar coordinates. We know that $z = re^{i\theta}$ where $r = \sqrt{\left(\frac{1}{2}\right)^2 + \left(\frac{\sqrt{3}}{2}\right)^2} = 1$ and $\tan\theta = \sqrt{3}$. Thus, $\theta$ is either $\frac{\pi}{3}$ or $\frac{4\pi}{3}$. Since $z$ lies in the first quadrant of the complex plane, it must be the case that $\theta = \frac{\pi}{3}$, so $z = e^{\pi i/3}$. Then, $z^{100} = \left(e^{\pi i/3}\right)^{100} = e^{100\pi i/3}$. In Cartesian coordinates, $z^{100} = \cos\frac{100\pi}{3} + i\sin\frac{100\pi}{3} = -\frac{1}{2} - \frac{\sqrt{3}}{2}i$. ❖

**Example 5.** Let's find all complex numbers $z$ such that $z^3 = 1$. We know we can write any complex number $z$ as $re^{i\theta}$ for some $r$ and $\theta$. Then, $z^3 = r^3e^{3i\theta}$, so $z^3 = 1$ if and only if $r^3 = 1$ and $3\theta$ is a multiple of $2\pi$. So, the solutions of $z^3 = 1$ are $z = 1$, $e^{2i\pi/3}$, and $e^{4i\pi/3}$. These lie on a circle of radius 1 in the complex plane:




In Cartesian coordinates, $e^{2\pi i/3} = \cos\frac{2\pi}{3} + i\sin\frac{2\pi}{3} = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$ and $e^{4\pi i/3} = \cos\frac{4\pi}{3} + i\sin\frac{4\pi}{3} = -\frac{1}{2} - i\frac{\sqrt{3}}{2}$. ❖

## The Fundamental Theorem of Algebra

As we discussed earlier, not all polynomials have real roots, so not all matrices have real eigenvalues. For complex numbers, the situation is much better.

**The Fundamental Theorem of Algebra.** *Let $p(x)$ be a polynomial whose coefficients are complex numbers (or just real numbers), and let $n$ be the degree of $p(x)$. Then, $p(x)$ factors as $p(x) = c(x-a_1)\cdots(x-a_n)$ for some complex numbers $c, a_1, \ldots, a_n$. In particular, $p(x)$ has exactly $n$ roots $a_1, \ldots, a_n$ (if the roots are counted with multiplicity).*

**Example 6.** The polynomial $x^2 + 1$ can be written as $(x - i)(x + i)$. That is, it has two roots, $i$ and $-i$. ❖

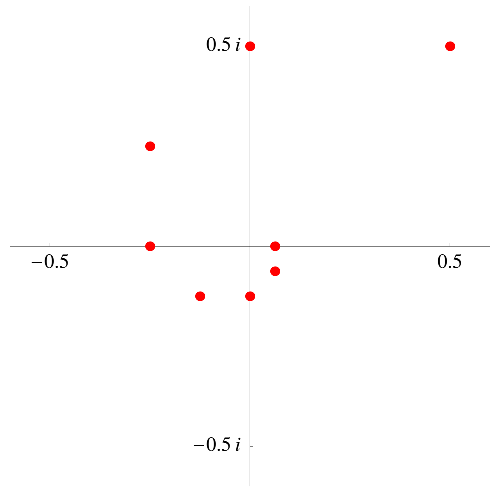
**Example 7.** The polynomial $x^2 + 2ix - 1$ can be written as $(x + i)^2$, so $-i$ is a root of multiplicity 2. ❖

The Fundamental Theorem of Algebra tells us that every $n \times n$ matrix has *exactly* $n$ complex eigenvalues. Don't worry if you don't understand the theorem perfectly; we will talk about it more in class.

## Practice Problems

1. *Let $z = \frac{1}{2} + \frac{1}{2}i$. Write $z$ in polar coordinates. Draw $z, z^2, z^3, \ldots, z^8$ in the complex plane. Write $z^5$ in Cartesian coordinates.*

   **Solution.** We can write $z = re^{i\theta}$ where $r = \sqrt{\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2} = \frac{1}{\sqrt{2}}$ and $\tan\theta = 1$. Since $z$ is in the first quadrant of the complex plane, $\theta$ must be $\frac{\pi}{4}$. Thus, $z = \frac{1}{\sqrt{2}}e^{\pi i/4}$, so $z^n = \frac{1}{(\sqrt{2})^n}e^{n\pi i/4}$. Therefore, $z, z^2, \ldots, z^8$ form a sort of spiral inwards:




In particular,

$$
\begin{aligned}
z^5 &= \frac{1}{(\sqrt{2})^5} e^{5\pi i/4} \\
&= \frac{1}{4\sqrt{2}} e^{5\pi i/4} \\
&= \frac{1}{4\sqrt{2}} \cos \frac{5\pi}{4} + \left( \frac{1}{4\sqrt{2}} \sin \frac{5\pi}{4} \right) i \\
&= -\frac{1}{8} - \frac{1}{8} i
\end{aligned}
$$

2. *Explain why $\mathbb{C}$ is a linear space. What is its dimension?*

   **Solution.** Adding two complex numbers gives us another complex number; after all, we said that $(x_1 + iy_1) + (x_2 + iy_2) = (x_1 + x_2) + i(y_1 + y_2)$; thus, $\mathbb{C}$ is closed under addition.

   If $z = x + iy$ is a complex number and $c$ is a real number (i.e., a scalar), then $cz = cx + i(cy)$ is a complex number, so $\mathbb{C}$ is closed under scalar multiplication.

   Therefore, $\mathbb{C}$ is a linear space. Any element of $\mathbb{C}$ can be written uniquely as $x + iy = x \cdot 1 + y \cdot i$, so 1 and $i$ form a basis of $\mathbb{C}$. Therefore, $\mathbb{C}$ has dimension 2 (which makes sense intuitively since $\mathbb{C}$ looks like a plane).

3. *If $z = re^{i\theta}$ is a nonzero complex number, write $\frac{1}{z}$ in polar coordinates.*

   **Solution.** If $z = re^{i\theta}$, then $\frac{1}{z} = \frac{1}{r}\frac{1}{e^{i\theta}} = \frac{1}{r}e^{-i\theta}$.

4. *Find the square roots of $i$.*

   **Solution.** We want to find $z = re^{i\theta}$ such that $z^2 = i$. In polar coordinates, $i = e^{\pi i/2}$, so we want $r^2 e^{2i\theta} = e^{\pi i/2}$. Therefore, we must have $r = 1$ and $2\theta$ equal to $\frac{\pi}{2}$ plus a multiple of $2\pi$. Thus, the square roots of $i$ are $e^{\pi i/4}$ and $e^{5\pi i/4}$. In Cartesian coordinates, the square roots are $\pm\left(\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2} i\right)$.

## Julia Sets (Optional but Super Cool!)

In class, we talked about a dynamical system involving wolves and deer. Essentially, we came up with a matrix $A$ that represented how wolf and deer populations change over time. Then, if $T$ was the linear transformation $T(\vec{x}) = A\vec{x}$ and $\vec{x}_0$ was a vector representing the starting populations, we found:

- the populations after one year are given by $\vec{x}_1 = T(\vec{x}_0)$,
- the populations after two years are given by $\vec{x}_2 = T(\vec{x}_1)$,
- and so on.

We might try doing the same thing with a different type of transformation; it turns out that, if we make $T$ a quadratic function, we get very interesting objects called Julia sets. You won't need to know about Julia sets for this class, but they are pretty amazing things.

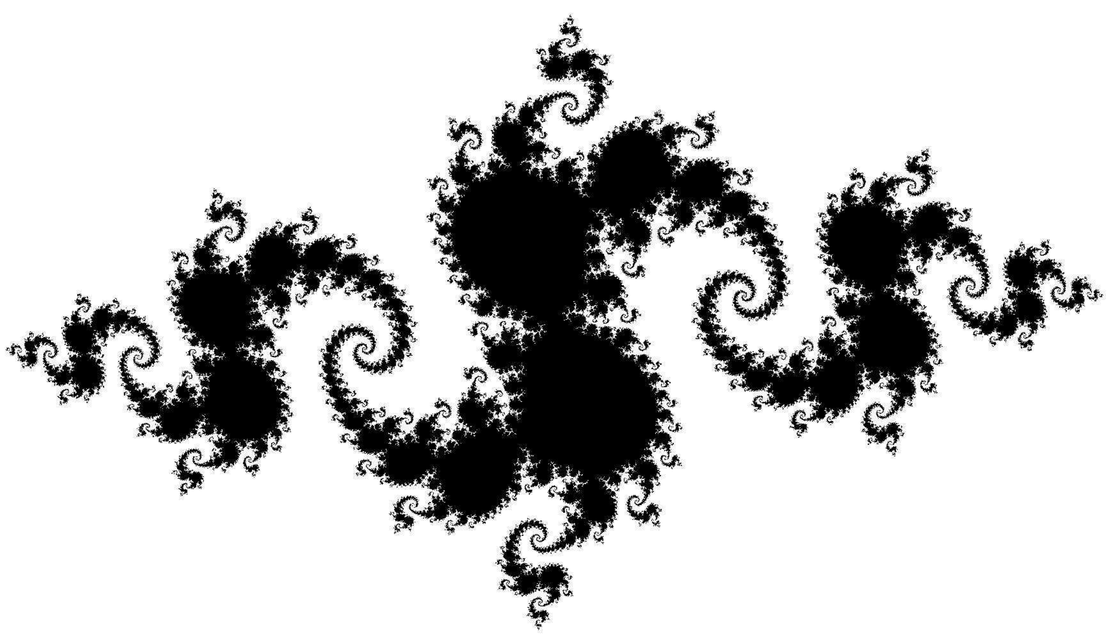
The way to make a Julia set is pretty easy. We start by picking a complex number $c$, like $c = -0.8 - 0.15i$ (every choice of $c$ gives a different Julia set). Then, we define a function $f(z) = z^2 + c$. If $z_0$ is any complex number, we define:

$$\begin{aligned} z_1 &= f(z_0) \\ z_2 &= f(z_1) \\ z_3 &= f(z_2) \\ &\vdots \end{aligned}$$

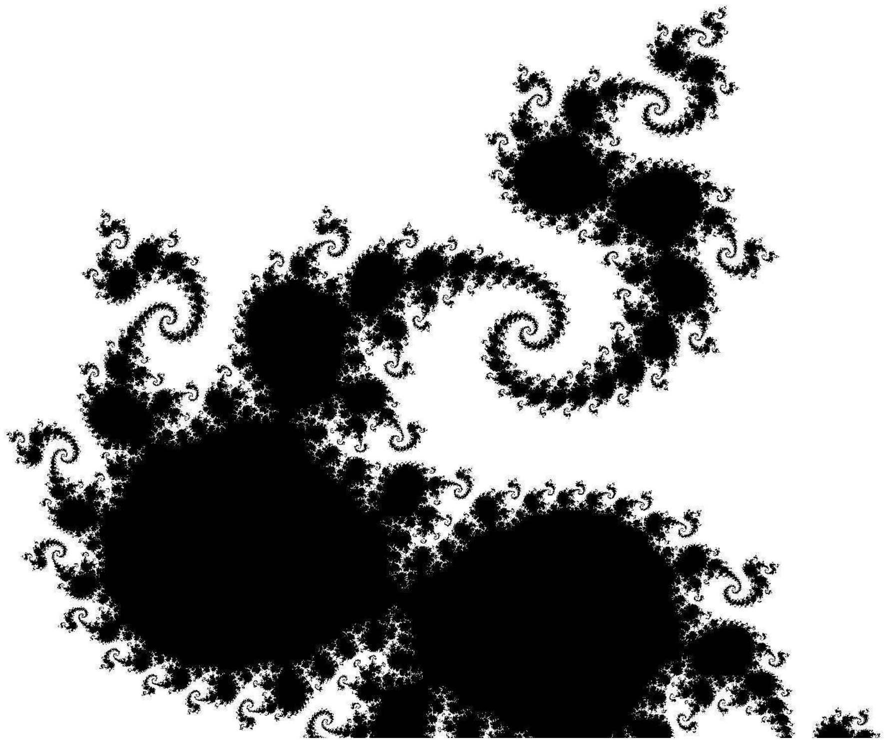
This is just a list of complex numbers $z_0, z_1, z_2, \ldots$; we call this list the orbit of $z_0$. In general, such orbits may be really wacky; rather than trying to understand them very deeply, we just ask a simple question: do the $z_n$ stay close to the origin? There are two possibilities:

1. Every $z_n$ satisfies $|z_n| < 2$. That is, all elements of the orbit stay within 2 units of the origin. (We say the orbit "does not escape.")
2. Some $z_n$ satisfies $|z_n| \geq 2$. (We say the orbit "escapes.")

Now, we are ready to define the Julia set (for $c$) — it is simply the set of all complex numbers $z_0$ such that the orbit of $z_0$ does not escape. Although this definition is relatively straightforward, Julia sets are extremely complicated objects. For example, here is the Julia set for $c = -0.8 - 0.15i$; as you can see, it is not simple at all!



Julia sets have lots of interesting properties. One of the most striking is that tiny pieces of a Julia set look identical to the entire Julia set. For instance, here is a teeny piece of the previous image, magnified about a million times:



The pictures get even more interesting if you color the white points somehow; you can see some examples at http://math.harvard.edu/~jjchen/fractals/

# Eigenvectors

1. *True or false: if $(\vec{v}_1, \ldots, \vec{v}_n)$ is an eigenbasis for $A$, then $(\vec{v}_1, \ldots, \vec{v}_n)$ is an eigenbasis for $A^2$.*

   **Solution.** True. Since $(\vec{v}_1, \ldots, \vec{v}_n)$ is an eigenbasis for $A$, $(\vec{v}_1, \ldots, \vec{v}_n)$ is a basis of $\mathbb{R}^n$ and each $\vec{v}_i$ is an eigenvector of $A$. In particular, for each $i$, there is some scalar $\lambda_i$ such that $A\vec{v}_i = \lambda_i \vec{v}_i$. Then, $A^2\vec{v}_i = A(A\vec{v}_i) = A(\lambda_i \vec{v}_i) = \lambda_i(A\vec{v}_i) = \lambda_i^2 \vec{v}_i$, so $\vec{v}_i$ is an eigenvector of $A^2$. Thus, we have shown that $(\vec{v}_1, \ldots, \vec{v}_n)$ is a basis for $\mathbb{R}^n$ and that each $\vec{v}_i$ is an eigenvector of $A^2$. Therefore, $(\vec{v}_1, \ldots, \vec{v}_n)$ is an eigenbasis of $A^2$.

2. *Let $V$ be a plane in $\mathbb{R}^3$ which contains the origin, and let $A$ be the matrix of $\operatorname{proj}_V$ (that is, $\operatorname{proj}_V(\vec{x}) = A\vec{x}$). Find the eigenvalues of $A$, and describe the eigenspaces geometrically. Is there an eigenbasis for $A$?*

   **Solution.** If $\vec{x} \in V$, then $A\vec{x} = \operatorname{proj}_V(\vec{x}) = \vec{x}$, so $\vec{x}$ is an eigenvector of $A$ with eigenvalue 1. $E_1$, the 1-eigenspace of $A$, consists of all vectors $\vec{x} \in \mathbb{R}^3$ such that $\operatorname{proj}_V(\vec{x}) = \vec{x}$; this is simply the plane $V$.

   If $\vec{x} \in V^\perp$, then $A\vec{x} = \operatorname{proj}_V(\vec{x}) = \vec{0}$, so $\vec{x}$ is an eigenvector of $A$ with eigenvalue 0. $E_0$, the 0-eigenspace of $A$, consists of all vectors $\vec{x} \in \mathbb{R}^3$ such that $\operatorname{proj}_V(\vec{x}) = \vec{0}$; this is simply the line $V^\perp$.

   There is an eigenbasis for $A$: if we choose two linearly independent vectors $\vec{v}_1, \vec{v}_2$ in $E_0$ and a nonzero vector $\vec{v}_3$ in $E_1$, then $(\vec{v}_1, \vec{v}_2, \vec{v}_3)$ is an eigenbasis for $A$.

3. *Let $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, which represents a shear. Find the eigenvalues of $A$ and their algebraic and geometric multiplicities. Is there an eigenbasis for $A$?*

   **Solution.** The characteristic polynomial of $A$ is

   $$f_A(\lambda) = \det \begin{bmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{bmatrix} = (1-\lambda)^2.$$

   Thus, the only eigenvalue of $A$ is 1, and it has algebraic multiplicity 2.

   The 1-eigenspace of $A$ is $\ker(A - I)$, which is spanned by the vector $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Therefore, 1 has geometric multiplicity 1, and there is no eigenbasis for $A$.

4. *Let $A$ be a noninvertible $n \times n$ matrix. Explain why 0 must be an eigenvalue of $A$. Find the geometric multiplicity of the eigenvalue 0 in terms of $\operatorname{rank}(A)$.*

   **Solution.** If $A$ is not invertible, then $\det(A - 0I_n) = \det A = 0$, so 0 is an eigenvalue of $A$. The geometric multiplicity of 0 is $\dim[\ker(A - 0I_n)] = \dim(\ker A) = n - \operatorname{rank}(A)$.

5. *Let $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Then, $A$ and $B$ are similar. Find the characteristic polynomials, eigenvalues, and eigenvectors of $A$ and $B$.*

   **Solution.** The characteristic polynomial of $A$ is $\lambda^2 - 1 = (\lambda - 1)(\lambda + 1)$, so the eigenvalues of $A$ are 1 and $-1$. The 1-eigenspace of $A$ is the kernel of $A - I = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$, which is spanned by $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. The $-1$-eigenspace of $A$ is the kernel of $A + I = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, which is spanned by $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

   The characteristic polynomial of $B$ is $(1 - \lambda)(-1 - \lambda) = \lambda^2 - 1$, which is the same as the characteristic polynomial of $A$. Therefore, $B$ has the same eigenvalues, 1 and $-1$. The 1-eigenspace of $B$ is the kernel of $B - I = \begin{bmatrix} 0 & 0 \\ 0 & -2 \end{bmatrix}$, which is spanned by $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. The $-1$-eigenspace of $B$ is the kernel of $B + I = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$, which is spanned by $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

Thus, although $A$ and $B$ have the same characteristic polynomials and the same eigenvalues, they do not have the same eigenvectors.

6. *Is the matrix* $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ *similar to the matrix* $\begin{bmatrix} 1 & 4 \\ 2 & 6 \end{bmatrix}$*?*

   **Solution.** The two matrices have different traces, so they can't be similar.

# Diagonalization

Let $L$ be the line $y = 2x$ in $\mathbb{R}^2$. Let $\mathrm{ref}_L$ be reflection over $L$, and let $A$ be the standard matrix of $\mathrm{ref}_L$.

1. *Find an eigenbasis* $\mathfrak{B}$ *for* $A$.

   **Solution.** Let $\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\vec{v}_2 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$. Then, $\vec{v}_1$ lies on the line $L$, so $A\vec{v}_1 = \mathrm{ref}_L(\vec{v}_1) = \vec{v}_1$. On the other hand, $\vec{v}_2$ is perpendicular to $L$, so $A\vec{v}_2 = \mathrm{ref}_L(\vec{v}_2) = -\vec{v}_2$. Thus, $\vec{v}_1$ and $\vec{v}_2$ are both eigenvectors of $A$. Since $(\vec{v}_1, \vec{v}_2)$ is clearly a basis of $\mathbb{R}^2$, $\mathfrak{B} = (\vec{v}_1, \vec{v}_2)$ is an eigenbasis for $A$.

2. *Find the* $\mathfrak{B}$*-matrix of* $\mathrm{ref}_L$.

   **Solution.** If $D$ is the $\mathfrak{B}$-matrix of $\mathrm{ref}_L$, the columns of $D$ are $[\mathrm{ref}_L(\vec{v}_1)]_{\mathfrak{B}}$ and $[\mathrm{ref}_L(\vec{v}_2)]_{\mathfrak{B}}$. Since $\mathrm{ref}_L(\vec{v}_1) = \vec{v}_1 = 1 \cdot \vec{v}_1 + 0 \cdot \vec{v}_2$, $[\mathrm{ref}_L(\vec{v}_1)]_{\mathfrak{B}} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Since $\mathrm{ref}_L(\vec{v}_2) = -\vec{v}_2 = 0 \cdot \vec{v}_1 + (-1) \cdot \vec{v}_2$, $[\mathrm{ref}_L(\vec{v}_2)]_{\mathfrak{B}} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$. So, the $\mathfrak{B}$-matrix of $\mathrm{ref}_L$ is $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$.

3. *Find* $A$ *(the standard matrix of* $\mathrm{ref}_L$*).*

   **Solution.** If $S = [\vec{v}_1 \ \vec{v}_2]$, then $A = SDS^{-1}$. In this case, this says that

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

**True / False**

1. *If* $A$ *is diagonalizable, then* $A^2$ *is diagonalizable.*

   **Solution.** True. Since $A$ is diagonalizable, there is an invertible matrix $S$ such that $S^{-1}AS$ is a diagonal matrix. Then, $S^{-1}A^2S = (S^{-1}AS)(S^{-1}AS)$ is the product of two diagonal matrices, which is a diagonal matrix. Therefore, $A^2$ is diagonalizable.

2. *If* $A$ *and* $B$ *are* $n \times n$ *diagonalizable matrices, then* $A + B$ *is diagonalizable.*

   **Solution.** False. For example, $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$ are both diagonalizable, but $A + B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ is not.

3. *If* $A$ *and* $B$ *are* $n \times n$ *diagonalizable matrices with the same eigenvectors, then* $AB$ *is diagonalizable.*

   **Solution.** True. Since $A$ is diagonalizable, there is an eigenbasis for $A$, say $(\vec{v}_1, \dots, \vec{v}_n)$. Since $B$ has the same eigenvectors as $A$, $(\vec{v}_1, \dots, \vec{v}_n)$ is also an eigenbasis for $B$. Therefore, if $S = [\vec{v}_1 \ \dots, \vec{v}_n]$, $S^{-1}AS$ and $S^{-1}BS$ are both diagonal matrices. If we multiply two diagonal matrices, we get another diagonal matrix. Thus, $(S^{-1}AS)(S^{-1}BS) = S^{-1}ABS$ is a diagonal matrix, so $AB$ is diagonalizable.

4. *If* $A$ *is diagonalizable, then* $A^T$ *is diagonalizable.*

   **Solution.** True. Since $A$ is diagonalizable, there is an invertible matrix $S$ such that $S^{-1}AS$ is a diagonal matrix $D$. Then, $(S^{-1}AS)^T$ is equal to $D^T$, which is the same as $D$. On the other hand, $(S^{-1}AS)^T$ is just $S^TA^T(S^{-1})^T = S^TA^T(S^T)^{-1}$. Thus, if $R = (S^T)^{-1}$, then $R^{-1}A^TR = D$, so $A^T$ is diagonalizable.

5. *If $A$ is an $n \times n$ matrix with $n$ distinct eigenvalues, then $A$ is diagonalizable.*

   **Solution.** True. The geometric multiplicity of any eigenvalue is at least 1, so, if $A$ has $n$ distinct eigenvalues, then the sum of the geometric multiplicities of the eigenvalues is $n$. Therefore, $A$ has a basis of eigenvectors, so $A$ is diagonalizable.

6. *If $A$ is a diagonalizable matrix and $\lambda$ is an eigenvalue of $A$, then the algebraic multiplicity of $\lambda$ is equal to the geometric multiplicity of $\lambda$.*

   **Solution.** True. Let's say $A$ is an $n \times n$ matrix, and let $\lambda$ be an eigenvalue. Then, we know:

   - The algebraic multiplicity of $\lambda$ is greater than or equal to the geometric multiplicity of $\lambda$.
   - The sum of the algebraic multiplicities of all eigenvalues is at most $n$ (the degree of the characteristic polynomial of $A$).
   - Since $A$ is diagonalizable, the geometric multiplicities of the eigenvalues of $A$ must add up to $n$.

   Thus, the algebraic multiplicity of $\lambda$ must be the same as the geometric multiplicity of $\lambda$ (otherwise the sum of the algebraic multiplicities would be greater than $n$).

7. *If $A$ and $B$ are both diagonalizable and if $A$ and $B$ have the same eigenvalues with the same geometric multiplicities, then $A$ is similar to $B$.*

   **Solution.** True. Since $A$ is diagonalizable, the geometric multiplicities of its eigenvalues add up to $n$. That is, $A$ has $n$ eigenvalues $\lambda_1, \ldots, \lambda_n$ (if we count the eigenvalues with their geometric multiplicities). Then, $A$ is similar to the diagonal matrix

$$D = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}.$$

   Since $B$ has the same eigenvalues with the same geometric multiplicities and $B$ is diagonalizable, $B$ is also similar to $D$. Thus, $A$ and $B$ are both similar to $D$, so they must be similar to each other.

# Stability

Let $A$ be the matrix $\begin{bmatrix} 1/2 & 1/2 \\ -1/2 & 1/2 \end{bmatrix}$. You are given the following information:

- $\vec{v}_1 = \begin{bmatrix} 1 \\ i \end{bmatrix}$ is an eigenvector with eigenvalue $\frac{1+i}{2}$.
- $\vec{v}_2 = \begin{bmatrix} 1 \\ -i \end{bmatrix}$ is an eigenvector with eigenvalue $\frac{1-i}{2}$.

1. *Write $A$ as $SDS^{-1}$ where $S$ is an invertible matrix and $D$ is a diagonal matrix (both with complex entries).*

   **Solution.** We can take $S = [\vec{v}_1 \ \vec{v}_2] = \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix}$ and $D = \begin{bmatrix} (1+i)/2 & 0 \\ 0 & (1-i)/2 \end{bmatrix}$.

2. *Find a closed formula for $A^t$.*

   **Solution.** Since $A = SDS^{-1}$, $A^t = SD^tS^{-1}$. First, observe that

$$D = \begin{bmatrix} (1+i)/2 & 0 \\ 0 & (1-i)/2 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}}e^{\pi i/4} & 0 \\ 0 & \frac{1}{\sqrt{2}}e^{-\pi i/4} \end{bmatrix}.$$

   Therefore,

$$D^t = \begin{bmatrix} 2^{-t/2}e^{\pi it/4} & 0 \\ 0 & 2^{-t/2}e^{-\pi it/4} \end{bmatrix} = 2^{-t/2}\begin{bmatrix} e^{\pi it/4} & 0 \\ 0 & e^{-\pi it/4} \end{bmatrix},$$

   and

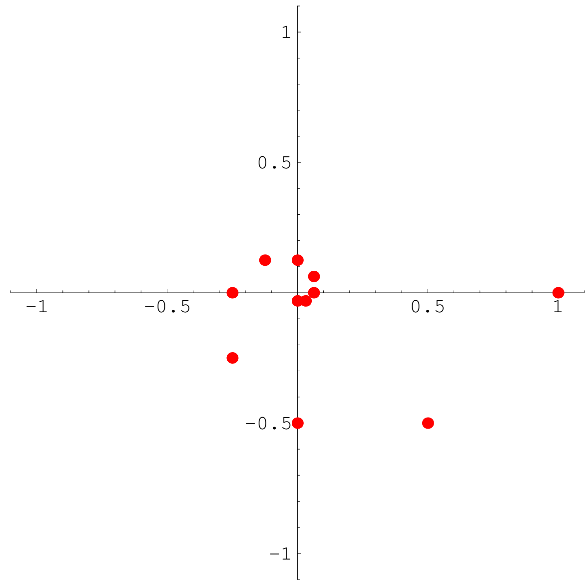
$$\begin{aligned} A^t &= 2^{-t/2}\begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix}\begin{bmatrix} e^{\pi it/4} & 0 \\ 0 & e^{-\pi it/4} \end{bmatrix}\begin{bmatrix} 1/2 & -i/2 \\ 1/2 & i/2 \end{bmatrix} \\ &= 2^{-t/2}\begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix}\begin{bmatrix} \frac{1}{2}e^{\pi it/4} & -\frac{i}{2}e^{\pi it/4} \\ \frac{1}{2}e^{-\pi it/4} & \frac{i}{2}e^{-\pi it/4} \end{bmatrix} \\ &= 2^{-t/2}\begin{bmatrix} \frac{1}{2}e^{\pi it/4} + \frac{1}{2}e^{-\pi it/4} & -\frac{i}{2}e^{\pi it/4} + \frac{i}{2}e^{-\pi it/4} \\ \frac{i}{2}e^{\pi it/4} - \frac{i}{2}e^{-\pi it/4} & \frac{1}{2}e^{\pi it/4} + \frac{1}{2}e^{-\pi it/4} \end{bmatrix} \end{aligned}$$

   Now, we use the fact that $e^{\pi it/4} = \cos\frac{\pi t}{4} + i\sin\frac{\pi t}{4}$ and $e^{-\pi it/4} = \cos\frac{\pi t}{4} - i\sin\frac{\pi t}{4}$ to rewrite this as

$$A^t = 2^{-t/2}\begin{bmatrix} \cos\frac{\pi t}{4} & \sin\frac{\pi t}{4} \\ -\sin\frac{\pi t}{4} & \cos\frac{\pi t}{4} \end{bmatrix}.$$

3. *If $\vec{x}(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, what do the trajectories of the dynamical system $\vec{x}(t+1) = A\vec{x}(t)$ look like? What is the long-term behavior of $\vec{x}(t)$? What if $\vec{x}(0)$ is some other vector?*

   **Solution.** If $\vec{x}(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, the vectors $\vec{x}(0), \ldots, \vec{x}(10)$ are shown:



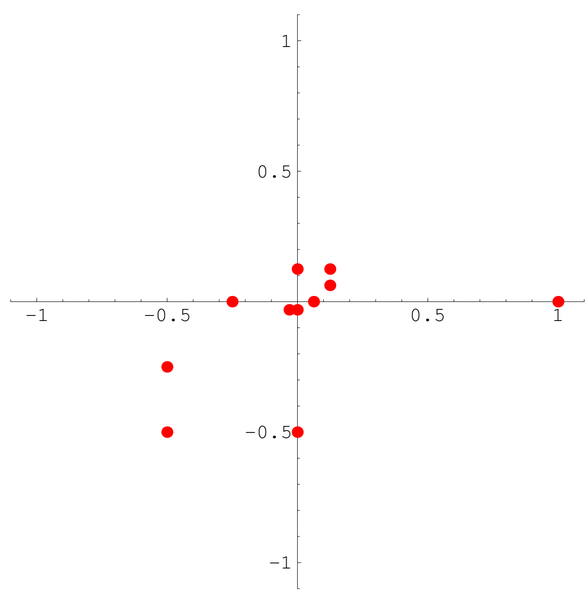

As $t$ gets large, $A^t$ tends toward the zero matrix, so $\vec{x}(t)$ tends to $\vec{0}$.

The same thing is true even if $\vec{x}(0)$ is another vector. $A^t$ is a matrix which rotates by $\frac{\pi t}{4}$ and scales by $2^{-t/2}$. Thus, the trajectories of $\vec{x}(t+1) = A\vec{x}(t)$ always spiral inward and tend to $\vec{0}$.

4. *If $B$ is any $2 \times 2$ matrix with the same eigenvalues as $A$, what do the trajectories of $\vec{x}(t+1) = B\vec{x}(t)$ look like?*

   **Solution.** Since $A$ and $B$ have two distinct eigenvalues, they are both diagonalizable. Since $A$ and $B$ have the same eigenvalues, they are similar. That is, $B = P^{-1}AP$ for some invertible matrix $P$. The trajectories of the dynamical system $\vec{x}(t+1) = B\vec{x}(t)$ are given by $\vec{x}(t) = B^t\vec{x}(0) = P^{-1}A^tP\vec{x}(0)$. We know that the vectors $A^tP\vec{x}(0)$ spiral inward as $t$ increases, so the vectors $P^{-1}A^tP\vec{x}(0)$ must spiral inward as well. In particular, every trajectory tends toward $\vec{0}$.

   For example, here is the trajectory when $P = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$, $B = P^{-1}AP$, and $\vec{x}(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.



# Symmetric Matrices

1. *If $A$ is orthogonally diagonalizable, what is the relationship between $A^T$ and $A$?*

   **Solution.** If $A$ is orthogonaly diagonalizable, then there exists an orthogonal matrix $S$ and a diagonal matrix $D$ such that $S^{-1}AS = D$. So, we can write $A = SDS^{-1} = SDS^T$ (remember that $S^{-1} = S^T$ since $S$ is orthogonal). Therefore,

$$\begin{aligned} A^T &= (S^T)^T D^T S^T \\ &= SD^T S^T \\ &= SDS^T \text{ since } D \text{ is diagonal} \\ &= A \end{aligned}$$

   Thus, $A^T = A$, so $A$ is symmetric.

2. (a) *Let $A$ be a matrix, $\vec{v}$ be any vector, and $\vec{w}$ be an eigenvector of $A$ with eigenvalue $\mu$. What is the relationship between $\vec{v}^T A\vec{w}$ and $\vec{v} \cdot \vec{w}$?*

   **Solution.** We know that $A\vec{w} = \mu\vec{w}$, so $\vec{v}^T A\vec{w} = \mu(\vec{v}^T \vec{w}) = \mu(\vec{v} \cdot \vec{w})$.

   (b) *Let $A$ be a matrix, $\vec{w}$ be any vector, and $\vec{v}$ be an eigenvector of $A^T$ with eigenvalue $\lambda$. What is the relationship between $\vec{v}^T A\vec{w}$ and $\vec{v} \cdot \vec{w}$?*

   **Solution.** We know that $A^T\vec{v} = \lambda\vec{v}$. Taking the transpose of both sides, $\vec{v}^T A = \lambda\vec{v}^T$. Thus, $\vec{v}^T A\vec{w} = \lambda\vec{v}^T\vec{w} = \lambda(\vec{v} \cdot \vec{w})$.

   (c) *If $A$ is a symmetric matrix and $\vec{v}$ and $\vec{w}$ are eigenvectors with different eigenvalues, what can you say about $\vec{v} \cdot \vec{w}$?*

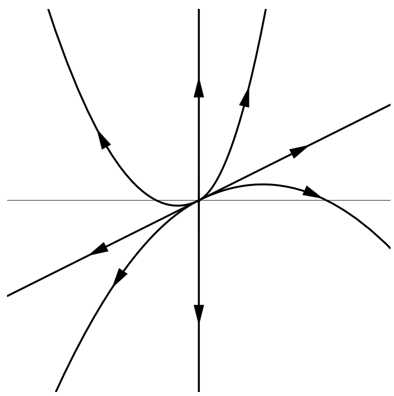
   **Solution.** Let $\lambda$ be the eigenvalue of $\vec{v}$ and $\mu$ be the eigenvalue of $\vec{w}$. That is, $A\vec{v} = \lambda\vec{v}$ and $A\vec{w} = \mu\vec{w}$. Since $A\vec{w} = \mu\vec{w}$, $\vec{v}^T A\vec{w} = \mu(\vec{v} \cdot \vec{w})$ by part (a). Since $A$ is symmetric, $A = A^T$, so $A^T\vec{v} = A\vec{v} = \lambda\vec{v}$. Therefore, by part (b), $\vec{v}^T A\vec{w} = \lambda(\vec{v} \cdot \vec{w})$.

   Thus, we see that $\vec{v}^T A\vec{w}$ is equal to both $\mu(\vec{v} \cdot \vec{w})$ and $\lambda(\vec{v} \cdot \vec{w})$. Therefore, $(\mu - \lambda)(\vec{v} \cdot \vec{w}) = 0$. Since $\lambda \neq \mu$, it must be the case that $\vec{v} \cdot \vec{w} = 0$.

# Phase Portraits

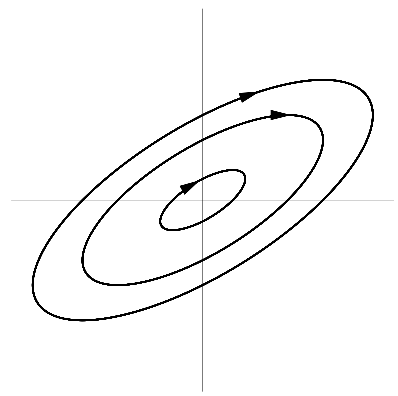
Each of the following is the phase portrait of a continuous dynamical system $\frac{d\vec{x}}{dt} = A\vec{x}$ where $A$ is a real $2 \times 2$ matrix. What can you say about the eigenvalues, determinant, and trace of $A$? In which cases is $\vec{0}$ an asymptotically stable equilibrium?

1.



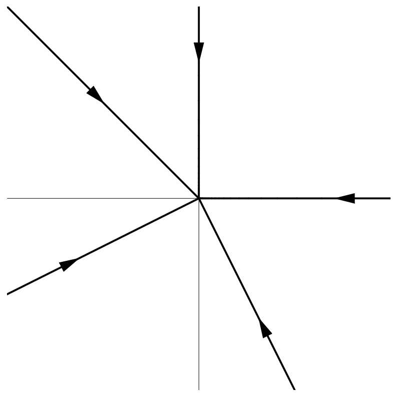
**Solution.** We have straight trajectories along two different lines, so $A$ has 2 distinct real eigenvalues. All trajectories tend away from zero, so the eigenvalues must be positive. Since $\det A$ is the product of the eigenvalues, $\det A$ must be positive. Since $\operatorname{tr} A$ is the sum of the eigenvalues, $\operatorname{tr} A$ must also be positive. The trajectories don't all tend toward $\vec{0}$, so $\vec{0}$ is not an asymptotically stable equilibrium.

2.



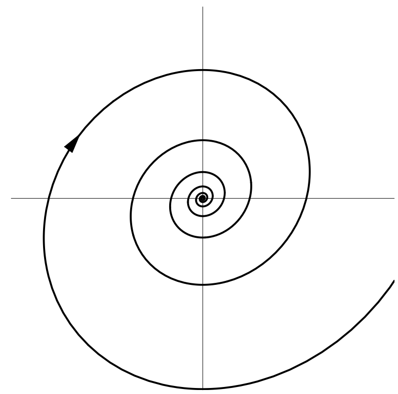
**Solution.** Since there are no straight trajectories, $A$ has no real eigenvalues; therefore, $A$ must have complex eigenvalues. Since complex eigenvalues come in conjugate pairs, the eigenvalues of $A$ must be $\lambda = p + iq$ and $\overline{\lambda} = p - iq$ for some $p, q$ with $q \neq 0$. If $\vec{v}_1$ and $\vec{v}_2$ are the corresponding eigenvectors, we know the solutions of $\frac{d\vec{x}}{dt} = A\vec{x}$ look like $\vec{x}(t) = c_1 e^{pt} e^{iqt} \vec{v}_1 + c_2 e^{pt} e^{-iqt} \vec{v}_2$ for some constants $c_1$ and $c_2$. Since the trajectories are ellipses, it must be the case that $p = 0$. (If $p > 0$, then $e^{pt}$ will grow as $t$ gets large, so the trajectories will get further and further from the origin; if $p < 0$, then $e^{pt}$ will get closer and closer to 0.) Thus, the eigenvalues of $A$ have the form $\pm iq$ for some $q \neq 0$. Then, $\det A = (iq)(-iq) = q^2 > 0$ and $\operatorname{tr} A = (iq) + (-iq) = 0$. The trajectories don't tend toward $\vec{0}$, so $\vec{0}$ is not an asymptotically stable equilibrium.

3.



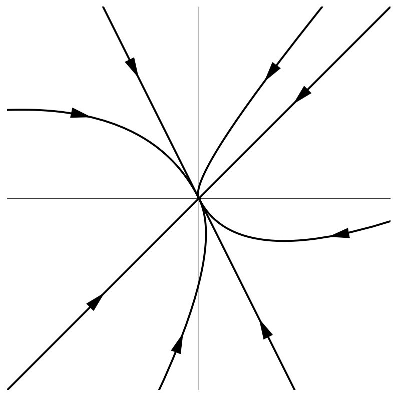
**Solution.** We have straight trajectories along every line, which says that the whole plane must be an eigenspace; that is, $A$ has only 1 real eigenvalue $\lambda$, and this eigenvalue has geometric multiplicity 2. Since the trajectories tend toward $\vec{0}$, $\lambda < 0$. Thus, $\det A = \lambda^2 > 0$ and $\operatorname{tr} A = 2\lambda < 0$. Also, $\vec{0}$ is an asymptotically stable equilibrium because all trajectories tend toward $\vec{0}$.

4.



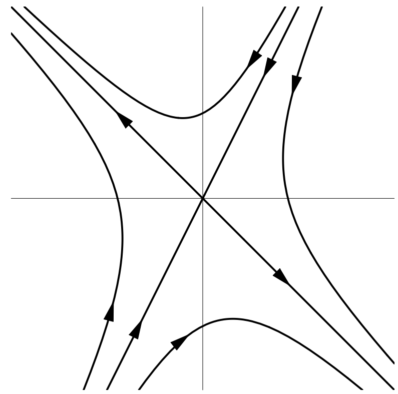
**Solution.** By the same argument as in #2, $A$ has a pair of complex eigenvalues $\lambda = p + iq$ and $\overline{\lambda} = p - iq$. Since the trajectories spiral inward and tend to $\vec{0}$, it must be the case that $p < 0$. Thus, $\det A = (p + iq)(p - iq) = p^2 + q^2 > 0$ and $\det A = (p + iq) + (p - iq) = 2p < 0$. Since all trajectories tend toward $\vec{0}$, $\vec{0}$ is an asymptotically stable equilibrium.

5.



**Solution.** We have straight trajectories along two different lines, so $A$ has 2 distinct real eigenvalues. All trajectories tend toward $\vec{0}$, so the eigenvalues must be negative. Therefore, $\det A > 0$ and $\operatorname{tr} A < 0$. Also, $\vec{0}$ is an asymptotically stable equilibrium.

6.

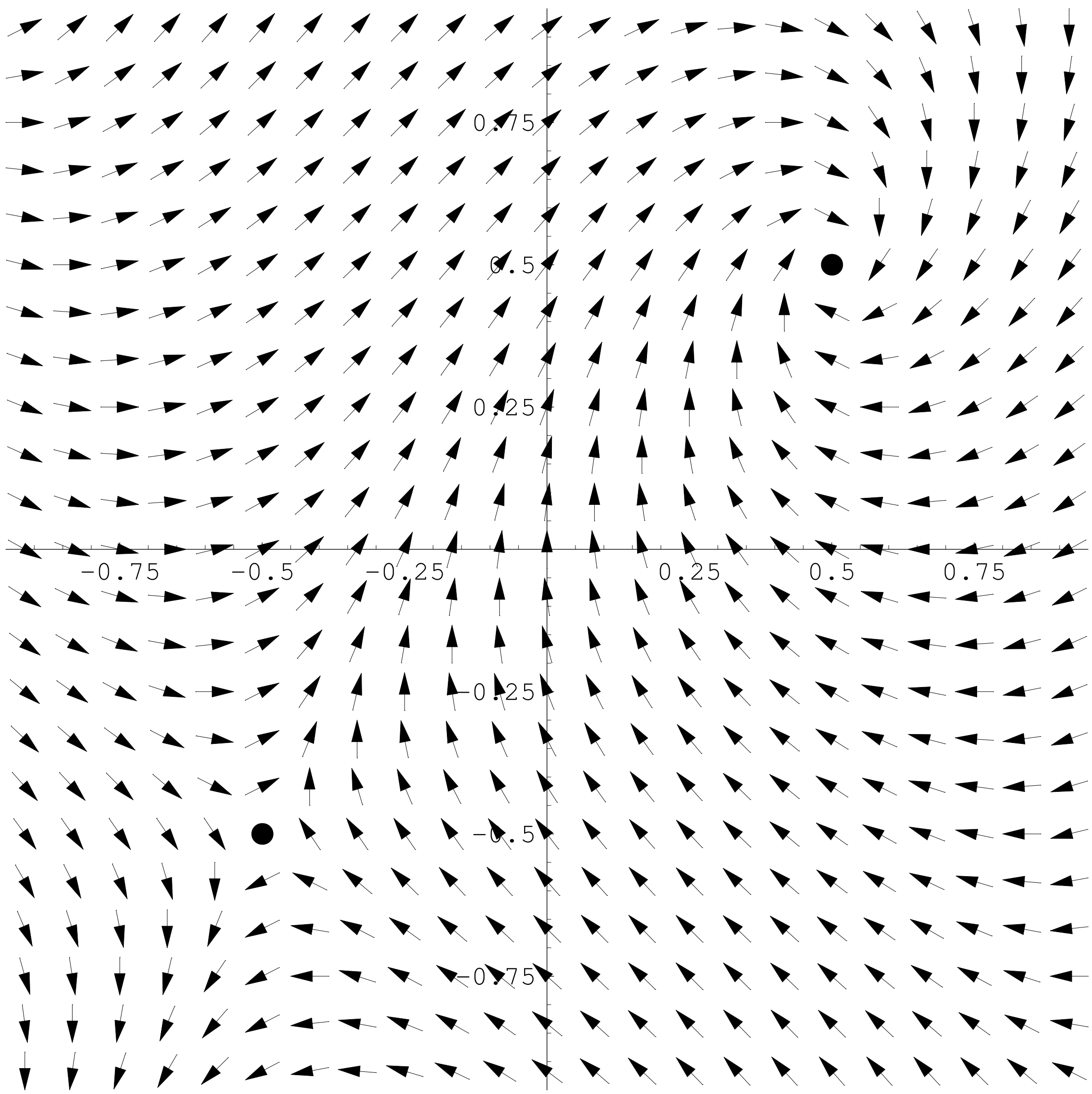


**Solution.** We have straight trajectories along two different lines, so $A$ has 2 distinct real eigenvalues. Along one straight line, the trajectories tend toward zero; therefore, the corresponding eigenvalue must be negative. Along the other straight line, the trajectories tend away from zero, so the corresponding eigenvalue must be positive. Thus, $A$ has two real eigenvalues, one positive and one negative. In particular, $\det A$ is negative. However, we can't say anything about $\operatorname{tr} A$ — it is the sum of a positive number and a negative number, which could be anything. $\vec{0}$ is not an asymptotically stable equilibrium since not all trajectories tend toward $\vec{0}$.

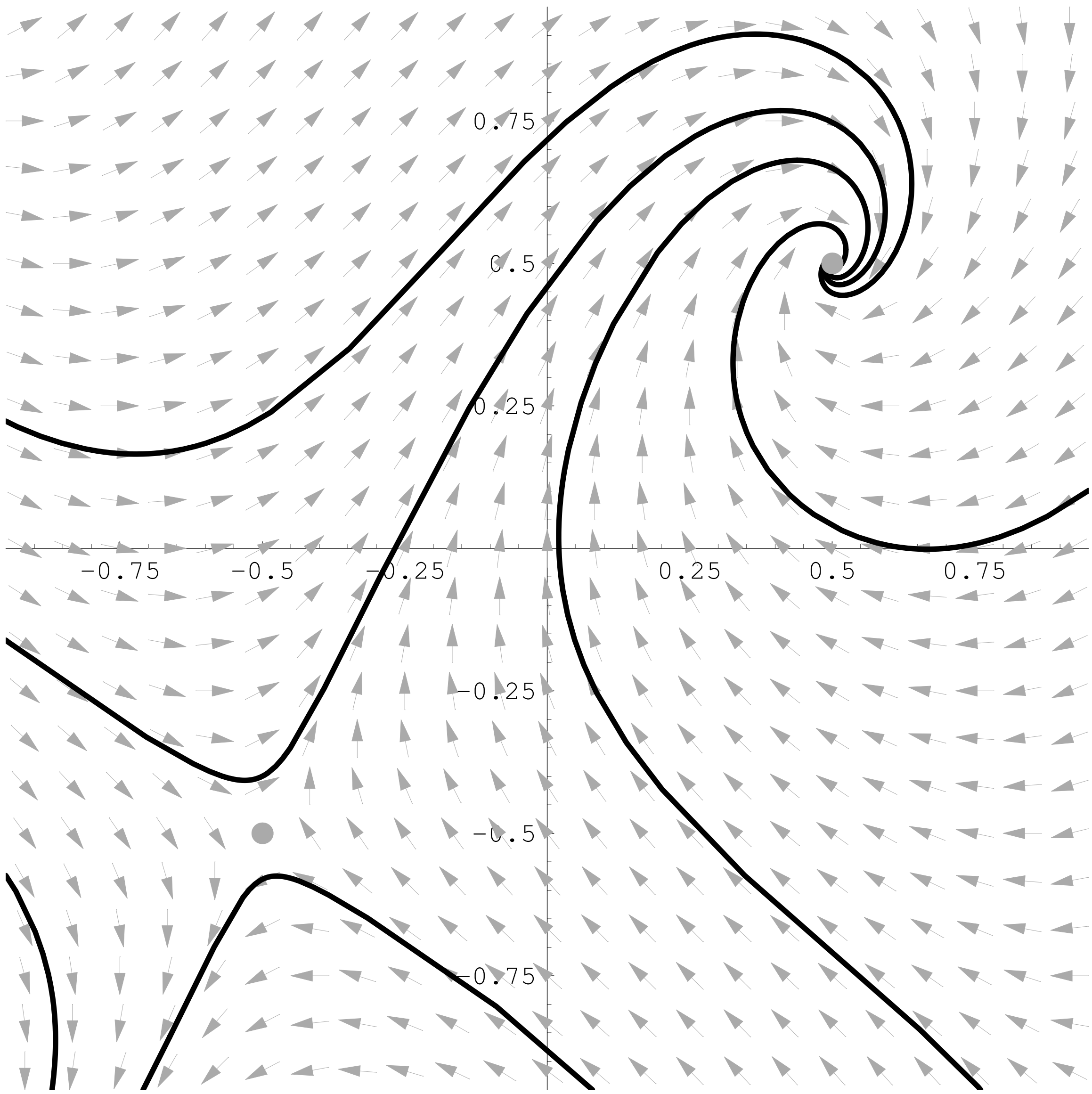
# Nonlinear Dynamical Systems

Here is the direction field for the nonlinear system

$$\begin{aligned}\frac{dx}{dt} &= \sin\frac{\pi(y-x)}{2}\\ \frac{dy}{dt} &= \cos\frac{\pi(y+3x)}{4}\end{aligned}$$



Here is the direction field with some trajectories.



# Linear Transformations

Which of the following are linear spaces over $\mathbb{R}$? Over $\mathbb{C}$?

1. *The set $P$ of polynomials with real coefficients.*

   **Solution.** This is a linear space over $\mathbb{R}$:

   - If $f, g \in P$, then $f$ and $g$ are polynomials with real coefficients, so $f+g$ is also a polynomial with real coefficients; that is, $f+g \in P$.

   - If $c \in \mathbb{R}$ and $f \in P$, then $cf$ is a polynomial with real coefficients, so $cf \in P$.

   However, $P$ is not a linear space over $\mathbb{C}$: if $f$ is a polynomial with real coefficients and $c$ is a complex number, then $cf$ does not have real coefficients, so $cf \notin P$.

2. $V = \{f \in C^\infty(\mathbb{R}) : f(1) = i\}$.

   **Solution.** This is not a linear space: if $f, g \in V$, then $f(1) = i$ and $g(1) = i$. However, $(f+g)(1) = f(1) + g(1) = 2i$, so $f + g \notin V$. Thus, $V$ is not closed under addition, so $V$ is not a linear space (over $\mathbb{R}$ or over $\mathbb{C}$).

3. *The complex numbers $\mathbb{C}$.*

   **Solution.** $\mathbb{C}$ is a linear space over $\mathbb{R}$:

   - If $x, y \in \mathbb{C}$, then $x + y \in \mathbb{C}$. (That is, if we add two complex numbers, we get another complex number.)

   - If $x \in \mathbb{C}$ and $c \in \mathbb{R}$, then $cx \in \mathbb{C}$. (If we multiply a complex number by a real number, we get another complex number.)

   $\mathbb{C}$ is also a linear space over $\mathbb{C}$:

   - We already explained that $\mathbb{C}$ is closed under addition.

   - If $x \in \mathbb{C}$ and $c \in \mathbb{C}$, then $cx \in \mathbb{C}$.

Which of the following are linear transformations?

4. *$D$ from $C^\infty(\mathbb{R})$ to $C^\infty(\mathbb{R})$ defined by $D(f) = f'$ (over $\mathbb{C}$).*

   **Solution.** $D$ is linear:

   - $D(f+g) = (f+g)' = f' + g' = D(f) + D(g)$.

   - For any $c \in \mathbb{C}$, $D(cf) = (cf)' = cf' = cD(f)$.

   $D$ is not invertible: although every function in $C^\infty(\mathbb{R})$ has an antiderivative, the antiderivative is not unique (because there is a constant of integration when you integrate).

   The kernel of $D$ is the set of constant functions. The image of $D$ is $C^\infty(\mathbb{R})$ (because every function in $C^\infty(\mathbb{R})$ has an antiderivative).

5. *$T$ from $\mathbb{R}^2$ to $\mathbb{C}$ defined by* $T\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) = a + bi$ *(over $\mathbb{R}$).*

**Solution.** $T$ is a linear transformation:

- $T\left(\begin{bmatrix} a_1 \\ b_1 \end{bmatrix}\right) + T\left(\begin{bmatrix} a_2 \\ b_2 \end{bmatrix}\right) = (a_1 + b_1 i) + (a_2 + b_2 i) = (a_1 + a_2) + (b_1 + b_2)i = T\left(\begin{bmatrix} a_1 + a_2 \\ b_1 + b_2 \end{bmatrix}\right)$.

- For any $c \in \mathbb{R}$, $T\left(c\begin{bmatrix} a \\ b \end{bmatrix}\right) = T\left(\begin{bmatrix} ca \\ cb \end{bmatrix}\right) = ca + (cb)i = c(a + bi) = cT\left(\begin{bmatrix} a \\ b \end{bmatrix}\right)$.

$T$ is invertible; its inverse is $S(a + bi) = \begin{bmatrix} a \\ b \end{bmatrix}$. The kernel of $T$ is $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$, and the image of $T$ is $\mathbb{C}$.

6. *$T$ from $C^\infty(\mathbb{R})$ to $C^\infty(\mathbb{R})$ defined by* $T(f) = f' + \sin x$ *(over $\mathbb{C}$).*

**Solution.** $T$ is not a linear transformation: $T(cf) = (cf)' + \sin x = cf' + \sin x$ but $cT(f) = cf' + c\sin x$.

7. *$I$ from $C^\infty(\mathbb{R})$ to $\mathbb{C}$ defined by* $I(f) = \int_0^1 f(x)\ dx$ *(over $\mathbb{C}$).*

**Solution.** $I$ is a linear transformation:

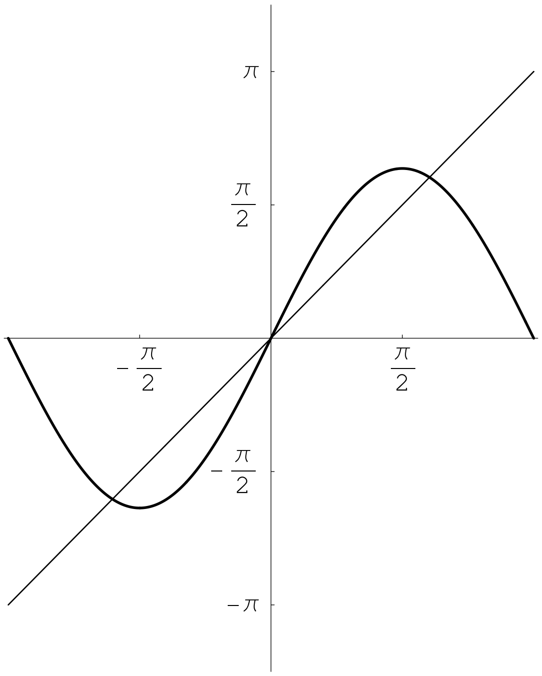
- $I(f + g) = \int_0^1 (f + g)(x)\ dx = \int_0^1 f(x)\ dx + \int_0^1 g(x)\ dx = I(f) + I(g)$.

- If $c \in \mathbb{C}$, then $I(cf) = \int_0^1 (cf)(x)\ dx = c\int_0^1 f(x)\ dx = cI(f)$.

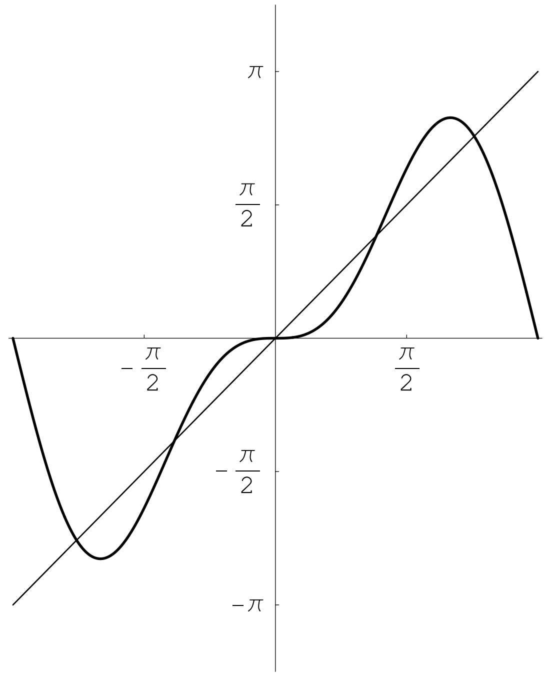
$I$ is not invertible; given a complex number $a$, there is not a unique $f \in C^\infty(\mathbb{R})$ such that $\int_0^1 f(x)\ dx = a$.

# Fourier Series

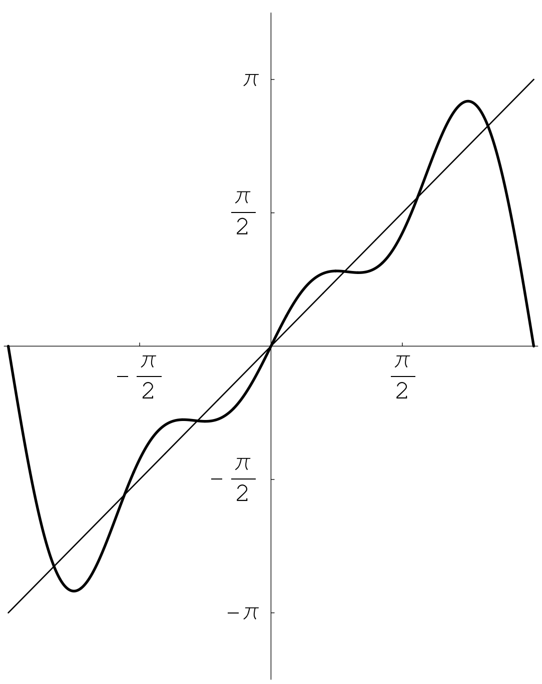
Let $f(t) = t$. Here are the graphs of $f_n = \mathrm{proj}_{T_n} f$ for various values of $n$.



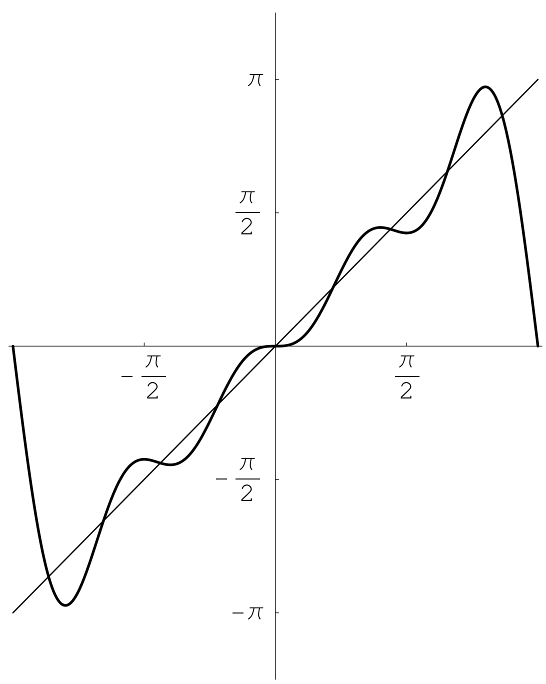

$n = 1$



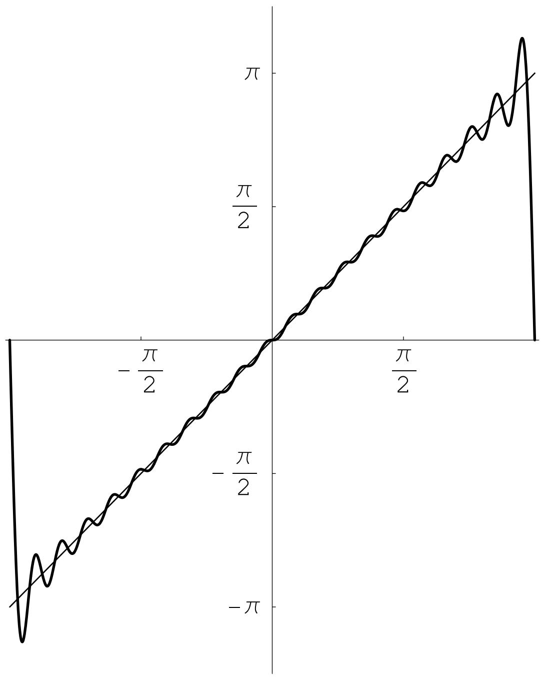

$n = 2$



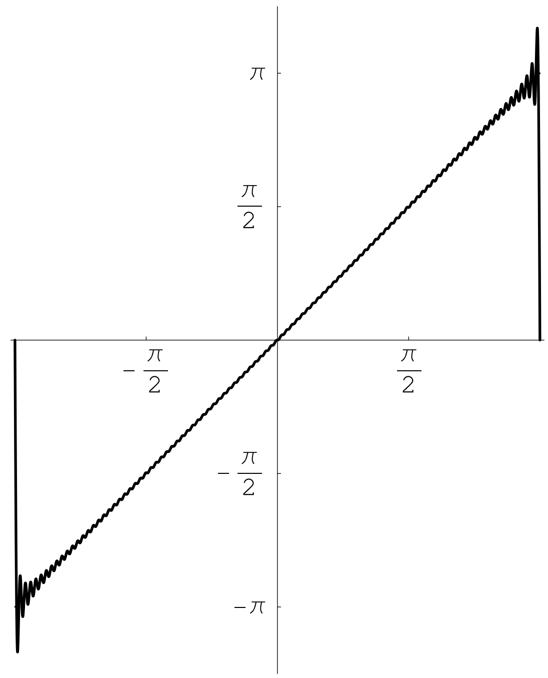

$n = 3$




$n = 4$




$n = 20$




$n = 100$

## Useful Trigonometric Identities

When finding Fourier series, you will sometimes need to integrate products of sines and cosines. There are some trigonometric identities which are useful for this; these identities essentially all come from the fact that $e^{i(A+B)} = e^{iA}e^{iB}$. Using the fact that $e^{i\theta} = \cos\theta + i\sin\theta$, we have

$$\begin{aligned}
\cos(A+B) + i\sin(A+B) &= e^{i(A+B)} \\
&= e^{iA}e^{iB} \\
&= (\cos A + i\sin A)(\cos B + i\sin B) \\
&= (\cos A\cos B - \sin A\sin B) + i(\sin A\cos B + \cos A\sin B)
\end{aligned}$$

Equating the real and imaginary parts, we have

$$\begin{aligned}
\cos(A+B) &= \cos A\cos B - \sin A\sin B \\
\sin(A+B) &= \sin A\cos B + \cos A\sin B
\end{aligned}$$

If we replace $B$ by $-B$ and use the fact that $\cos(-B) = \cos B$ while $\sin(-B) = -\sin B$, we have

$$\begin{aligned}
\cos(A-B) &= \cos A\cos B + \sin A\sin B \\
\sin(A-B) &= \sin A\cos B - \cos A\sin B
\end{aligned}$$

We can now use these to find identities for products. For instance, $\sin(A+B) + \sin(A-B) = 2\sin A\cos B$, so

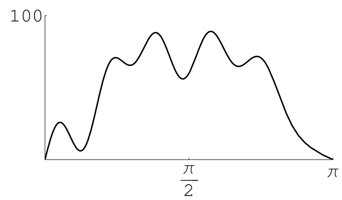
$$\sin A\cos B = \frac{1}{2}\sin(A+B) + \frac{1}{2}\sin(A-B),$$

which is a convenient formula if you need to integrate $\sin A\cos B$.

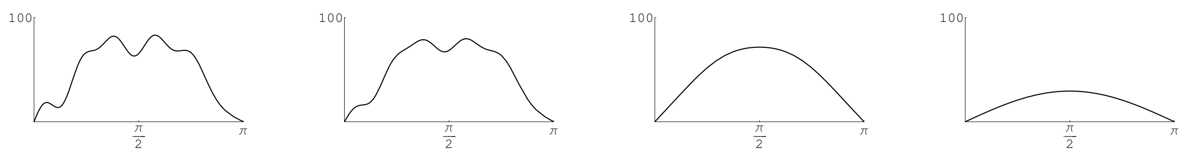
# Partial Differential Equations

## The Heat Equation

Imagine that we have a metal bar of length $\pi$ which is thermally insulated (meaning that it doesn't lose heat to its surroundings). Suppose that, at time 0, the bar is heated so that a graph of the temperature along the bar looks like this:




Here, the $x$-axis represents length along the bar and the $y$-axis represents temperature. For example, the temperature at both ends of the bar is 0. Suppose that the ends of the bar are kept at temperature 0 while nothing is done to the rest of the bar. Since the ends of the bar are kept cold, we can guess that the temperature along the whole bar will eventually tend to 0. In fact, we will see that this is exactly what happens.




time 0.005 time 0.01 time 0.1 time 1

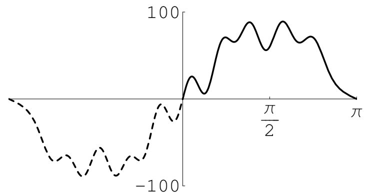
At any particular time $t$, we can represent the temperature along the bar as a function of $x$. So, we should really think of the temperature as a function of both time and position. Let $f(t,x)$ be the temperature of the bar at time $t$ and position $x$. Since the ends of the bar are always kept at temperature 0, we know that $f(t,0) = f(t,\pi) = 0$ for all $t$. Also, we are given a graph of $f(0,x)$ — the initial temperature distribution of the bar. In order to find out what $f(t,x)$ is at any time, we need to understand how the bar cools.

It turns out that the way the bar cools can be modeled by the <u>heat equation</u> $\frac{\partial f}{\partial t} = \mu \frac{\partial^2 f}{\partial x^2}$; here, $\mu$ is a positive constant indicating how well the bar conducts heat (metal would have a much larger $\mu$ than plastic). Since this differential equation involves partial derivatives, we call it a <u>partial differential equation</u>.

## Fourier Series on $[0,\pi]$

We would like to use Fourier series to solve the heat equation. At any fixed time $t$, the temperature is just a function of $x$, so we should be able to write the Fourier series for that function. The only problem is that $x$ goes from 0 to $\pi$ rather than from $-\pi$ to $\pi$. That is, we have a function $g$ on $[0,\pi]$ (representing the temperature at a given time), but we only know how to write Fourier series for functions on $[-\pi,\pi]$. To fix this problem, we simply extend $g$ to a function $\tilde{g}$ on $[-\pi,\pi]$.

There are lots of ways to pick $\tilde{g}$; however, if we're smart about picking $\tilde{g}$, we can save ourselves a lot of work when we compute the Fourier series of $\tilde{g}$. Recall that the Fourier coefficients of $\tilde{g}$ are the inner products $\left\langle \tilde{g}(x), \frac{1}{\sqrt{2}} \right\rangle$, $\langle \tilde{g}(x), \sin(nx) \rangle$, and $\langle \tilde{g}(x), \cos(nx) \rangle$. Notice that $\frac{1}{\sqrt{2}}$ and $\cos(nx)$ are both even functions. In problem #2(c) of your Monday homework, you found that the inner product of an even function and an odd function is 0; thus, if we can make $\tilde{g}$ an odd function, then it will only have Fourier coefficients for the $\sin(nx)$ terms. Visually, here is how we extend $g$ to an odd function $\tilde{g}$ on $[-\pi,\pi]$:




Now, $\left\langle \tilde{g}(x), \frac{1}{\sqrt{2}} \right\rangle = 0$ and $\langle \tilde{g}(x), \cos(nx) \rangle = 0$, so $\tilde{g}(x)$ has a Fourier series $\sum_{n=1}^{\infty} a_n \sin(nx)$ where $a_n = \langle \tilde{g}(x), \sin(nx) \rangle$. Since both $\tilde{g}$ and $\sin(nx)$ are odd functions, #2(b) of your Monday homework says that

$$a_n = \frac{2}{\pi} \int_0^{\pi} \tilde{g}(x) \sin(nx)\, dx.$$

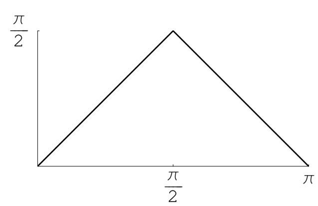
But on $[0,\pi]$, $\tilde{g}(x)$ is just $g(x)$! So, we really have the following.

**Fact 1.** *If $g(x)$ is a (reasonable) continuous function on $[0,\pi]$ such that $g(0) = g(\pi) = 0$, then we can write $g(x) = \sum_{n=1}^{\infty} a_n \sin(nx)$ where*

$$a_n = \frac{2}{\pi} \int_0^{\pi} g(x) \sin(nx)\, dx.$$

Note: Without the condition $g(0) = g(\pi) = 0$, we can still write the Fourier series for $g$, but the Fourier series of $g$ won't actually be equal to $g$. (The condition $g(0) = 0$ means that $\tilde{g}$ is continuous, and the condition $g(\pi) = 0$ means that $\tilde{g}(-\pi) = \tilde{g}(\pi)$, so $\tilde{g}$ is equal to its Fourier series.)

**Example 2.** Let $g(x)$ be the function on $[0,\pi]$ with the following graph.




That is,

$$g(x) = \begin{cases} x & \text{if } 0 \le x \le \frac{\pi}{2} \\ \pi - x & \text{if } \frac{\pi}{2} \le x \le \pi \end{cases}$$

In particular, $g(0) = g(\pi) = 0$. By Fact 1, we can write $g(x) = \sum_{n=1}^{\infty} a_n \sin(nx)$ where

$$\begin{aligned} a_n &= \frac{2}{\pi} \int_0^{\pi} g(x) \sin(nx)\, dx \\ &= \frac{2}{\pi} \left( \int_0^{\pi/2} x \sin(nx)\, dx + \int_{\pi/2}^{\pi} (\pi - x) \sin(nx)\, dx \right) \\ &= \frac{2}{\pi} \left( \left[ -\frac{x\cos(nx)}{n} + \frac{\sin(nx)}{n^2} \right]_0^{\pi/2} + \left[ \frac{(x-\pi)\cos(nx)}{n} - \frac{\sin(nx)}{n^2} \right]_{\pi/2}^{\pi} \right) \\ &= \frac{2}{\pi} \left( \frac{2\sin\left(\frac{n\pi}{2}\right)}{n^2} - \frac{\sin(n\pi)}{n^2} \right) \\ &= \frac{4\sin\left(\frac{n\pi}{2}\right)}{\pi n^2} \text{ because } \sin(n\pi) = 0 \text{ for all } n \end{aligned}$$

Therefore,

$$g(x) = \sum_{n=1}^{\infty} \frac{4\sin\left(\frac{n\pi}{2}\right)}{\pi n^2} \sin(nx).$$

(Again, remember this statement is only true because $g(0) = g(\pi) = 0$.) ❖

## Back to the Heat Equation

Now, we're ready to solve the heat equation. Remember that we were given a function representing the temperature at time 0; let's call this function $g(x)$. We were looking for a function $f(t,x)$ such that $\frac{\partial f}{\partial t} = \mu \frac{\partial^2 f}{\partial x^2}$, $f(t,0) = f(t,\pi) = 0$, and $f(0,x) = g(x)$.

At any fixed time $t$, $f(t,x)$ is just a function of $x$ on $[0,\pi]$ satisfying $f(t,0) = f(t,\pi) = 0$. Therefore, by Fact 1, we can write

$$f(t,x) = \sum_{n=1}^{\infty} a_n(t) \sin(nx) \tag{1}$$

for some $a_n(t)$. Then,

$$\begin{aligned}
\frac{\partial f}{\partial t}(t,x) &= \sum_{n=1}^{\infty} a_n'(t) \sin(nx) \\
\frac{\partial f}{\partial x}(t,x) &= \sum_{n=1}^{\infty} n a_n(t) \cos(nx) \\
\frac{\partial^2 f}{\partial x^2}(t,x) &= \sum_{n=1}^{\infty} (-n^2) a_n(t) \sin(nx)
\end{aligned}$$

So, the equation $\frac{\partial f}{\partial t} = \mu \frac{\partial^2}{\partial x^2}$ can be rewritten as

$$\sum_{n=1}^{\infty} a_n'(t) \sin(nx) = \sum_{n=1}^{\infty} (-\mu n^2) a_n(t) \sin(nx).$$

Equating the coefficients of $\sin(nx)$ on each side, we have $a_n'(t) = -\mu n^2 a_n(t)$. This is a first-order linear differential equations whose solutions are $a_n(t) = c_n e^{-\mu n^2 t}$. Plugging this into (1), we obtain $f(t,x) = \sum_{n=1}^{\infty} c_n e^{-\mu n^2 t} \sin(nx)$. Plugging $t = 0$ into this equation gives $f(0,x) = \sum_{n=1}^{\infty} c_n \sin(nx)$. Thus, the constants $c_n$ are simply the Fourier coefficients of the initial condition $f(0,x) = g(x)$. To summarize:

**Fact 3.** *Consider the heat equation with initial conditions*

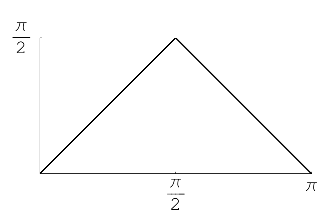
$$\begin{cases} \frac{\partial f}{\partial t} = \mu \frac{\partial^2 f}{\partial x^2} \\ f(t,0) = f(t,\pi) = 0 \text{ for all } t \\ f(0,x) = g(x) \end{cases}$$

*By Fact 1, we can write $g(x) = \sum_{n=1}^{\infty} c_n \sin(nx)$. Then, the solution of the heat equation is*

$$f(t,x) = \sum_{n=1}^{\infty} c_n e^{-\mu n^2 t} \sin(nx).$$

Note: We've glossed over a lot of details — it actually takes quite a lot of math to make the above argument work!

**Example 4.** Let $f(t,x)$ be the temperature of a bar at time $t$ and point $x$. Suppose the bar is initially heated so that its temperature distribution looks like this:



That is, the initial temperature is given by

$$g(x) = \begin{cases} x & \text{if } 0 \le x \le \frac{\pi}{2} \\ \pi - x & \text{if } \frac{\pi}{2} \le x \le \pi \end{cases}$$

The ends of the bar are kept at temperature 0 at all times, so $f(t,0) = f(t,\pi) = 0$ for all $t$.

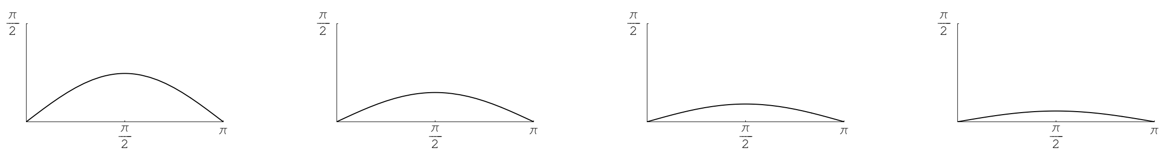
To find a formula for $f(t,x)$, we need to first write the Fourier series for $g(x)$. We did this in Example 2 and found that

$$g(x) = \sum_{n=1}^{\infty} \frac{4\sin\left(\frac{n\pi}{2}\right)}{\pi n^2} \sin(nx).$$

Therefore, the solution of the heat equation is

$$f(t,x) = \sum_{n=1}^{\infty} \frac{4\sin\left(\frac{n\pi}{2}\right)}{\pi n^2} e^{-\mu n^2 t} \sin(nx).$$

Here are graphs of $f(t,x)$ for $t = 1,2,3,4$ (using $\mu = 0.5$).



In particular, it is clear from our formula that $f(t,x) \to 0$ as $t \to \infty$. ❖

## Connection with Continuous Dynamical Systems

Earlier in the course, we looked at continuous dynamical systems of the form $\frac{d\vec{x}}{dt} = A\vec{x}$. As we saw, the solutions of these dynamical systems could be written in terms of the eigenvectors of $A$: if $\vec{v}_1, \ldots, \vec{v}_m$ was an eigenbasis for $A$ and $\lambda_1, \ldots, \lambda_m$ were the corresponding eigenvalues, we found that the solutions of $\frac{d\vec{x}}{dt} = A\vec{x}$ were $\vec{x}(t) = c_1 e^{\lambda_1 t}\vec{v}_1 + \cdots + c_m e^{\lambda_m t}\vec{v}_m$. In particular, $\vec{x}(0) = c_1\vec{v}_m + \cdots + c_m\vec{v}_m$. That is, the constants $c_i$ came from writing the initial condition $\vec{x}(0)$ as a linear combination of the eigenvectors.

When we solve the heat equation $\frac{\partial f}{\partial t} = \mu \frac{\partial^2 f}{\partial x^2}$ using Fourier series, we are really doing something very similar. If we let $D$ be the linear transformation $D(f) = \frac{\partial f}{\partial x}$, then the heat equation can be written as $\frac{\partial f}{\partial t} = \mu D^2 f$. Thus, $\mu D^2$ now plays the role of the matrix $A$, and we expect the eigenfunctions of $\mu D^2$ to figure into the solution. In fact, this is exactly what happens — the functions $\sin(nx)$ are eigenfunctions of $\mu D^2$ because $(\mu D^2)[\sin(nx)] = \mu D[n\cos(nx)] = -\mu n^2 \sin(nx)$. That is, $\sin(nx)$ is an eigenfunction of $\mu D^2$ with eigenvalue $-\mu n^2$. Now, compare $\frac{d\vec{x}}{dt} = A\vec{x}$ to $\frac{\partial f}{\partial t} = \mu \frac{\partial^2 f}{\partial x^2}$:

| | | |
|---|---|---|
| differential equation | $\frac{d\vec{x}}{dt} = A\vec{x}$ | $\frac{\partial f}{\partial t} = \mu D^2 f$ |
| linear transformation (LT) | $A$ | $\mu D^2$ |
| eigenvectors / eigenfunctions of the LT | $\vec{v}_i$ | $\sin(nx)$ |
| corresponding eigenvalues | $\lambda_i$ | $-\mu n^2$ |
| solutions of the differential equation | $\vec{x}(t) = \sum_{i=1}^{m} c_i e^{\lambda_i t}\vec{v}_i$ | $f(t,x) = \sum_{n=1}^{\infty} c_n e^{-\mu n^2 t}\sin(nx)$ |
| initial condition | $\vec{x}(0) = \sum_{i=1}^{m} c_i \vec{v}_i$ | $f(0,x) = \sum_{n=1}^{\infty} c_n \sin(nx)$ |

As you can see, using Fourier series to solve partial differential equations is really exactly analogous to using eigenvectors to solve continuous dynamical systems!

## The Wave Equation

Now, we'll look at a different partial differential equation. Suppose we have a string of length $\pi$ whose ends are fixed but whose middle can move around (imagine a guitar string, for example). Let $f(t,x)$ be the height of point $x$ of the string at time $t$. Then, $f$ satisfies a differential equation $\frac{\partial^2 f}{\partial t^2} = c^2 \frac{\partial^2 f}{\partial x^2}$, called the wave equation. Just as before, we can solve this partial differential equation by writing $f$ as a Fourier series $f(t,x) = \sum_{n=1}^{\infty} r_n(t)\sin(nx)$. Differentiating,

$$\begin{aligned}
\frac{\partial f}{\partial t}(t,x) &= \sum_{n=1}^{\infty} r_n'(t)\sin(nx) \\
\frac{\partial^2 f}{\partial t^2}(t,x) &= \sum_{n=1}^{\infty} r_n''(t)\sin(nx) \\
\frac{\partial f}{\partial x}(t,x) &= \sum_{n=1}^{\infty} n r_n(t)\cos(nx) \\
\frac{\partial^2 f}{\partial x^2}(t,x) &= \sum_{n=1}^{\infty} (-n^2) r_n(t)\sin(nx)
\end{aligned}$$

Equating the Fourier coefficients of $\frac{\partial^2 f}{\partial t^2}$ and $c^2 \frac{\partial^2 f}{\partial x^2}$, we have $r_n''(t) = -n^2 c^2 r_n(t)$. This is a second-order linear differential equation, and its solutions are $r_n(t) = a_n \cos(nct) + b_n \sin(nct)$. Thus,

$$f(t,x) = \sum_{n=1}^{\infty} [a_n \cos(nct) + b_n \sin(nct)]\sin(nx) \tag{2}$$

for some constants $a_n, b_n$. Since the wave equation is a second-order equation in $t$, we need two initial conditions to get a unique solution. Usually, the initial conditions given are the initial position of the string, $f(0,x) = g(x)$, and the initial velocity of the string, $\frac{\partial f}{\partial t}(0,x) = h(x)$, which must satisfy $h(0) = h(\pi) = 0$. Plugging $t = 0$ into (2) gives

$$g(x) = f(0,x) = \sum_{n=1}^{\infty} a_n \sin(nx).$$

On the other hand, if we differentiate (2) and plug in $t = 0$, we get

$$h(x) = \frac{\partial f}{\partial t}(0,x) = \sum_{n=1}^{\infty} ncb_n \sin(nx).$$

Therefore, the $a_n$ are the Fourier coefficients of $g(x)$ and the $ncb_n$ are the Fourier coefficients of $h(x)$. To summarize:

**Fact 5.** *Consider the wave equation with initial conditions*

$$\begin{cases}
\frac{\partial^2 f}{\partial t^2} = c^2 \frac{\partial^2 f}{\partial x^2} \\
f(t,0) = f(t,\pi) = 0 \text{ for all } t \\
f(0,x) = g(x) \\
\frac{\partial f}{\partial t}(0,x) = h(x) \text{ where } h(0) = h(\pi) = 0
\end{cases}$$

*By Fact 1, we can write*

$$g(x) = \sum_{n=1}^{\infty} a_n \sin(nx), h(x) = \sum_{n=1}^{\infty} b_n \sin(nx).$$

*Then, the solution of the wave equation is*

$$f(t,x) = \sum_{n=1}^{\infty} \left[a_n \cos(nct) + \frac{b_n}{nc}\sin(nct)\right]\sin(nx).$$

In particular, we see that the constant $c$ determines the period of oscillation of the string.

**Example 6.** Let $f(t,x)$ be the height of a guitar string at time $t$ and point $x$. Suppose the string initially has height $g(x) = \sin x + \sin 3x - 2\sin 4x$ and velocity $h(x) = -\sin x + \sin 3x$.

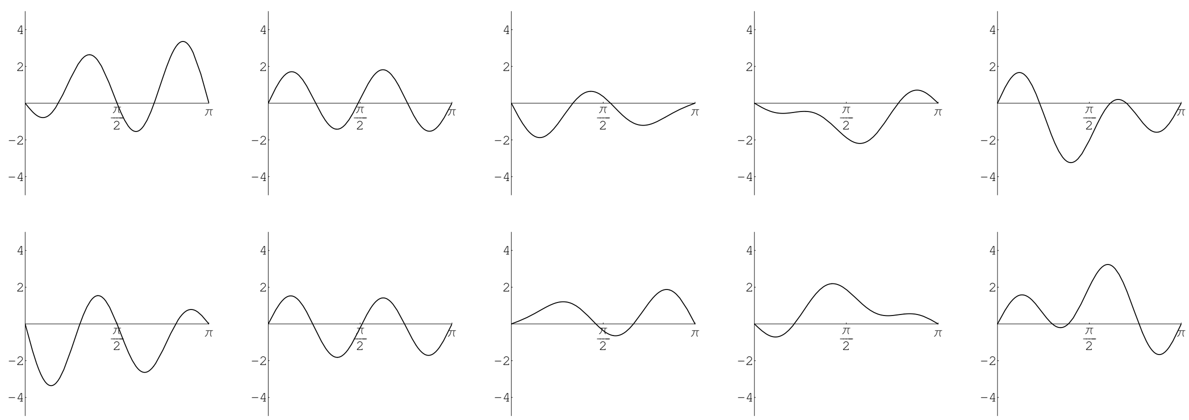
Notice that $g(x)$ and $h(x)$ are already written as Fourier series:

- $g(x) = \sum_{n=1}^{\infty} a_n \sin(nx)$ where $a_1 = 1$, $a_2 = 0$, $a_3 = 1$, $a_4 = -2$, and $a_n = 0$ for $n \geq 5$.
- $h(x) = \sum_{n=1}^{\infty} b_n \sin(nx)$ where $b_1 = -1$, $b_2 = 0$, $b_3 = 1$, and $b_n = 0$ for $n \geq 4$.

Therefore, the height of the string at any time is given by

$$f(t,x) = \left[\cos(ct) - \frac{1}{c}\sin(ct)\right]\sin(x) + \left[\cos(3ct) + \frac{1}{3c}\sin(3ct)\right]\sin(3x) + [-2\cos(4ct)]\sin(4x).$$

Here are graphs of $f(t,x)$ for $t = 0, \frac{\pi}{5}, \frac{2\pi}{5}, \ldots, \frac{9\pi}{5}$ (using $c = 1$).




❖