2

Here and There: Points and Vectors in 2D

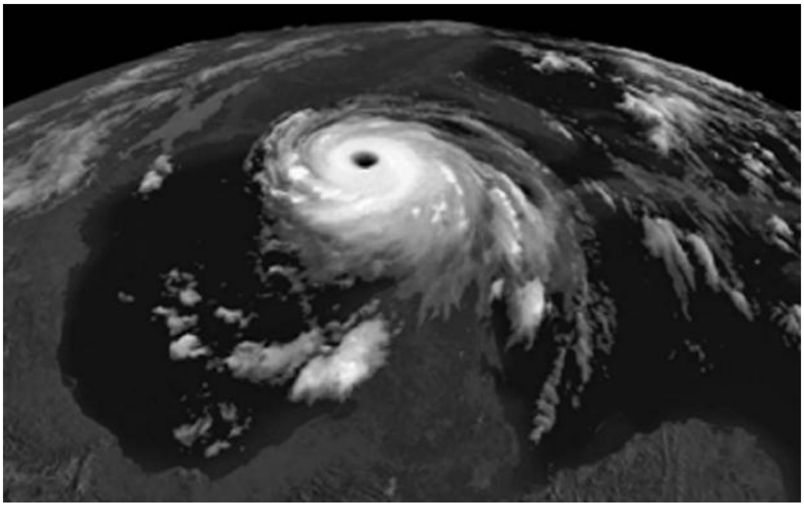
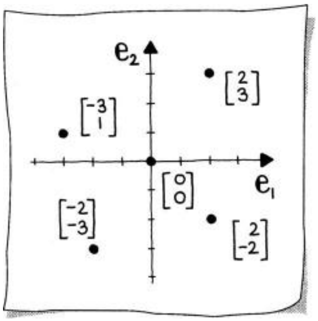


Figure 2.1.
Hurricane Katrina: the hurricane is shown here approaching south Louisiana. (Image courtesy of NOAA, katrina.noaa.gov.)

In 2005 Hurricane Katrina caused flooding and deaths as it made its way from the Bahamas to south Florida as a category 1 hurricane. Over the warm waters of the Gulf, it grew into a category 5 hurricane, and even though at landfall in southeast Louisiana it had weakened to a category 3 hurricane, the storm surges and destruction it created rates it as the most expensive hurricane to date, causing more than $45 billion of damage. Sadly it was also one of the deadliest, particularly for residents of New Orleans. In the hurricane image (Figure 2.1), air is moving rapidly, spiraling in a counterclockwise fashion. What isn't so clear from this image is that the air moves faster as it approaches the eye of the hurricane. This air movement is best described by points and vectors: at any location (point), air moves in a certain direction and with a certain speed (velocity vector).

This hurricane image is a good example of how helpful 2D geometry can be in a 3D world. Of course a hurricane is a 3D phenomenon; however, by analyzing 2D slices, or cross sections, we can develop a very informative analysis. Many other applications call for 2D geometry only. The purpose of this chapter is to define the two most fundamental tools we need to work in a 2D world: points and vectors.



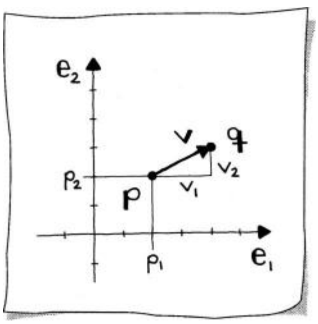
Sketch 2.1.
Points and their coordinates.

2.1 Points and Vectors

The most basic geometric entity is the *point*. A point is a reference to a *location*. Sketch 2.1 illustrates examples of points. In the text, boldface lowercase letters represent points, e.g.,

$$\mathbf{p} = \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}. \tag{2.1}$$

The location of $\mathbf{p}$ is p_1-units along the $\mathbf{e}_1$-axis and p_2-units along the $\mathbf{e}_2$-axis. Thus a point's *coordinates*, p_1 and p_2, are dependent upon the location of the coordinate origin. We use the boldface notation so there is a noticeable difference between a one-dimensional (1D) number, or *scalar* p. To clearly identify $\mathbf{p}$ as a point, the notation $\mathbf{p} \in \mathbb{E}^2$ is used. This means that a 2D point "lives" in 2D Euclidean space $\mathbb{E}^2$.



Sketch 2.2.
Two points and a vector.

Now let's move away from our reference point. Following Sketch 2.2, suppose the reference point is $\mathbf{p}$, and when moving along a straight path, our target point is $\mathbf{q}$. The directions from $\mathbf{p}$ would be to follow the *vector* $\mathbf{v}$. Our notation for a vector is the same as for a point: boldface lowercase letters. To get to $\mathbf{q}$ we say,

$$\mathbf{q} = \mathbf{p} + \mathbf{v}. \tag{2.2}$$

To calculate this, add each component separately; that is,

$$\begin{bmatrix} q_1 \\ q_2 \end{bmatrix} = \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} p_1 + v_1 \\ p_2 + v_2 \end{bmatrix}.$$

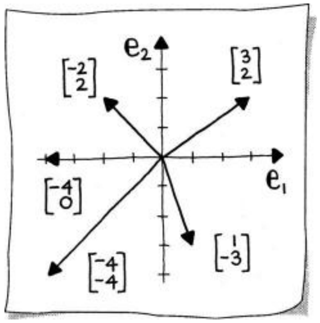
For example, in Sketch 2.2, we have

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

The *components* of $\mathbf{v}$, v_1 and v_2, indicate how many units to move along the $\mathbf{e}_1$- and $\mathbf{e}_2$-axis, respectively. This means that $\mathbf{v}$ can be defined as

$$\mathbf{v} = \mathbf{q} - \mathbf{p}. \tag{2.3}$$

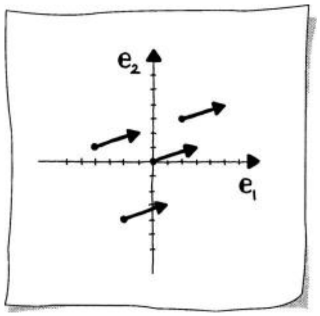
This defines a vector as a difference of two points, which describes a *direction and a distance*, or a *displacement*. Examples of vectors are illustrated in Sketch 2.3.



Sketch 2.3.
Vectors and their components.

How to determine a vector's length is covered in Section 2.4. Above we described this length as a distance. Alternatively, this length can be described as speed: then we have a *velocity vector*.[1] Yet another interpretation is that the length represents acceleration: then we have a *force vector*.

A vector has a *tail* and a *head*. As in Sketch 2.2, the tail is typically displayed positioned at a point, or *bound to a point* in order to indicate the geometric significance of the vector. However, unlike a point, a vector does *not* define a position. Two vectors are equal if they have the same component values, just as points are equal if they have the same coordinate values. Thus, considering a vector as a difference of two points, there are any number of vectors with the same direction and length. See Sketch 2.4 for an illustration.



Sketch 2.4.
Instances of one vector.

A special vector worth mentioning is the *zero vector*,

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

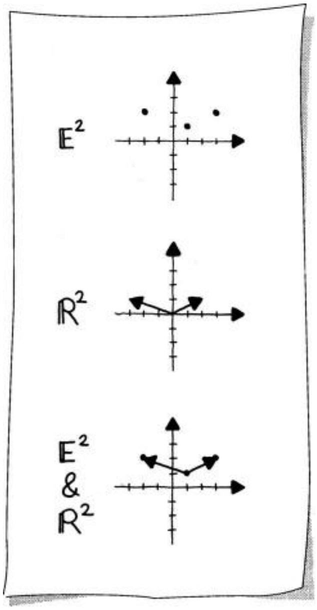
This vector has no direction or length. Other somewhat special vectors include

$$\mathbf{e}_1 = \begin{bmatrix}1\\0\end{bmatrix} \quad \text{and} \quad \mathbf{e}_2 = \begin{bmatrix}0\\1\end{bmatrix}.$$

In the sketches, these vectors are not always drawn true to length to prevent them from obscuring the main idea.

To clearly identify $\mathbf{v}$ as a vector, we write $\mathbf{v} \in \mathbb{R}^2$. This means that a 2D vector "lives" in a 2D *linear space* $\mathbb{R}^2$. (Other names for $\mathbb{R}^2$ are *real* or *vector spaces*.)

[1]This is what we'll use to continue the Hurricane Katrina example.



Sketch 2.5.
Euclidean and linear spaces illustrated separately and together.

2.2 What's the Difference?

When writing a point or a vector we use boldface lowercase letters; when programming we use the same data structure, e.g., arrays. This makes it appear that points and vectors can be treated in the same manner. Not so!

Points and vectors are different geometric entities. This is reiterated by saying they live in different spaces, $\mathbb{E}^2$ and $\mathbb{R}^2$. As shown in Sketch 2.5, for convenience and clarity elements of Euclidean and linear spaces are typically displayed together.

The primary reason for differentiating between points and vectors is to achieve geometric constructions that are *coordinate independent*. Such constructions are manipulations applied to geometric objects that produce the same result regardless of the location of the coordinate origin (for example, the midpoint of two points). This idea becomes clearer by analyzing some fundamental manipulations of points and vectors. In what follows, let's use $\mathbf{p}, \mathbf{q} \in \mathbb{E}^2$ and $\mathbf{v}, \mathbf{w} \in \mathbb{R}^2$.

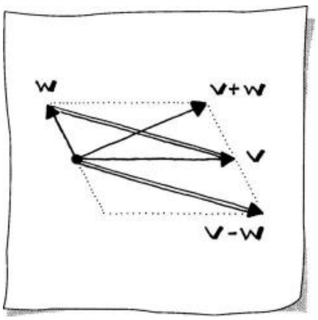
Coordinate Independent Operations:

- Subtracting a point from another $(\mathbf{p} - \mathbf{q})$ yields a vector, as depicted in Sketch 2.2 and Equation (2.3).

- Adding or subtracting two vectors yields another vector. See Sketch 2.6, which illustrates the *parallelogram rule:* the vectors $\mathbf{v} - \mathbf{w}$ and $\mathbf{v} + \mathbf{w}$ are the diagonals of the parallelogram defined by $\mathbf{v}$ and $\mathbf{w}$. This is a coordinate independent operation since vectors are defined as a difference of points.

- Multiplying by a scalar s is called *scaling*. Scaling a vector is a well-defined operation. The result $s\mathbf{v}$ adjusts the length by the scaling factor. The direction is unchanged if $s > 0$ and reversed for $s < 0$. If $s = 0$ then the result is the zero vector. Sketch 2.7 illustrates some examples of scaling a vector.

- Adding a vector to a point $(\mathbf{p} + \mathbf{v})$ yields another point, as in Sketch 2.2 and Equation (2.2).

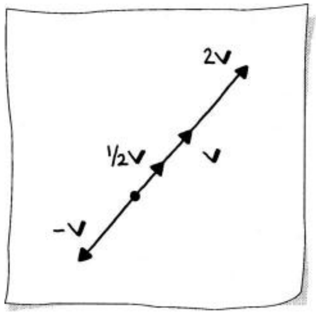


Sketch 2.6.
Parallelogram rule.

Any coordinate independent combination of two or more points and/or vectors can be grouped to fall into one or more of the items above. See the Exercises for examples.

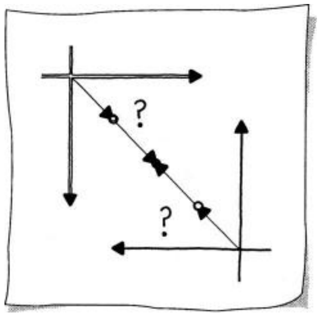
Coordinate Dependent Operations:

- Scaling a point ($s\mathbf{p}$) is not a well-defined operation because it is not coordinate independent. Sketch 2.8 illustrates that the result of scaling the solid black point by one-half with respect to two different coordinate systems results in two different points.
- Adding two points ($\mathbf{p}+\mathbf{q}$) is not a well-defined operation because it is not coordinate independent. As depicted in Sketch 2.9, the result of adding the two solid black points is dependent on the coordinate origin. (The parallelogram rule is used here to construct the results of the additions.)



Sketch 2.7.
Scaling a vector.

Some special combinations of points are allowed; they are defined in Section 2.5.



Sketch 2.8.
Scaling of points is ambiguous.

2.3 Vector Fields

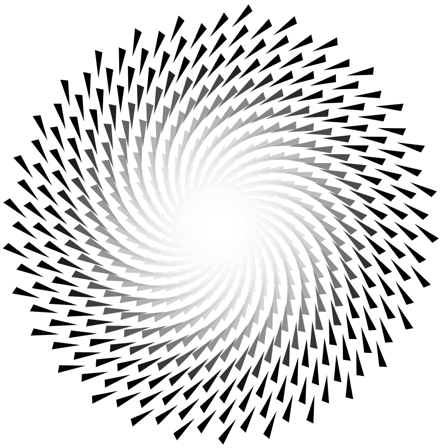
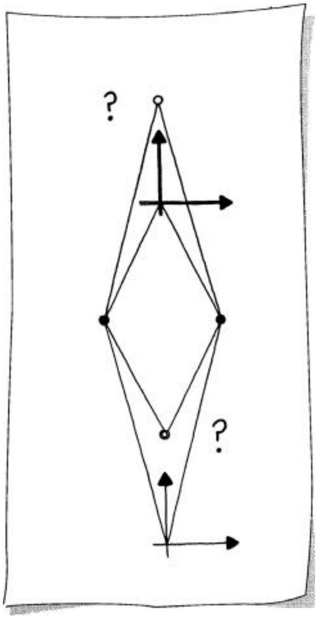


Figure 2.2.
Vector field: simulating hurricane air velocity. Lighter gray indicates greater velocity.

A good way to visualize the interplay between points and vectors is through the example of *vector fields*. In general, we speak of a vector field if every point in a given region is assigned a vector. We have already encountered an example of this in Figure 2.1: Hurricane Katrina! Recall that at each location (point) we could describe the air velocity (vector). Our previous image did not actually tell us anything about the air speed, although we could presume something about the direction. This is where a vector field is helpful. Shown in Figure 2.2 is a vector field simulating Hurricane Katrina. By plotting all the



Sketch 2.9.
Addition of points is ambiguous.

vectors the same length and using *gray scale* or varying shades of gray to indicate speed, the vector field can be more informative than the photograph. (Visualization of a vector field requires *discretizing* it: a finite number of point and vector pairs are selected from a continuous field or from sampled measurements.)

Other important applications of vector fields arise in the areas of automotive and aerospace design: before a car or an airplane is built, it undergoes extensive aerodynamic simulations. In these simulations, the vectors that characterize the flow around an object are computed from complex differential equations. In Figure 2.3 we have another example of a vector field.

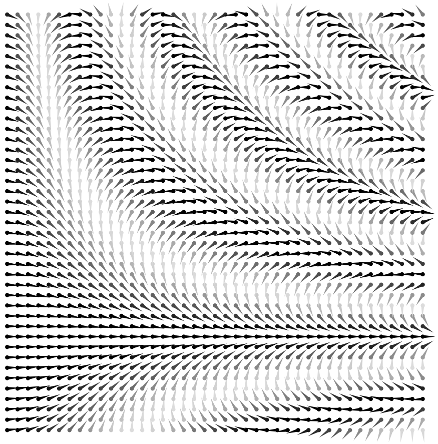
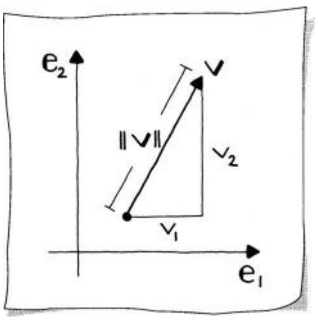


Figure 2.3.
Vector field: every sampled point has an associated vector. Lighter gray indicates greater vector length.



Sketch 2.10.
Length of a vector.

2.4 Length of a Vector

As mentioned in Section 2.1, the length of a vector can represent distance, velocity, or acceleration. We need a method for finding the length of a vector, or the *magnitude*. As illustrated in Sketch 2.10, a vector defines the displacement necessary (with respect to the $\mathbf{e}_1$- and $\mathbf{e}_2$-axis) to get from a point at the tail of the vector to a point at the head of the vector.

In Sketch 2.10 we have formed a right triangle. The square of the length of the hypotenuse of a right triangle is well known from the *Pythagorean theorem*. Denote the *length* of a vector $\mathbf{v}$ as $\|\mathbf{v}\|$. Then

$$\|\mathbf{v}\|^2 = v_1^2 + v_2^2.$$

Therefore, the magnitude of $\mathbf{v}$ is

$$\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2}. \tag{2.4}$$

This is also called the *Euclidean norm*. Notice that if we scale the vector by an amount k then

$$\|k\mathbf{v}\| = |k|\|\mathbf{v}\|. \tag{2.5}$$

A *normalized vector* $\mathbf{w}$ has *unit length*, that is

$$\|\mathbf{w}\| = 1.$$

Normalized vectors are also known as *unit vectors*. To *normalize* a vector simply means to scale a vector so that it has unit length. If $\mathbf{w}$ is to be our unit length version of $\mathbf{v}$ then

$$\mathbf{w} = \frac{\mathbf{v}}{\|\mathbf{v}\|}.$$

Each component of $\mathbf{v}$ is divided by the scalar value $\|\mathbf{v}\|$. This scalar value is always *nonnegative*, which means that its value is zero or greater. It can be zero! You must check the value before dividing to be sure it is greater than your *zero divide tolerance*. The zero divide tolerance is the absolute value of the smallest number by which you can divide confidently. (When we refer to checking that a value is greater than this number, it means to check the absolute value.)

In Figures 2.2 and 2.3, we display vectors of varying magnitudes. But instead of plotting them using different lengths, their magnitude is indicated by gray scales.

Example 2.1

Start with

$$\mathbf{v} = \begin{bmatrix} 5 \\ 0 \end{bmatrix}.$$

Applying (2.4), $\|\mathbf{v}\| = \sqrt{5^2 + 0^2} = 5$. Then the normalized version of $\mathbf{v}$ is defined as

$$\mathbf{w} = \begin{bmatrix} 5/5 \\ 0/5 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Clearly $\|\mathbf{w}\| = 1$, so this is a normalized vector. Since we have only scaled $\mathbf{v}$ by a positive amount, the direction of $\mathbf{w}$ is the same as $\mathbf{v}$.

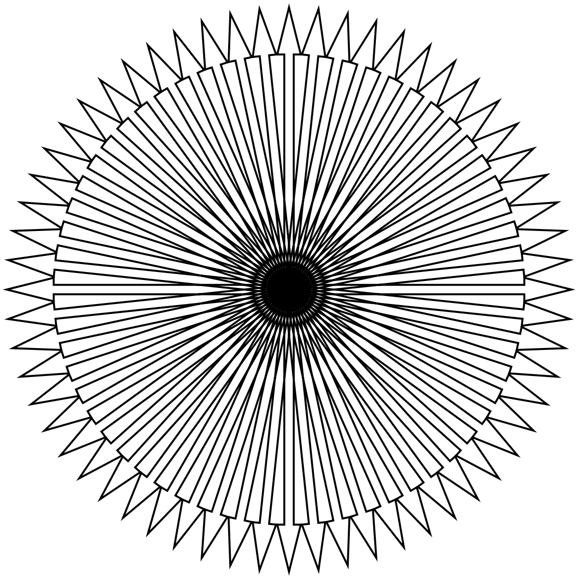
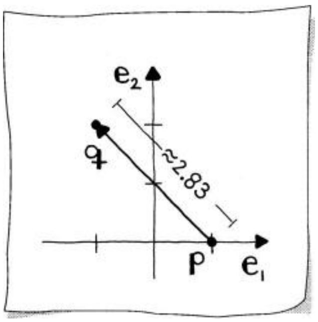


Figure 2.4.
Unit vectors: they define a circle.

There are infinitely many unit vectors. Imagine drawing them all, emanating from the origin. The figure that you will get is a circle of radius one! See Figure 2.4.

To find the *distance between two points* we simply form a vector defined by the two points, e.g., $\mathbf{v} = \mathbf{q} - \mathbf{p}$, and apply (2.4).



Sketch 2.11.
Distance between two points.

Example 2.2

Let

$$\mathbf{q} = \begin{bmatrix} -1 \\ 2 \end{bmatrix} \quad \text{and} \quad \mathbf{p} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Then

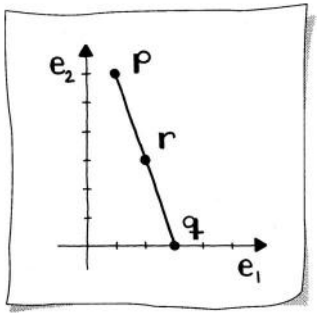
$$\mathbf{q} - \mathbf{p} = \begin{bmatrix} -2 \\ 2 \end{bmatrix}$$

and

$$\|\mathbf{q} - \mathbf{p}\| = \sqrt{(-2)^2 + 2^2} = \sqrt{8} \approx 2.83.$$

Sketch 2.11 illustrates this example.

2.5 Combining Points



Sketch 2.12.
The midpoint of two points.

Seemingly contrary to Section 2.2, there actually is a way to combine two points such that we get a (meaningful) third one. Take the example of the midpoint $\mathbf{r}$ of two points $\mathbf{p}$ and $\mathbf{q}$; more specifically, take

$$\mathbf{p} = \begin{bmatrix} 1 \\ 6 \end{bmatrix}, \quad \mathbf{r} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \quad \mathbf{q} = \begin{bmatrix} 3 \\ 0 \end{bmatrix},$$

as shown in Sketch 2.12.

Let's start with the known coordinate independent operation of adding a vector to a point. Define $\mathbf{r}$ by adding an appropriately scaled version of the vector $\mathbf{v} = \mathbf{q} - \mathbf{p}$ to the point $\mathbf{p}$:

$$\mathbf{r} = \mathbf{p} + \frac{1}{2}\mathbf{v}$$
$$\begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 6 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 2 \\ -6 \end{bmatrix}.$$

Expanding, this shows that $\mathbf{r}$ can also be defined as

$$\mathbf{r} = \frac{1}{2}\mathbf{p} + \frac{1}{2}\mathbf{q}$$
$$\begin{bmatrix} 2 \\ 3 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 \\ 6 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 3 \\ 0 \end{bmatrix}.$$

This is a legal expression for a combination of points.

There is nothing magical about the factor 1/2, however. Adding a (scaled) vector to a point is a well-defined, coordinate independent operation that yields another point. Any point of the form

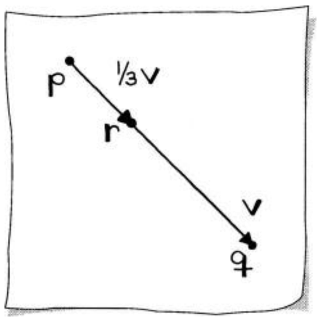
$$\mathbf{r} = \mathbf{p} + t\mathbf{v} \tag{2.6}$$

is on the line through $\mathbf{p}$ and $\mathbf{q}$. Again, we may rewrite this as

$$\mathbf{r} = \mathbf{p} + t(\mathbf{q} - \mathbf{p})$$

and then

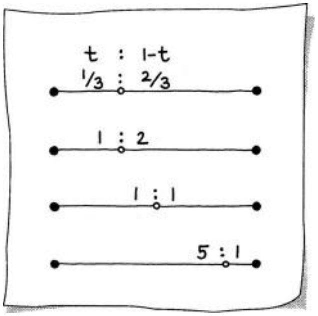
$$\mathbf{r} = (1 - t)\mathbf{p} + t\mathbf{q}. \tag{2.7}$$



Sketch 2.13.
Barycentric combinations: $t = 1/3$.

Sketch 2.13 gives an example with $t = 1/3$.

The scalar values $(1 - t)$ and t are *coefficients*. A weighted sum of points where the coefficients sum to one is called a *barycentric combination*. In this special case, where one point $\mathbf{r}$ is being expressed in terms of two others, $\mathbf{p}$ and $\mathbf{q}$, the coefficients $1 - t$ and t are called the *barycentric coordinates* of $\mathbf{r}$.



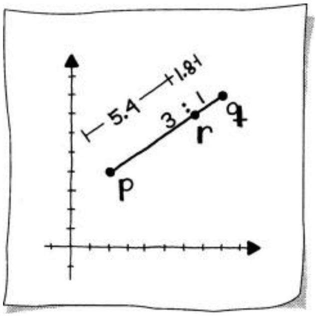
Sketch 2.14.
Examples of ratios.

A barycentric combination allows us to construct $\mathbf{r}$ anywhere on the line defined by $\mathbf{p}$ and $\mathbf{q}$. This is why (2.7) is also called linear interpolation. If we would like to restrict $\mathbf{r}$'s position to the *line segment* between $\mathbf{p}$ and $\mathbf{q}$, then we allow only *convex combinations*: t must satisfy $0 \le t \le 1$. To define points outside of the line segment between $\mathbf{p}$ and $\mathbf{q}$, we need values of $t < 0$ or $t > 1$.

The position of $\mathbf{r}$ is said to be in the *ratio* of $t : (1-t)$ or $t/(1-t)$. In physics, $\mathbf{r}$ is known as the *center of gravity* of two points $\mathbf{p}$ and $\mathbf{q}$ with weights $1-t$ and t, respectively. From a constructive approach, the ratio is formed from the quotient

$$\text{ratio} = \frac{||\mathbf{r}-\mathbf{p}||}{||\mathbf{q}-\mathbf{r}||}.$$

Some examples are illustrated in Sketch 2.14.



Sketch 2.15.
Barycentric coordinates in relation to lengths.

Example 2.3

Suppose we have three collinear points, $\mathbf{p}$, $\mathbf{q}$, and $\mathbf{r}$ as illustrated in Sketch 2.15. The points have the following locations.

$$\mathbf{p} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}, \quad \mathbf{r} = \begin{bmatrix} 6.5 \\ 7 \end{bmatrix}, \quad \mathbf{q} = \begin{bmatrix} 8 \\ 8 \end{bmatrix}.$$

What are the barycentric coordinates of $\mathbf{r}$ with respect to $\mathbf{p}$ and $\mathbf{q}$?

To answer this, recall the relationship between the ratio and the barycentric coordinates. The barycentric coordinates t and $(1-t)$ define $\mathbf{r}$ as

$$\begin{bmatrix} 6.5 \\ 7 \end{bmatrix} = (1-t)\begin{bmatrix} 2 \\ 4 \end{bmatrix} + t\begin{bmatrix} 8 \\ 8 \end{bmatrix}.$$

The ratio indicates the location of $\mathbf{r}$ relative to $\mathbf{p}$ and $\mathbf{q}$ in terms of relative distances. Suppose the ratio is $s_1 : s_2$. If we scale s_1 and s_2 such that they sum to one, then s_1 and s_2 are the barycentric coordinates t and $(1-t)$, respectively. By calculating the distances between points:

$$\begin{aligned} l_1 &= \|\mathbf{r}-\mathbf{p}\| \approx 5.4, \\ l_2 &= \|\mathbf{q}-\mathbf{r}\| \approx 1.8, \\ l_3 &= l_1 + l_2 \approx 7.2, \end{aligned}$$

we find that

$$\begin{aligned} t &= l_1/l_3 = 0.75 \quad \text{and} \\ (1-t) &= l_2/l_3 = 0.25. \end{aligned}$$

These are the barycentric coordinates. Let's verify this:

$$\begin{bmatrix} 6.5 \\ 7 \end{bmatrix} = 0.25 \times \begin{bmatrix} 2 \\ 4 \end{bmatrix} + 0.75 \times \begin{bmatrix} 8 \\ 8 \end{bmatrix}.$$

The barycentric coordinate t is also called a *parameter*. (See Section 3.2 for more details.) This parameter is defined by the quotient

$$t = \frac{||\mathbf{r} - \mathbf{p}||}{||\mathbf{q} - \mathbf{p}||}.$$

We have seen how useful this quotient can be in Section 1.1 for the construction of a point in the global system that corresponded to a point with parameter t in the local system.

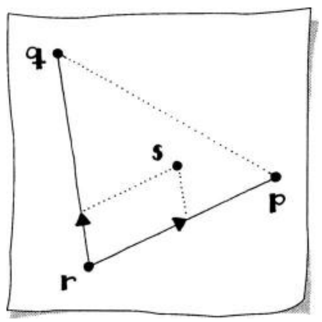
We can create barycentric combinations with *more than two points*. Let's look at three points $\mathbf{p}$, $\mathbf{q}$, and $\mathbf{r}$, which are not collinear. Any point $\mathbf{s}$ can be formed from

$$\mathbf{s} = \mathbf{r} + t_1(\mathbf{p} - \mathbf{r}) + t_2(\mathbf{q} - \mathbf{r}).$$

This is a coordinate independent operation of point + vector + vector. Expanding and regrouping, we can also define $\mathbf{s}$ as

$$\begin{aligned} \mathbf{s} &= t_1\mathbf{p} + t_2\mathbf{q} + (1 - t_1 - t_2)\mathbf{r} \\ &= t_1\mathbf{p} + t_2\mathbf{q} + t_3\mathbf{r}. \end{aligned} \tag{2.8}$$

Thus, the point $\mathbf{s}$ is defined by a barycentric combination with coefficients t_1, t_2, and $t_3 = 1 - t_1 - t_2$ with respect to $\mathbf{p}$, $\mathbf{q}$, and $\mathbf{r}$, respectively. This is another special case where the barycentric combination coefficients correspond to *barycentric coordinates*. Sketch 2.16 illustrates this. We will encounter barycentric coordinates in more detail in Chapter 17.



Sketch 2.16.

A barycentric combination of three points.

We can also *combine points* so that the result is a vector. For this, we need the coefficients to sum to zero. We encountered a simple case of this in (2.3). Suppose we have the equation

$$\mathbf{e} = \mathbf{r} - 2\mathbf{p} + \mathbf{q}, \qquad \mathbf{r}, \mathbf{p}, \mathbf{q} \in \mathbb{E}^2.$$

Does $\mathbf{e}$ have a geometric meaning? Looking at the sum of the coefficients, $1 - 2 + 1 = 0$, we would conclude by the rule above that $\mathbf{e}$ is a vector. How to see this? By rewriting the equation as

$$\mathbf{e} = (\mathbf{r} - \mathbf{p}) + (\mathbf{q} - \mathbf{p}),$$

it is clear that $\mathbf{e}$ is a vector formed from (vector + vector).

2.6 Independence

Two vectors $\mathbf{v}$ and $\mathbf{w}$ describe a parallelogram, as shown in Sketch 2.6. It may happen that this parallelogram has zero area; then the two vectors are parallel. In this case, we have a relationship of the form $\mathbf{v} = c\mathbf{w}$. If two vectors are parallel, then we call them *linearly dependent*. Otherwise, we say that they are *linearly independent*.

Two linearly independent vectors may be used to write any other vector $\mathbf{u}$ as a *linear combination*:

$$\mathbf{u} = r\mathbf{v} + s\mathbf{w}.$$

How to find r and s is described in Chapter 5. Two linearly independent vectors in 2D are also called a *basis* for $\mathbb{R}^2$. If $\mathbf{v}$ and $\mathbf{w}$ are linearly dependent, then you cannot write all vectors as a linear combination of them, as the following example shows.

Example 2.4

Let

$$\mathbf{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \text{and} \quad \mathbf{w} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}.$$

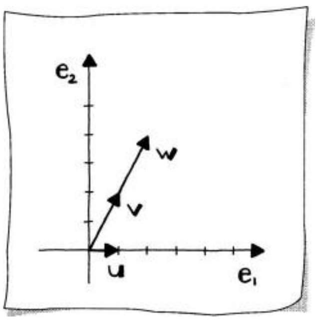
If we tried to write the vector

$$\mathbf{u} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

as $\mathbf{u} = r\mathbf{v} + s\mathbf{w}$, then this would lead to

$$1 = r + 2s, \tag{2.9}$$
$$0 = 2r + 4s. \tag{2.10}$$

If we multiply the first equation by a factor of 2, the two right-hand sides will be equal. Equating the new left-hand sides now results in the expression $2 = 0$. This shows that $\mathbf{u}$ cannot be written as a linear combination of $\mathbf{v}$ and $\mathbf{w}$. (See Sketch 2.17.)



Sketch 2.17.
Dependent vectors.

2.7 Dot Product

Given two vectors $\mathbf{v}$ and $\mathbf{w}$, we might ask:

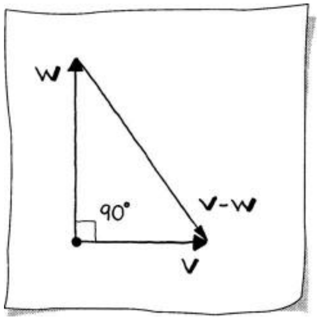
- Are they the *same* vector?

- Are they *perpendicular* to each other?
- What *angle* do they form?

The *dot product* is the tool to resolve these questions. Assume that $\mathbf{v}$ and $\mathbf{w}$ are not the zero vector.

To motivate the dot product, let's start with the Pythagorean theorem and Sketch 2.18. There, we see two perpendicular vectors $\mathbf{v}$ and $\mathbf{w}$; we conclude

$$\|\mathbf{v} - \mathbf{w}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2. \tag{2.11}$$



Sketch 2.18.
Perpendicular vectors.

Writing the components in (2.11) explicitly

$$(v_1 - w_1)^2 + (v_2 - w_2)^2 = (v_1^2 + v_2^2) + (w_1^2 + w_2^2),$$

and then expanding, bringing all terms to the left-hand side of the equation yields

$$(v_1^2 - 2v_1w_1 + w_1^2) + (v_2^2 - 2v_2w_2 + w_2^2) - (v_1^2 + v_2^2) - (w_1^2 + w_2^2) = 0,$$

which reduces to

$$v_1w_1 + v_2w_2 = 0. \tag{2.12}$$

We find that perpendicular vectors have the property that the sum of the products of their components is zero. The short-hand vector notation for (2.12) is

$$\mathbf{v} \cdot \mathbf{w} = 0. \tag{2.13}$$

This result has an immediate application: a vector $\mathbf{w}$ perpendicular to a given vector $\mathbf{v}$ can be formed as

$$\mathbf{w} = \begin{bmatrix} -v_2 \\ v_1 \end{bmatrix}$$

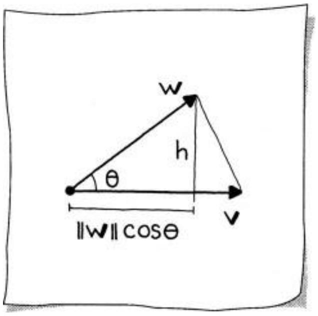
(switching components and negating the sign of one). Then $\mathbf{v} \cdot \mathbf{w}$ becomes $v_1(-v_2) + v_2v_1 = 0$.

If we take two arbitrary vectors $\mathbf{v}$ and $\mathbf{w}$, then $\mathbf{v} \cdot \mathbf{w}$ will in general not be zero. But we can compute it anyway, and define

$$s = \mathbf{v} \cdot \mathbf{w} = v_1w_1 + v_2w_2 \tag{2.14}$$

to be the *dot product* of $\mathbf{v}$ and $\mathbf{w}$. Notice that the dot product returns a scalar s, which is why it is also called a *scalar product*. (Mathematicians have yet another name for the dot product—an *inner product*. See Section 14.3 for more on these.) From (2.14) it is clear that

$$\mathbf{v} \cdot \mathbf{w} = \mathbf{w} \cdot \mathbf{v}.$$



Sketch 2.19.
Geometry of the dot product.

This is called the *symmetry property*. Other properties of the dot product are given in the Exercises.

In order to understand the geometric meaning of the dot product of two vectors, let's construct a triangle from two vectors $\mathbf{v}$ and $\mathbf{w}$ as illustrated in Sketch 2.19.

From trigonometry, we know that the height h of the triangle can be expressed as

$$h = \|\mathbf{w}\| \sin(\theta).$$

Squaring both sides results in

$$h^2 = \|\mathbf{w}\|^2 \sin^2(\theta).$$

Using the identity

$$\sin^2(\theta) + \cos^2(\theta) = 1,$$

we have

$$h^2 = \|\mathbf{w}\|^2(1 - \cos^2(\theta)). \tag{2.15}$$

We can also express the height h with respect to the other right triangle in Sketch 2.19 and by using the Pythagorean theorem:

$$h^2 = \|\mathbf{v} - \mathbf{w}\|^2 - (\|\mathbf{v}\| - \|\mathbf{w}\| \cos\theta)^2. \tag{2.16}$$

Equating (2.15) and (2.16) and simplifying, we have the expression,

$$\|\mathbf{v} - \mathbf{w}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 - 2\|\mathbf{v}\|\|\mathbf{w}\| \cos\theta. \tag{2.17}$$

We have just proved the *Law of Cosines*, which generalizes the Pythagorean theorem by correcting it for triangles with an opposing angle different from 90°.

We can formulate another expression for $\|\mathbf{v} - \mathbf{w}\|^2$ by explicitly writing out

$$\begin{aligned} \|\mathbf{v} - \mathbf{w}\|^2 &= (\mathbf{v} - \mathbf{w}) \cdot (\mathbf{v} - \mathbf{w}) \\ &= \|\mathbf{v}\|^2 - 2\mathbf{v} \cdot \mathbf{w} + \|\mathbf{w}\|^2. \end{aligned} \tag{2.18}$$

By equating (2.17) and (2.18) we find that

$$\mathbf{v} \cdot \mathbf{w} = \|\mathbf{v}\|\|\mathbf{w}\| \cos\theta. \tag{2.19}$$

Here is another expression for the *dot product*—it is a very useful one! Rearranging (2.19), the cosine of the angle between the two vectors can be determined as

$$\cos\theta = \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{v}\|\|\mathbf{w}\|}. \tag{2.20}$$

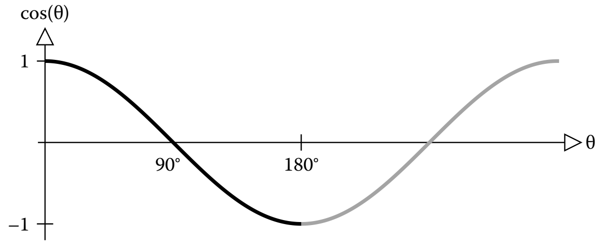


Figure 2.5.
Cosine function: its values at $\theta = 0°$, $\theta = 90°$, and $\theta = 180°$ are important to remember.

By examining a plot of the cosine function in Figure 2.5, some sense can be made of (2.20).

First we consider the special case of perpendicular vectors. Recall the dot product was zero, which makes $\cos(90°) = 0$, just as it should be.

If $\mathbf{v}$ has the same (or opposite) direction as $\mathbf{w}$, that is $\mathbf{v} = k\mathbf{w}$, then (2.20) becomes

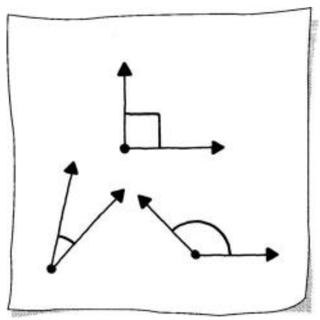
$$\cos\theta = \frac{k\mathbf{w} \cdot \mathbf{w}}{\|k\mathbf{w}\|\|\mathbf{w}\|}.$$

Using (2.5), we have

$$\cos\theta = \frac{k\|\mathbf{w}\|^2}{|k|\|\mathbf{w}\|\|\mathbf{w}\|} = \pm 1.$$

Again, examining Figure 2.5, we see this corresponds to either $\theta = 0°$ or $\theta = 180°$, for vectors of the same or opposite direction, respectively.

The cosine values from (2.20) range between ± 1; this corresponds to angles between 0° and 180° (or 0 and π radians). Thus, the smaller angle between the two vectors is measured. This is clear from the derivation: the angle θ enclosed by completing the triangle defined by the two vectors must be less than 180°. Three types of angles can be formed:



Sketch 2.20.
Three types of angles.

- *right*: $\cos(\theta) = 0 \rightarrow \mathbf{v} \cdot \mathbf{w} = 0$;
- *acute*: $\cos(\theta) > 0 \rightarrow \mathbf{v} \cdot \mathbf{w} > 0$;
- *obtuse*: $\cos(\theta) < 0 \rightarrow \mathbf{v} \cdot \mathbf{w} < 0$.

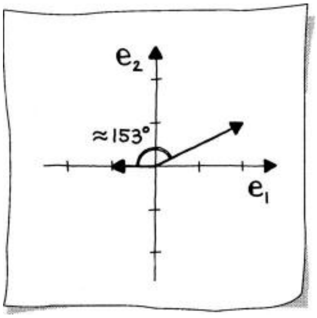
These are illustrated in counterclockwise order from twelve o'clock in Sketch 2.20.

If the actual angle θ needs to be calculated, then the arccosine function has to be invoked: let

$$s = \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{v}\|\|\mathbf{w}\|}$$

then $\theta = \mathrm{acos}(s)$ where acos is short for arccosine. One word of warning: in some math libraries, if $s > 1$ or $s < -1$ then an error occurs and a nonusable result (NaN—Not a Number) is returned.

Thus, if s is calculated, it is best to check that its value is within the appropriate range. It is not uncommon that an intended value of $s = 1.0$ is actually something like $s = 1.0000001$ due to *round-off*. Thus, the arccosine function should be used with caution. In many instances, as in comparing angles, the cosine of the angle is all you need! Additionally, computing the cosine or sine is 40 times more expensive than a multiplication, meaning that a cosine operation might take 200 cycles (operations) and a multiplication might take 5 cycles. Arccosine and arcsine are yet more expensive.



Sketch 2.21.
The angle between two vectors.

Example 2.5

Let's calculate the angle between the two vectors illustrated in Sketch 2.21, forming an obtuse angle:

$$\mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \quad \text{and} \quad \mathbf{w} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}.$$

Calculate the length of each vector,

$$\|\mathbf{v}\| = \sqrt{2^2 + 1^2} = \sqrt{5}$$
$$\|\mathbf{w}\| = \sqrt{-1^2 + 0^2} = 1.$$

The cosine of the angle between the vectors is calculated using (2.20) as

$$\cos(\theta) = \frac{(2 \times -1) + (1 \times 0)}{\sqrt{5} \times 1} = \frac{-2}{\sqrt{5}} \approx -0.8944.$$

Then

$$\arccos(-0.8944) \approx 153.4^\circ.$$

To convert an angle given in degrees to radians multiply by $\pi/180^\circ$. (Recall that $\pi \approx 3.14159$ radians.) This means that

$$2.677 \text{ radians} \approx 153.4^\circ \times \frac{\pi}{180^\circ}.$$

2.8 Orthogonal Projections

Sketch 2.19 illustrates that the projection of the vector $\mathbf{w}$ onto $\mathbf{v}$ creates a footprint of length $b = ||\mathbf{w}|| \cos(\theta)$. This we derive from basic trigonometry: $\cos(\theta) = b/\text{hypotenuse}$. The *orthogonal projection* of $\mathbf{w}$ onto $\mathbf{v}$ is then the vector

$$\mathbf{u} = (||\mathbf{w}|| \cos(\theta)) \frac{\mathbf{v}}{||\mathbf{v}||} = \frac{\mathbf{v} \cdot \mathbf{w}}{||\mathbf{v}||^2} \mathbf{v}. \tag{2.21}$$

Sometimes this projection is expressed as

$$\mathbf{u} = \text{proj}_{\mathcal{V}_1} \mathbf{w},$$

where $\mathcal{V}_1$ is the set of all 2D vectors $k\mathbf{v}$ and it is referred to as a one-dimensional *subspace* of $\mathbb{R}^2$. Therefore, $\mathbf{u}$ is the *best approximation* to $\mathbf{w}$ in the subspace $\mathcal{V}_1$. This concept of closest or best approximation will be needed for several problems, such as finding the point at the end of the footprint in Section 3.7 and for least squares approximations in Section 12.7. We will revisit subspaces with more rigor in Chapter 14.

Using the orthogonal projection, it is easy to decompose the 2D vector $\mathbf{w}$ into a sum of two perpendicular vectors, namely $\mathbf{u}$ and $\mathbf{u}^\perp$ (a vector perpendicular to $\mathbf{u}$), such that

$$\mathbf{w} = \mathbf{u} + \mathbf{u}^\perp. \tag{2.22}$$

Another way to state this: we have *resolved* $\mathbf{w}$ into components with respect to two other vectors. Already having found the vector $\mathbf{u}$, we now set

$$\mathbf{u}^\perp = \mathbf{w} - \frac{\mathbf{v} \cdot \mathbf{w}}{||\mathbf{v}||^2} \mathbf{v}.$$

This can also be written as

$$\mathbf{u}^\perp = \mathbf{w} - \text{proj}_{\mathcal{V}_1} \mathbf{w},$$

and thus $\mathbf{u}^\perp$ is the component of $\mathbf{w}$ orthogonal to the space of $\mathbf{u}$.

See the Exercises of Chapter 8 for a 3D version of this decomposition. Orthogonal projections and vector decomposition are at the core of constructing the Gram-Schmidt orthonormal coordinate frame in Section 11.8 for 3D and in Section 14.4 for higher dimensions. An application that uses this frame is discussed in Section 20.7.

The ability to decompose a vector into its component parts is key to Fourier analysis, quantum mechanics, digital audio, and video recording.

2.9 Inequalities

Here are two important inequalities when dealing with vector lengths.

Let's start with the expression from (2.19), i.e.,

$$\mathbf{v} \cdot \mathbf{w} = \|\mathbf{v}\| \|\mathbf{w}\| \cos\theta.$$

Squaring both sides gives

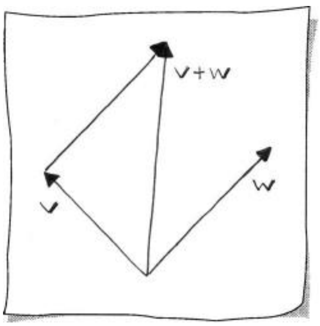
$$(\mathbf{v} \cdot \mathbf{w})^2 = \|\mathbf{v}\|^2 \|\mathbf{w}\|^2 \cos^2\theta.$$

Noting that $0 \leq \cos^2\theta \leq 1$, we conclude that

$$(\mathbf{v} \cdot \mathbf{w})^2 \leq \|\mathbf{v}\|^2 \|\mathbf{w}\|^2. \tag{2.23}$$

This is the *Cauchy-Schwartz inequality.* Equality holds if and only if $\mathbf{v}$ and $\mathbf{w}$ are linearly dependent. This inequality is fundamental in the study of more general vector spaces, which are presented in Chapter 14.

Suppose we would like to find an inequality that describes the relationship between the length of two vectors $\mathbf{v}$ and $\mathbf{w}$ and the length of their sum $\mathbf{v} + \mathbf{w}$. In other words, how does the length of the third side of a triangle relate to the lengths of the other two? Let's begin with expanding $\|\mathbf{v} + \mathbf{w}\|^2$:



Sketch 2.22.
The triangle inequality.

$$\begin{aligned} \|\mathbf{v} + \mathbf{w}\|^2 &= (\mathbf{v} + \mathbf{w}) \cdot (\mathbf{v} + \mathbf{w}) \\ &= \mathbf{v} \cdot \mathbf{v} + 2\mathbf{v} \cdot \mathbf{w} + \mathbf{w} \cdot \mathbf{w} \\ &\leq \mathbf{v} \cdot \mathbf{v} + 2|\mathbf{v} \cdot \mathbf{w}| + \mathbf{w} \cdot \mathbf{w} \\ &\leq \mathbf{v} \cdot \mathbf{v} + 2\|\mathbf{v}\|\|\mathbf{w}\| + \mathbf{w} \cdot \mathbf{w} \\ &= \|\mathbf{v}\|^2 + 2\|\mathbf{v}\|\|\mathbf{w}\| + \|\mathbf{w}\|^2 \\ &= (\|\mathbf{v}\| + \|\mathbf{w}\|)^2. \end{aligned} \tag{2.24}$$

Taking square roots gives

$$\|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|,$$

which is known as the *triangle inequality.* It states the intuitively obvious fact that the sum of any two edge lengths in a triangle is never smaller than the length of the third edge; see Sketch 2.22 for an illustration.

- point versus vector
- coordinates versus components
- $\mathbb{E}^2$ versus $\mathbb{R}^2$
- coordinate independent
- vector length
- unit vector
- zero divide tolerance
- Pythagorean theorem
- distance between two points
- parallelogram rule
- scaling
- ratio
- barycentric combination
- linear interpolation
- convex combination
- barycentric coordinates
- linearly dependent vectors
- linear combination
- basis for $\mathbb{R}^2$
- dot product
- Law of Cosines
- perpendicular vectors
- angle between vectors
- orthogonal projection
- vector decomposition
- Cauchy-Schwartz inequality
- triangle inequality

2.10 Exercises

1. Illustrate the parallelogram rule applied to the vectors

$$\mathbf{v} = \begin{bmatrix} -2 \\ 1 \end{bmatrix} \quad \text{and} \quad \mathbf{w} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

2. The parallelogram rule states that adding or subtracting two vectors, $\mathbf{v}$ and $\mathbf{w}$, yields another vector. Why is it called the parallelogram rule?

3. Define your own $\mathbf{p}, \mathbf{q} \in \mathbb{E}^2$ and $\mathbf{v}, \mathbf{w} \in \mathbb{R}^2$. Determine which of the following expressions are geometrically meaningful. Illustrate those that are.

(a)	$\mathbf{p} + \mathbf{q}$	(b)	$\frac{1}{2}\mathbf{p} + \frac{1}{2}\mathbf{q}$
(c)	$\mathbf{p} + \mathbf{v}$	(d)	$3\mathbf{p} + \mathbf{v}$
(e)	$\mathbf{v} + \mathbf{w}$	(f)	$2\mathbf{v} + \frac{1}{2}\mathbf{w}$
(g)	$\mathbf{v} - 2\mathbf{w}$	(h)	$\frac{3}{2}\mathbf{p} - \frac{1}{2}\mathbf{q}$

4. Suppose we are given $\mathbf{p}, \mathbf{q} \in \mathbb{E}^2$ and $\mathbf{v}, \mathbf{w} \in \mathbb{R}^2$. Do the following operations result in a point or a vector?

(a)	$\mathbf{p} - \mathbf{q}$	(b)	$\frac{1}{2}\mathbf{p} + \frac{1}{2}\mathbf{q}$
(c)	$\mathbf{p} + \mathbf{v}$	(d)	$3\mathbf{v}$
(e)	$\mathbf{v} + \mathbf{w}$	(f)	$\mathbf{p} + \frac{1}{2}\mathbf{w}$

5. What barycentric combination of the points $\mathbf{p}$ and $\mathbf{q}$ results in the midpoint of the line through these two points?

6. Illustrate a point with barycentric coordinates $(1/2, 1/4, 1/4)$ with respect to three other points.
7. Consider two points. Form the set of all convex combinations of these points. What is the geometry of this set?
8. Consider three noncollinear points. Form the set of all convex combinations of these points. What is the geometry of this set?
9. What is the length of the vector
$$\mathbf{v} = \begin{bmatrix} -4 \\ -3 \end{bmatrix}?$$
10. What is the magnitude of the vector
$$\mathbf{v} = \begin{bmatrix} 3 \\ -3 \end{bmatrix}?$$
11. If a vector $\mathbf{v}$ is length 10, then what is the length of the vector $-2\mathbf{v}$?
12. Find the distance between the points
$$\mathbf{p} = \begin{bmatrix} 3 \\ 3 \end{bmatrix} \quad \text{and} \quad \mathbf{q} = \begin{bmatrix} -2 \\ -3 \end{bmatrix}.$$
13. Find the distance between the points
$$\mathbf{p} = \begin{bmatrix} -3 \\ -3 \end{bmatrix} \quad \text{and} \quad \mathbf{q} = \begin{bmatrix} -2 \\ -3 \end{bmatrix}.$$
14. What is the length of the unit vector $\mathbf{u}$?
15. Normalize the vector $\mathbf{v} = \begin{bmatrix} -4 \\ -3 \end{bmatrix}$.
16. Normalize the vector $\mathbf{v} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$.
17. Given points
$$\mathbf{p} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{q} = \begin{bmatrix} 7 \\ 7 \end{bmatrix}, \quad \mathbf{r} = \begin{bmatrix} 3 \\ 3 \end{bmatrix},$$
what are the barycentric coordinates of $\mathbf{r}$ with respect to $\mathbf{p}$ and $\mathbf{q}$?
18. Given points
$$\mathbf{p} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{q} = \begin{bmatrix} 7 \\ 7 \end{bmatrix}, \quad \mathbf{r} = \begin{bmatrix} 5 \\ 5 \end{bmatrix},$$
what are the barycentric coordinates of $\mathbf{r}$ with respect to $\mathbf{p}$ and $\mathbf{q}$?
19. If $\mathbf{v} = 4\mathbf{w}$, are $\mathbf{v}$ and $\mathbf{w}$ linearly independent?
20. If $\mathbf{v} = 4\mathbf{w}$, what is the area of the parallelogram spanned by $\mathbf{v}$ and $\mathbf{w}$?

21. Do the vectors
$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 4 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$$
form a basis for $\mathbb{R}^2$?

22. What linear combination allows us to express $\mathbf{u}$ with respect to $\mathbf{v}_1$ and $\mathbf{v}_2$, where
$$\mathbf{u} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}, \quad \mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 4 \\ 4 \end{bmatrix}?$$

23. Show that the dot product has the following properties for vectors $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^2$.
$$\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u} \qquad \text{symmetric}$$
$$\mathbf{v} \cdot (s\mathbf{w}) = s(\mathbf{v} \cdot \mathbf{w}) \qquad \text{homogeneous}$$
$$(\mathbf{v} + \mathbf{w}) \cdot \mathbf{u} = \mathbf{v} \cdot \mathbf{u} + \mathbf{w} \cdot \mathbf{u} \qquad \text{distributive}$$
$$\mathbf{v} \cdot \mathbf{v} > 0 \quad \text{if} \quad \mathbf{v} \neq \mathbf{0} \quad \text{and} \quad \mathbf{v} \cdot \mathbf{v} = 0 \quad \text{if} \quad \mathbf{v} = \mathbf{0} \quad \text{positive}$$

24. What is $\mathbf{v} \cdot \mathbf{w}$ where
$$\mathbf{v} = \begin{bmatrix} 5 \\ 4 \end{bmatrix} \quad \text{and} \quad \mathbf{w} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}?$$
What is the scalar product of $\mathbf{w}$ and $\mathbf{v}$?

25. Compute the angle (in degrees) formed by the vectors
$$\mathbf{v} = \begin{bmatrix} 5 \\ 5 \end{bmatrix} \quad \text{and} \quad \mathbf{w} = \begin{bmatrix} 3 \\ -3 \end{bmatrix}.$$

26. Compute the cosine of the angle formed by the vectors
$$\mathbf{v} = \begin{bmatrix} 5 \\ 5 \end{bmatrix} \quad \text{and} \quad \mathbf{w} = \begin{bmatrix} 0 \\ 4 \end{bmatrix}.$$
Is the angle less than or greater than $90°$?

27. Are the following angles acute, obtuse, or right?
$$\cos\theta_1 = -0.7 \quad \cos\theta_2 = 0 \qquad \cos\theta_3 = 0.7$$

28. Given the vectors
$$\mathbf{v} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \text{ and } \mathbf{w} = \begin{bmatrix} 3 \\ 2 \end{bmatrix},$$
find the orthogonal projection $\mathbf{u}$ of $\mathbf{w}$ onto $\mathbf{v}$. Decompose $\mathbf{w}$ into components $\mathbf{u}$ and $\mathbf{u}^\perp$.

29. For
$$\mathbf{v} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \quad \text{and} \quad \mathbf{w} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$
find
$$\mathbf{u} = \operatorname{proj}_{\mathcal{V}_1} \mathbf{w},$$
where $\mathcal{V}_1$ is the set of all 2D vectors $k\mathbf{v}$, and find
$$\mathbf{u}^{\perp} = \mathbf{w} - \operatorname{proj}_{\mathcal{V}_1} \mathbf{w}.$$

30. Given vectors $\mathbf{v}$ and $\mathbf{w}$, is it possible for $(\mathbf{v} \cdot \mathbf{w})^2$ to be greater than $\|\mathbf{v}\|^2 \|\mathbf{w}\|^2$?
31. Given vectors $\mathbf{v}$ and $\mathbf{w}$, under what conditions is $(\mathbf{v} \cdot \mathbf{w})^2$ equal to $\|\mathbf{v}\|^2 \|\mathbf{w}\|^2$? Give an example.
32. Given vectors $\mathbf{v}$ and $\mathbf{w}$, can the length of $\|\mathbf{v} + \mathbf{w}\|$ be longer than the length of $\|\mathbf{v}\| + \|\mathbf{w}\|$?
33. Given vectors $\mathbf{v}$ and $\mathbf{w}$, under what conditions is $\|\mathbf{v}+\mathbf{w}\| = \|\mathbf{v}\|+\|\mathbf{w}\|$? Give an example.