These inner and outer approximations are analogous to upper and lower sets of a Dedekind cut. And like them, they make it easy to define the *sum* of angles. Recall that our purpose in defining arc length on the unit circle was to define angle measure. We now want to see whether the measure of a sum of angles is the sum of their measures. When angles are added by joining them along a common ray, the corresponding arcs are joined at a common point. One certainly expects the length of the combined arc to be the sum of the lengths of the two pieces, and in fact this follows from the definition of arc length as a least upper bound.

5.1.5. Show the following, for arcs $AB$, $BC$, and their combined arc $AC$:

- Any polygons drawn inside $AB$ and $BC$ have total length $<$ some polygon drawn inside $AC$.
- Any polygon drawn inside $AC$ has length $<$ the sum of the lengths of some polygons drawn inside $AB$ and $BC$.
- It follows that $\text{length}(AC) = \text{length}(AB) + \text{length}(BC)$.

## 5.2 Circular Functions

We first meet the circular functions sine and cosine at school, as ratios of sides of right-angled triangles (Figure 5.5). These ratios depend only on the size $\theta$ of the angle, and not the size of the triangle, because of the basic property of straight lines, in fact, by the *defining* property of straight lines in analytic geometry (Section 3.1):

$$\cos\theta = \frac{a}{c}, \qquad \sin\theta = \frac{b}{c}.$$

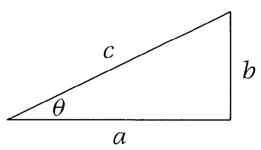


**FIGURE 5.5** Defining cos and sin via a triangle.

In this context, cos and sin are called *trigonometric* functions, because they assist in the measurement of triangles. However, this definition limits their domain to angles $\theta$ less than $\pi/2$, which is inconvenient for at least two reasons:

- There are formulas for $\cos(\theta+\phi)$ and $\sin(\theta+\phi)$ in terms of $\cos\theta$, $\cos\phi$, $\sin\theta$, $\sin\phi$, which suggest a meaning for $\cos(\theta+\phi)$ and $\sin(\theta+\phi)$ when $\theta+\phi > \pi/2$.
- The functions cos and sin not only give sides of triangles as functions of angle, but also amplitude of vibration as a function of time or the height of a wave as a function of distance—and the time or distance can be *any* real number.

This leads us to extend the definition of cos and sin so that they make sense for any real number $\theta$. It then becomes more appropriate to call them *circular* functions.

We take the unit circle (Figure 5.6) and view the coordinates $x$ and $y$ of any point $P$ on it as functions $\cos\theta$ and $\sin\theta$ of the angle $\theta$. This gives a meaning to $\cos\theta$ and $\sin\theta$ for all $\theta$ from 0 to $2\pi$, but there is no reason to stop there. It is natural to define cos and sin for $\theta$ outside this interval by the equations

$$\cos(\theta+2\pi) = \cos\theta,$$
$$\sin(\theta+2\pi) = \sin\theta,$$

because increasing $\theta$ by $2\pi$ means making a complete circuit, and hence returning to the same point $P$ on the circle. Likewise, there is no reason to distinguish between "angle $\theta$" and "angle $\theta+2\pi$." Any two real numbers that differ by $2\pi$ represent the same angle, so an

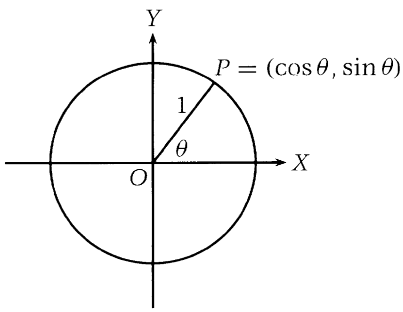


**FIGURE 5.6** Defining cos and sin via the circle.

*angle* is really a *set* of real numbers of the form

$$\{\theta + 2n\pi : n \in \mathbb{Z}\} = \{\ldots, \theta - 4\pi, \theta - 2\pi, \theta, \theta + 2\pi, \theta + 4\pi, \ldots\},$$

obtained by adding all integer multiples of $2\pi$ to $\theta$. As in previous cases (such as Dedekind cuts), there are advantages to defining a mathematical object as a set. We now have no problem defining the sum of angles:[1] the *sum* of the angle $\{\theta + 2n\pi : n \in \mathbb{Z}\}$ and the angle $\{\phi + 2n\pi : n \in \mathbb{Z}\}$ is simply the angle $\{\theta + \phi + 2n\pi : n \in \mathbb{Z}\}$.

When cos and sin are related to the circle in this way, it becomes obvious why they are relevant to rotation and vibration. If the point $P$ travels around the circle at constant angular velocity, so that $\theta$ measures time as well as angle, then $x = \cos\theta$ and $y = \sin\theta$ measure the horizontal and vertical displacements of the uniformly rotating point. The $x$-coordinate of $P$ can be viewed as the position of its shadow under a light shining vertically downward, and movement of this shadow is the simplest form of vibration. It is called *simple harmonic motion* because such vibration is the basis of musical tones.

It is also obvious, from the circular interpretation of cos and sin, that their graphs have the same shape and that the graph of sin lags $\pi/2$ behind the graph of cos. (Figure 5.7 shows $y = \sin x$ drawn heavily and $y = \cos x$ drawn dotted). The shape is known as the *sine wave*.

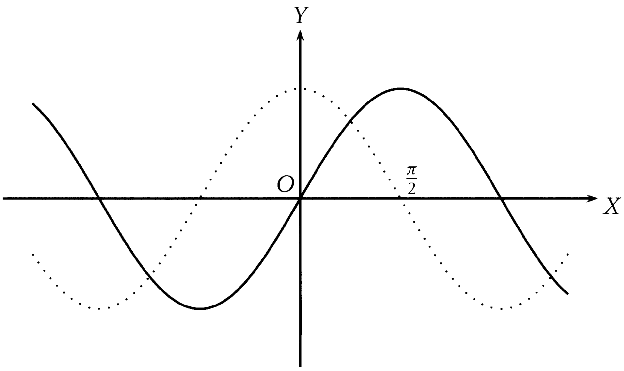


**FIGURE 5.7** Graphs of the cos and sin functions.

[1]Note, however, that when we speak of the "angle sum" of a triangle, quadrilateral, and so on, we take the angles to be real numbers between 0 and $2\pi$, and we use the sum of reals.

## Exercises

The properties of cos and sin bear out the claim that the concept of angle measure lies outside elementary geometry. In fact, cos and sin are *transcendental functions*, which lie outside the realm of *algebraic functions* we have considered so far. In general, a function $y(x)$ is called algebraic if $p(x, y) = 0$ for some polynomial $p$ in the two variables $x$ and $y$. The graph $p(x, y) = 0$ of an algebraic function is called an *algebraic curve*.

For example, $y = \sqrt{1 - x^2}$ is an algebraic function of $x$, because it satisfies the equation

$$x^2 + y^2 = 1,$$

which is of the form $p(x, y) = 0$, with $p(x, y) = x^2 + y^2 - 1$. In this case the algebraic curve is simply the unit circle.

The curve $y = \sin x$ is *not* an algebraic curve, and hence $\sin x$ is not an algebraic function. The reason is that the sine curve meets the line $y = 0$ infinitely often, namely, at the points $x = n\pi$ for all integers $n$. An algebraic curve $p(x, y) = 0$, on the other hand, meets the line $y = mx + c$ where $p(x, mx + c) = 0$, which is a polynomial equation whose roots are the $x$-coordinates of the points of intersection. Such an equation cannot have infinitely many roots. In fact, a polynomial equation of degree $n$ can have at most $n$ roots. The following exercises give one way to see this.

5.2.1. Check that $x^n - a^n = (x - a)(x^{n-1} + x^{n-2}a + \cdots + ax^{n-2} + x^{n-1})$.

5.2.2. Deduce from Exercise 5.2.1 that if $p(x)$ is polynomial of degree $n$, then

$$p(x) - p(a) = (x - a)q(x),$$

where $q(x)$ is polynomial of degree $n - 1$.

5.2.3. Deduce from Exercise 5.2.2 that if $p(x)$ is a polynomial of degree $n > 0$, then $p(a)$ cannot be zero for more than $n$ different numbers $a$.

Thus cos and sin are examples of transcendental functions. The same is true of any function $f$ that satisfies an equation of the form

$$f(x + \alpha) = f(x) \quad \text{for some } \alpha \neq 0,$$

because its graph meets the horizontal line $y = f(\alpha)$ for infinitely many values of $x$. We call such a function *periodic*, with *period* $\alpha$. Many other periodic functions can be built from cos and sin, for example,

$$\tan x = \frac{\sin x}{\cos x} \qquad \text{and} \qquad \cot x = \frac{\cos x}{\sin x}.$$

The cotangent function, cot, is noteworthy because Euler (1748) discovered a formula that "shows" its period, namely,

$$\pi \cot \pi x = \cdots + \frac{1}{x-2} + \frac{1}{x-1} + \frac{1}{x} + \frac{1}{x+1} + \frac{1}{x+2} + \cdots .$$

5.2.4. Give a geometric reason why $\cot \pi x$ has period 1. Why does Euler's formula show that $\pi \cot \pi(x+1) = \pi \cot \pi x$?

5.2.5. Euler's formula suggests that $\cot \pi x$ tends to infinity as $x$ approaches any integer value. Give a geometric explanation of this behavior.

Euler's formula for $\pi \cot \pi x$ is probably the simplest one can imagine that shows periodicity, so there may be a sense in which the cot function is more fundamental than cos and sin. In fact, we shall see in Section 5.3 that either tan or cot may be used as a "primitive" circular function, with both cos and sin defined in terms of it.

Formulas for circular functions are most easily derived using calculus, but it takes time to build calculus to the point where it works efficiently. Instead, we shall get by with a few *limit properties* of the circular functions, the most important of which is:

$$\lim_{\theta \to 0} \frac{\sin \theta}{\theta} = 1.$$

This means we can ensure that $\sin \theta / \theta$ is within any given positive distance (say, $\varepsilon$) of 1, by choosing $\theta$ sufficiently small (say, $< \delta$).

5.2.6. By referring to Figure 5.8 and the definition of angle measure, show that

$$\sin \theta < \theta < \tan \theta \quad \text{and hence that} \quad \cos \theta < \frac{\sin \theta}{\theta} < 1.$$

5.2.7. Deduce from Exercise 5.2.6 that $\lim_{\theta \to 0} (\sin \theta)/\theta = 1$.

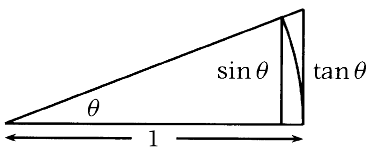


**FIGURE 5.8** Comparing sin, arc, and tan.

## 5.3 Addition Formulas

The functions cos and sin are necessarily complicated, inasmuch as they are transcendental, but there are still some simple relations between them. For example, because $x = \cos\theta$ and $y = \sin\theta$ are the coordinates of a point $(x, y)$ on the circle $x^2 + y^2 = 1$, we necessarily have

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

We usually write this

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

even though the notation $\cos^2\theta$ conflicts with the notation $\cos^{-1}\theta$ for the inverse cosine. When there is a danger of misunderstanding, it is wise to write $(\cos\theta)^2$ for the square of $\cos\theta$.

This relation enables us to express either one of $\cos\theta$ or $\sin\theta$ in terms of the other, namely,

$$\cos\theta = \sqrt{1 - \sin^2\theta} \quad \text{and} \quad \sin\theta = \sqrt{1 - \cos^2\theta}.$$

There is a cost, however, because now we have to worry about the sign of the square root. As we shall see in the next section, there are unambiguous formulas for both $\cos\theta$ and $\sin\theta$ in terms of $\tan\frac{\theta}{2}$.

After $\cos^2\theta + \sin^2\theta = 1$, the most important relations between cos and sin are the so-called *addition formulas*:

$$\cos(\theta + \phi) = \cos\theta\cos\phi - \sin\theta\sin\phi,$$
$$\sin(\theta + \phi) = \sin\theta\cos\phi + \cos\theta\sin\phi.$$

We prove just the first of these, because the second is similar (and in fact it follows from the first). The proof refers to Figure 5.9.

Looking first at the right-angled triangle $OAC$, we see

$$OA = \cos\phi \quad \text{and} \quad AC = \sin\phi.$$

Next, viewing $OA$ and $AC$ as the hypotenuses of the right-angled triangles $ODA$ and $ABC$, respectively, we see

$$OD = \cos\theta\cos\phi \quad \text{and} \quad AB = \sin\theta\sin\phi$$

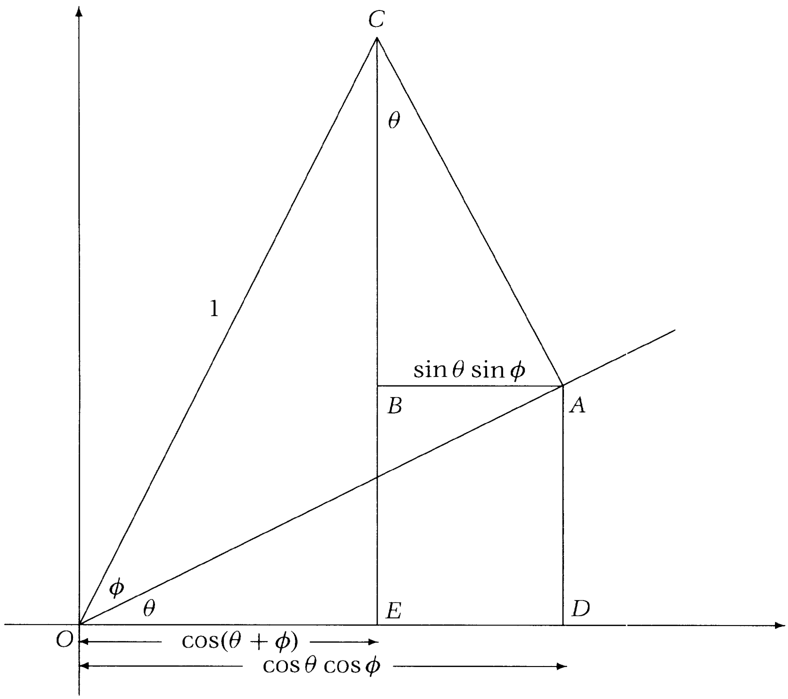


**FIGURE 5.9** Constructing the cosine of a sum.

(the latter because angle $ACB = \theta$, because angle $OAC = \pi/2$). This finally gives

$$\cos(\theta + \phi) = OE = OD - AB = \cos\theta\cos\phi - \sin\theta\sin\phi,$$

which is the required result. □

# Exercises

The addition formula for cosine is useful, but not quite as simple or memorable as one would like. The same goes for the addition formula for sine.

5.3.1. Prove the addition formula for sine, $\sin(\theta + \phi) = \sin\theta\cos\phi + \cos\theta\sin\phi$,

- by using Figure 5.9, but with appropriate vertical lengths instead of horizontal lengths, or
- by deducing it from the addition formula for cosine, with the help of the formulas $\sin\alpha = \cos(\frac{\pi}{2} - \alpha)$ and $\cos\alpha = \sin(\frac{\pi}{2} - \alpha)$.

By a kind of miracle, these two somewhat complicated formulas are parts of one simple formula. To state it, we have to use the "imaginary number" $i = \sqrt{-1}$, which will be explained more fully in Chapter 7. For the moment it is enough to know that $i^2 = -1$ and that if $\alpha + i\beta = \gamma + i\delta$ for real numbers $\alpha, \beta, \gamma, \delta$ then $\alpha = \gamma$ and $\beta = \delta$. Then the famous *de Moivre formula* (1730) is:

$$\cos(\theta + \phi) + i\sin(\theta + \phi) = (\cos\theta + i\sin\theta)(\cos\phi + i\sin\phi).$$

5.3.2. Verify the de Moivre formula, by multiplying out the right-hand side and using the addition formulas for cos and sin.

The de Moivre formula is so simple, compared with the cosine and sine formulas, that it seems that the function $\cos\theta + i\sin\theta$ is simpler than either its "real part" $\cos\theta$ or its "imaginary part" $\sin\theta$. Indeed, if we abbreviate $\cos\theta + i\sin\theta$ by $\operatorname{cis}\theta$, the addition formulas for cos and sin unite in the spectacularly simple addition formula for cis:

$$\operatorname{cis}(\theta + \phi) = \operatorname{cis}\theta \cdot \operatorname{cis}\theta.$$

If this reminds you of the exponential function, it should! However, we are getting ahead of the story. First, there are some important applications of the addition formulas to consider.

An easy special case is $\cos 2\theta = \cos^2\theta - \sin^2\theta$, which can be rewritten using $\cos^2\theta + \sin^2\theta = 1$ as $\cos 2\theta = 2\cos^2 - 1 = 1 - 2\sin^2\theta$.

5.3.3. From $\cos 2\theta = 1 - 2\sin^2\theta$ deduce $1 - \cos\theta = \sin^2\frac{\theta}{2}$, and hence show that $\lim_{\theta\to 0}(1 - \cos\theta)/\theta = 0$ with the help of the result $\lim_{\theta\to 0}(\sin\theta)/\theta = 1$ from Exercise 5.2.7.

By combining these limit results with the sine addition formula, we can find the tangent to the sine wave at any point, which in calculus is called "finding the derivative of $\sin x$." The tangent to any curve at a point $P$, if it exists, is found as the limiting position of a chord between $P$ and a point $Q \neq P$, as $Q$ approaches $P$. In the case of the sine wave $y = \sin x$, we take $P = (\alpha, \sin\alpha)$, $Q = (\alpha + \theta, \sin(\alpha + \theta))$, and let $\theta \to 0$.

5.3.4. Show that the slope of the chord between $P$ and $Q$ is

$$\frac{\sin(\alpha+\theta)-\sin\alpha}{\theta}=\frac{\sin\alpha(\cos\theta-1)+\cos\alpha\sin\theta}{\theta},$$

and deduce that the slope of the tangent at $x=\alpha$ is $\cos\alpha$.

5.3.5. Similarly use the cosine addition formula to show that the slope of the tangent to $y=\cos x$ at $x=\alpha$ is $-\sin\alpha$.

The de Moivre formula also makes it easy to find formulas for $\cos n\theta$ and $\sin n\theta$. It is as easy as expanding $(\cos\theta+i\sin\theta)^n$.

5.3.6. By expanding $(\cos\theta+i\sin\theta)^3$, show that

$$\cos 3\theta = 4\cos^3\theta - 3\cos\theta,$$
$$\sin 3\theta = 3\sin\theta - 4\sin^3\theta.$$

5.3.7. Show that $\cos n\theta$ is a polynomial in $\cos\theta$, for any natural number $n$. What is the situation for $\sin n\theta$?

These polynomials were discovered by Viète (1579) and were used by him to solve certain polynomial equations by circular functions. In 1593 he won a mathematical contest by noticing that a 45th-degree equation posed by his opponent was based on the polynomial for $\sin 45\theta$.

Another famous discovery of Viète is also based on addition formulas: his infinite product

$$\begin{aligned}\frac{2}{\pi}&=\cos\frac{\pi}{4}\cos\frac{\pi}{8}\cos\frac{\pi}{16}\cdots\\&=\frac{\sqrt{2}}{2}\frac{\sqrt{2+\sqrt{2}}}{2}\frac{\sqrt{2+\sqrt{2+\sqrt{2}}}}{2}\cdots.\end{aligned}$$

5.3.8.* Use the sine addition formula to show in turn that

$$\begin{aligned}\sin\theta&=2\sin\frac{\theta}{2}\cos\frac{\theta}{2},\\ \frac{\sin\theta}{2^n\sin(\theta/2^n)}&=\cos\frac{\theta}{2}\cos\frac{\theta}{2^2}\cdots\cos\frac{\theta}{2^n},\\ \frac{\sin\theta}{\theta}&=\cos\frac{\theta}{2}\cos\frac{\theta}{2^2}\cos\frac{\theta}{2^3}\cdots,\end{aligned}$$

and deduce Viète's product by substituting $\theta=\pi/2$.

## 5.4 A Rational Addition Formula

The addition formulas for cos and sin give the *double angle* formulas

$$\cos 2\theta = \cos^2\theta - \sin^2\theta,$$
$$\sin 2\theta = 2\sin\theta\cos\theta.$$

By rewriting $\cos 2\theta$ we find

$$\begin{aligned}\cos 2\theta &= \frac{\cos^2\theta - \sin^2\theta}{\cos^2\theta + \sin^2\theta} && \text{because } \cos^2\theta + \sin^2\theta = 1\\ &= \frac{1-\tan^2\theta}{1+\tan^2\theta}, && \text{dividing numerator and denominator by } \cos^2\theta.\end{aligned}$$

Similarly,

$$\begin{aligned}\sin 2\theta &= \frac{2\sin\theta\cos\theta}{\cos^2\theta + \sin^2\theta} && \text{because } \cos^2\theta + \sin^2\theta = 1\\ &= \frac{2\tan\theta}{1+\tan^2\theta}, && \text{dividing numerator and denominator by } \cos^2\theta.\end{aligned}$$

Finally, replacing $\theta$ by $\theta/2$, we get the *half angle formulas* expressing $\cos\theta$ and $\sin\theta$ rationally in terms of $\tan\frac{\theta}{2}$:

$$\cos\theta = \frac{1-\tan^2\frac{\theta}{2}}{1+\tan^2\frac{\theta}{2}}, \qquad \sin\theta = \frac{2\tan\frac{\theta}{2}}{1+\tan^2\frac{\theta}{2}}.$$

This supports our claim from Section 5.2 that the tan function may be considered more fundamental than either cos or sin. The surprise is that we already know these formulas! They are essentially the formulas used by Diophantus to find rational Pythagorean triples. Look again at the diagram we used in Section 4.2 to explain Diophantus' construction, and the role of $\tan\frac{\theta}{2}$ becomes clear (Figure 5.10).

If we take the angle to the point $(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2})$ to be $\theta$ as shown, then the line from $(-1, 0)$ to the same point is at angle $\frac{\theta}{2}$, by the theorem that the angle at the circumference is half the angle at the center (Section 2.4). It follows that the slope $t$ of the line is $\tan\frac{\theta}{2}$, and the coordinates $\frac{1-t^2}{1+t^2}$ and $\frac{2t}{1+t^2}$ are $\cos\theta$ and $\sin\theta$, respectively.

This prompts the thought that we should be able to add angles by calculating the corresponding slopes, and hence work with rational functions instead of the transcendental functions cos and sin. What