2 Instantaneous Rate of Change: The Derivative

2.1 The Slope of a Function

Suppose that \( y = f(x) \) is a function of \( x \), and \( x \) is small changes in \( x \).

\[ \text{Example 2.1.1} \]
Take, for example, \( y = f(x) = \sqrt{625 - x^2} \) (the upper semicircle of radius 25 centered at the origin). When \( x = 7 \), we find that \( y = \sqrt{625 - 49} = 24 \). Suppose we want to know how much \( y \) changes when \( x \) increases a little, say to 7.1 or 7.01.

In the case of a straight line \( y = mx + b \), the slope \( m = \frac{\Delta y}{\Delta x} \) measures the change in \( y \) per unit change in \( x \). This can be interpreted as a measure of “sensitivity”; for example, if \( y = 100x + 5 \), a small change in \( x \) corresponds to a change one hundred times as large in \( y \), so \( y \) is quite sensitive to changes in \( x \).

Let us look at the same ratio \( \Delta y/\Delta x \) for our function \( y = f(x) = \sqrt{625 - x^2} \) when \( x \) changes from 7 to 7.1. Here \( \Delta x = 7.1 - 7 = 0.1 \) is the change in \( x \), and

\[ \Delta y = (f(7 + \Delta x) - f(7)) = \left( \sqrt{625 - (7 + \Delta x)^2} \right) - \left( \sqrt{625 - 7^2} \right) \]

Thus, \( \Delta y/\Delta x \approx -0.0294 \). This means that \( y \) changes by less than 0.1 the change in \( x \). Perhaps \( y \) changes dramatically as \( x \) runs through the values from 7 to 7.1, but at \( x = 7 \) \( y \) just happens to be close to its value at 7. This is a case for this particular function, but we don’t yet know why.

2.1.2 The Slope of a Function

Instead of looking at more particular values of \( \Delta x \), let’s see what happens if we do some algebra with the difference quotient using just \( \Delta x \). The slope of a chord from \((7, 24)\) to a nearby point is given by

\[
\frac{\sqrt{625 - (7 + \Delta x)^2} - 24}{\Delta x} = \frac{\sqrt{625 - (7 + \Delta x)^2} - 24}{\Delta x}
\]

Now, can we tell by looking at this last formula what happens when \( \Delta x \) gets very close to zero? The numerator clearly gets very close to \( -2x \) while the denominator gets very close to \( \sqrt{625 - 7^2} + 48 \). It certainly seems reasonable, and in fact it is true: as \( \Delta x \) gets closer and closer to zero, the difference quotient does in fact get closer and closer to \(-2x\), and so the slope of the tangent line is exactly \(-2x\).

What about the slope of the tangent line at \( x = 127\)? Well, 127 can’t be at all different from 7, we just have to redo the calculation with 12 instead of 7. This won’t be hard, but it will be a bit tedious. What if we try to do all the algebra without using a specific value for \( x \)? Let’s copy from above, replacing 7 by \( x \). We’ll have to do a bit more than that—for example, the “24” in the calculation came from \( \sqrt{625 - 7^2} \). We’ll need to fix that too.

So far we have found the slopes of two chords that should be close to the slope of the tangent line, but what is the slope of the tangent line exactly? Since the tangent line touches the circle at just one point, we will never be able to calculate its slope directly, using two “known” points on the line. What we need is a way to capture what happens to the slopes of the chords as they get “closer and closer” to the tangent line.
We should note that in the particular case of a circle, there’s a simple way to find the derivative. Since the tangent to a circle at a point is perpendicular to the radius drawn to the point of contact, its slope is the negative reciprocal of the slope of the radius. The radius joining (0, 0) to (7, 24) has slope $-24/7$. Hence, the tangent line has slope $24/7$, so if the slope of the tangent line is $-\sqrt{24^2 + 7^2}$, as before. It is NOT always true that a tangent line is perpendicular to a line from the origin—don’t use this shortcut in any other circumstance.

As above, and as you might expect, for different values of $y$ we generally get different values of the derivative $f'(x)$. Could it be that the derivative always has the same value?

This would mean that the slope of $f$ or the slope of the tangent line, is the same everywhere. One curve that always has the same slope is a line; it seems odd to talk about the tangent to a circle at a point. But certainly we can do better. At $t = 2$ the height is $100 - 4\cdot 4.9 = 95.7$.

We started the last section by saying, “It is often necessary to know how sensitive the value of $y$ is to small changes in $x$.” We have seen one purely mathematical example of this: finding the “steepness” of a curve at a point is precisely this problem. Here is a more applied example.

With careful measurement it might be possible to discover that a dropped ball has been at $(h - x)^2 = 9\Delta t^2$. But if we neglect air resistance, the height is the initial height of the ball, when $t = 0$, and $k$ is some number determined by the experiment. A natural question is then, “How fast is the ball going at time $t$?” We can certainly get a pretty good idea with a little simple arithmetic. To make the calculation more concrete, let’s say $h_0 = 100$ meters and $k = 4.9$ and suppose we’re interested in the speed of $t = 2$. We know that when $t = 1$ the height is $100 - 4\cdot 4.9 = 55.9$, so in that second the ball has traveled $80.4 - 55.9 = 24.5$ meters. This means that the average speed during that time was $24.5$ meters per second. So we might guess that $24.5$ meters per second is not a terrible estimate of the speed at $t = 2$. But certainly we can do better. At $t = 2.5$ the height is $100 - 4\cdot (2.5)^2 = 69.75$. During the half second from $t = 2$ to $t = 2.5$ the ball dropped $30.4 - 69.75 = 39.35$ meters, at an average speed of $11.025/(1/2) = 22.05$ meters per second; this should be a better estimate of the speed at $t = 2$. So it’s clear now how to get better and better approximations: compute average speeds over shorter and shorter time intervals. Between $t = 2$ and $t = 2.01$, for example, the ball drops $0.19649$ meters in one hundredth of a second, at an average speed of $19.649$ meters per second.

We can’t do this forever, and we still might reasonably ask what the actual speed precisely at $t = 2$ is.

An object is traveling in a straight line so that its position (that is, distance from some fixed point) at time $t$ is given by the formula $f(t) = \frac{1}{2}gt^2$. If $\Delta t$ is some small number, and $\Delta y$ is the object’s change in position at time $t$ + $\Delta t$ compared with time $t$, the average velocity of the object over the time interval $\Delta t$ is $\frac{\Delta y}{\Delta t} = \frac{f(t + \Delta t) - f(t)}{\Delta t}$.

If the average speed of the object during the following time intervals: $[0, 1.0]$, $[0.0, 1.1]$, $[0.01, 1.01]$, $[0.001, 1.001]$. How good is this approximation?

If you had to guess the average speed at $t = 2$ just on the basis of these, what would you guess?

2. Let $y = f(t) = t^3$, where $t$ is the time in seconds and $y$ is the distance in meters that an object falls on a certain airless planet. Draw a graph of this function between $t = 0$ and $t = 3$. Make a table of the average speed of the falling object between $(a) 2$ sec and $1$ sec, $(b) 2$ sec and $2.1$ sec, $(c) 2$ sec and $2.01$ sec, and $(d) 2$ sec and $2.001$ sec. Then use algebra to find a simple formula for the average speed of time $t$ and $t + \Delta t$. (If you substitute $\Delta x = 1, 0.1, 0.01, 0.001$ in this formula you should also get the answers to parts (a)–(d).)

Next, in your formula for average speed (which should be in simplified form) determine what happens as $\Delta t$ approaches zero. This is the instantaneous speed. Finally, in your graph of $y = t^3$, draw the straight line through the point $(2.4)$ whose slope is the instantaneous velocity you just computed; it should of course be the tangent line. $\ddot{y}$

3. Find the average speed of the object during the following time intervals: $[0.0, 1.0]$, $[0.0, 1.1]$, $[0.01, 1.01]$, $[0.001, 1.001]$. How good is this approximation?

The speed of the object is the distance traveled divided by the time. If the object is moving at a constant speed of 10 meters per second, the distance traveled is 10 meters. If the object is moving at a constant speed of 10 meters per second, the time traveled is 1 second.

The slope of a function is the rate of change of a function at a point. It is the limit of the difference quotient as $\Delta t$ approaches zero. In other words, the slope of the tangent line to a curve at a point is the same as the derivative of the function at that point. The derivative of a function at a point is the instantaneous rate of change of the function at that point. It is the limit of the difference quotient as $\Delta t$ approaches zero.

The slope of a function is the derivative of a function. It is the limit of the difference quotient as $\Delta t$ approaches zero.

The slope of a function is the derivative of a function. It is the limit of the difference quotient as $\Delta t$ approaches zero.
which is meaningless. The quantity we are really interested in does not make sense "at zero," and this is why the answer to the original problem (finding a velocity or a slope) was not immediately obvious. In other words, we are generally going to want to figure out what a quantity "approaches" in situations where we can't merely plug in a value. If you would like to think about a hard example (which we will analyze later) consider what happens to \( \sin(1/x) \) as \( x \) approaches zero.

**EXAMPLE 2.3.1** Does \( \sqrt{2} \) approach 1.41 as \( x \) approaches 2? In this case it is possible to compute the actual value \( \sqrt{2} \) to a high precision to answer the question. But since in general we won't be able to do that, let's not. We might start by computing \( \sqrt{2} \) for values of \( x \) close to 2, as we did in the previous sections. Here are some values: \( \sqrt{2} = 1.414213562, \sqrt{2.000001} = 1.414213562, \sqrt{2.000000001} = 1.4142135626, \sqrt{2.000000000000001} = 1.41421356237 \) ... So it looks at least possible that indeed these values "approach" 1.41—already \( \sqrt{2} \) is quite close. If we continue this process, however, at some point we will appear to "stall." In fact, \( \sqrt{2} = 1.41421356237309504880168872420969807856967187537694807385386366571253\ldots \), so we will never even get as far as 1.4142, no matter how long we continue the sequence.

So in a fuzzy, everyday sort of sense, it is true that \( \sqrt{2} \) "gets close to" 1.41, but it does not "approach" 1.41 in the precise sense we want. To compute an exact slope or an exact velocity, what we want to know is that a given quantity becomes "arbitrarily close" to a fixed value, meaning that the first quantity can be made "as close as we like" to the fixed value. Consider again the quantities

\[
\frac{19.6 - 1.99}{x} - 19.6 < 4.9 \Delta x, \\
\Delta x \to 0
\]

Here is the actual, official definition of "limit".

**DEFINITION 2.3.2 Limit** Suppose \( f \) is a function. We say that \( \lim_{x \to a} f(x) = L \) if for every \( \varepsilon > 0 \) there is a \( \delta > 0 \) so that whenever \( 0 < |x - a| < \delta \), \( |f(x) - L| < \varepsilon \).

The \( \alpha \) and \( \delta \) here play exactly the role they did in the preceding discussion. The definition says, in a very precise way, that \( f(x) \) can be made as close as desired to \( L \) (that's the \( |f(x) - L| < \varepsilon \) part) by making \( x \) close enough to \( a \) (the \( 0 < |x - a| < \delta \) part). Note that we specifically make no mention of what must happen if \( x = a \), that is, if \( |x - a| = 0 \). This is because in the cases we are most interested in, substituting \( x = a \) doesn't even make sense.

Make sure you are not confused by the names of important quantities. The generic definition talks about \( f(x) \), but the function and the variable might have other names. In the discussion above, the function we analyzed was

\[
\frac{19.6 - 1.99}{x} - 19.6 < 4.9 \Delta x, \\
\Delta x \to 0
\]

and the variable of the limit was not \( x \) but \( \Delta x \). The \( x \) was the variable of the original function, when we were trying to compute a slope or a velocity; \( x \) was essentially a fixed quantity, telling us at what point we wanted the slope. (In the velocity problem, it was literally a fixed quantity, as we focused on the time 2.) The quantity \( \Delta x \) of the definition in all the examples was zero: we were always interested in what happened as \( \Delta x \) became very close to zero.

Armed with a precise definition, we can now prove that certain quantities behave in a particular way. The bad news is that even proofs for simple quantities can be quite tedious and complicated, the good news is that we rarely need to do such proofs, because most expressions act the way you would expect, and this can be proved once and for all.

**EXAMPLE 2.3.3** Let's show carefully that \( \lim_{x \to 4} x = 4 + 6 \). This is nothing we would "prove", since it is "obviously" true. But if we couldn't prove it using our official definition there would be something very wrong with the definition.

As is often the case in mathematical proofs, it helps to work backwards. We want to end up showing that under certain circumstances \( x = 4 \) is close to 6; precisely, we want to show that \( |x - 4| < \varepsilon \). Under what circumstances? We want this to be true whenever \( 0 < |x - 2| < \delta \). So the question becomes: can we choose a value for \( \delta \) that guarantees that \( |x - 2| < \delta \) if and only if \( |x - 1| < \delta / 2 \)?

**40 Chapter 2 Instantaneous Rate of Change: The Derivative**

We might start by computing \( \lim_{x \to 0} (x^2 - x) \). This is exactly what I did in the example: I picked \( \delta = 0.0000002 \) (or \( 0.000002 \)). A similar calculation again works for negative \( \Delta x \). The important fact is that this is now a completely general result— it shows that we can always win, no matter what "move" you make.

Now we can codify this by giving a precise definition to replace the fuzzy, "gets closer and closer" language we have used so far. Henceforward, we will say something like "the limit of \( (x^2 - x)/\Delta x \) as \( \Delta x \) goes to zero is \( 1/2 \)" and abbreviate this mouthful algebra:

\[
-19.6 - 4.9 \Delta x > -19.600001 \\
-4.9 \Delta x > -0.000001 \\
\Delta x < -0.000001 < -4.9 \\
\Delta x < 0.0000002408496327 \ldots
\]

Thus, we can say with certainty that if \( \Delta x \) is positive and less than 0.0000002, then \( \Delta x < 0.0000002408496327 \ldots \) and so \(-19.6 - 4.9 \Delta x > -19.600001 \). We could do a similar calculation if \( \Delta x \) is negative.

Now we know that we can make \(-19.6 - 4.9 \Delta x \) as close to one millionth of \(-19.6 \). But can we make it as close as we want? In this case, it is quite similar to see that the answer is yes, by modifying the calculation we've just done. It may be helpful to think of this as a game. I claim that I can make \(-19.6 - 4.9 \Delta x \) as close as you desire to \(-19.6 \) by making \( \Delta x \) "close enough" to zero. So the game is: you give me a number, like \( 10^{-6} \), and I have to come up with a number representing how close \( \Delta x \) can be to zero to guarantee that \(-19.6 - 4.9 \Delta x \) is at least close to \(-19.6 \) as you have requested.

Now if we actually play this game, I could redo the calculation above for each new number you provide. What I'd like to do is somehow see that I will always succeed, and even more, I'd like to have a simple strategy so that I don't have to do all that algebra every time. A strategy in this case would be a formula that gives me a correct answer no matter what you specify. So suppose the number you give me is \( \epsilon \). How close does \( \Delta x \) have to be to zero to guarantee that \(-19.6 - 4.9 \Delta x \) is in \((-19.6 - \epsilon, -19.6 + \epsilon)\)? If \( \Delta x \) is positive, we need:

\[
-19.6 - 4.9 \Delta x > -19.6 - \epsilon \\
-4.9 \Delta x > -\epsilon \\
\Delta x < -\epsilon / 4.9
\]

So if I pick any number \( \delta \) that is less than \( \epsilon/4.9 \), the algebra tells me that whenever \( \Delta x < \delta \) then \( \Delta x < -\epsilon/4.9 \) and so \(-19.6 - 4.9 \Delta x < -19.6 + \epsilon \). (This is exactly what I did in the example: I picked \( \delta = 0.0000002 \) (or \( 0.0000002408496327 \ldots \)). A similar calculation again works for negative \( \Delta x \). The important fact is that this is now a completely general result—it shows that I can always win, no matter what "move" you make.

Now we can codify this by giving a precise definition to replace the fuzzy, "gets closer and closer" language we have used so far. Henceforward, we will say something like "the limit of \( (x^2 - x)/\Delta x \) as \( \Delta x \) goes to zero is \( 1/2 \)" and abbreviate this mouthful
we don’t have to worry about it ever again. When we say that \( x \) might be “complicated” we really mean that in practice it might be a function. Here is then what we want to know:

**Theorem 2.3.5** Suppose \( \lim_{x \to a} f(x) = L \) and \( \lim_{x \to a} g(x) = M \). Then 
\[
\lim_{x \to a} (f(x)g(x)) = LM.
\]

**Proof.** We have to use the official definition of limit to make sense of this. So given any \( \varepsilon > 0 \) we need to find a \( \delta > 0 \) so that \( 0 < |x - a| < \delta \) implies \( |f(x)g(x) - LM| < \varepsilon \). What do we have to work with? We know that we can make \( f(x) \) close to \( L \) and \( g(x) \) close to \( M \), and we have to somehow connect these facts to make \( f(x)g(x) \) close to \( LM \).

We use, as is so often the case, a little algebraic trick:
\[
|f(x)g(x) - LM| = |f(x)(g(x) - M) + f(x)(M - L)| \leq |f(x)||g(x) - M| + |f(x)||M - L|.
\]

This is all straightforward except perhaps for the “\( \leq \)” . That is an example of the triangle inequality, which says that if \( a \) and \( b \) are any real numbers then \( |a + b| \leq |a| + |b| \). If you look at a few examples, using positive and negative numbers in various combinations for \( a \) and \( b \), you should quickly understand why this is true; we will not prove it formally.

Since \( \lim_{x \to a} f(x) = L \) there is a \( \delta_1 \) so that \( 0 < |x - a| < \delta_1 \) implies \( |f(x) - L| < \varepsilon/2M \). This means that \( 0 < |x - a| < \delta_1 \) implies \( |f(x) - L||M| < \varepsilon/2 \). You can see where this is going: if we can make \( |f(x)||M| < \varepsilon/2 \) or \( |f(x)| < \varepsilon/(2|M|) \), then we’re done. We can make \( |g(x) - M| \) smaller than any fixed number by making \( x \) close enough to \( a \); unfortunately, \( \varepsilon/(2|M|) \) is not a fixed number, since \( x \) is a variable. Here we need another little trick, just like the one we used in analyzing \( e^x \). We can find a \( \delta_2 \) so that \( |x - a| < \delta_2 \) implies that \( |f(x) - L| < \delta_1 \), meaning that \( L - \delta_1 < f(x) < L + \delta_1 \). This means that \( |f(x)| < N \), where \( N \) is either \( L - \delta_1 \) or \( L + \delta_1 \), depending on whether \( L \) is negative or positive. The important point is that \( N \) doesn’t depend on \( x \). Finally, we know that there is a \( \delta_3 \) so that \( 0 < |x - a| < \delta_3 \) implies \( |g(x) - M| < \varepsilon/(2|N|) \). Now we’re ready to put everything together. Let \( \delta \) be the smallest of \( \delta_1 \), \( \delta_2 \), and \( \delta_3 \). Then \( |x - a| < \delta \) implies that \( |f(x) - L|, |g(x) - M| < \varepsilon/(2|N|M) \). Then
\[
|f(x)g(x) - LM| \leq |f(x)||g(x) - M| + |f(x)||M - L| < \frac{\varepsilon}{2} \cdot \frac{\varepsilon}{2|M|} + \frac{\varepsilon}{2|N|M} = \varepsilon.
\]

This is just what we needed, so by the official definition, \( \lim_{x \to a} f(x)g(x) = LM \).

**Example 2.3.7** Compute \( \lim_{x \to 2} \frac{x^2 - 3x + 5}{x - 2} \).

If we apply the theorem in all its gory detail, we get
\[
\lim_{x \to 2} \frac{x^2 - 3x + 5}{x - 2} = \lim_{x \to 2} (x^2 - 3x + 5) - \lim_{x \to 2} (x - 2) = (2^2 - 3 \cdot 2 + 5) - (2 - 2) = 1.
\]

It is worth commenting on the trivial limit \( \lim \). From one point of view this might seem meaningless, as the number 5 can’t “approach” any value, since it is simply a fixed number. But 5 can, and should, be interpreted here as the height of the function that has value 5 everywhere, \( f(x) = 5 \), with graph a horizontal line. From this point of view it makes sense to ask what happens to the height of the function as \( x \) approaches 1.

Of course, as we’ve already seen, we’re primarily interested in limits that aren’t so easy, namely, limits in which a denominator approaches zero. There are a handful of algebraic tricks that work on many of these limits.

**Example 2.3.8** Compute \( \lim_{x \to 1} \frac{x^2 + 2x - 3}{x - 1} \). We can’t simply plug in \( x = 1 \) because that makes the denominator zero. However:
\[
\lim_{x \to 1} \frac{x^2 + 2x - 3}{x - 1} = \lim_{x \to 1} \frac{(x - 1)(x + 3)}{x - 1} = \lim_{x \to 1} (x + 3) = 4.
\]

While theorem 2.3.6 is very helpful, we need a bit more to work easily with limits. Since the theorem applies when some limits are already known, we need to know the behavior of some functions that cannot themselves be constructed from the simple arithmetic operations of the theorem, such as \( \sqrt{x} \). Also, there is one other extraordinarily useful way to put functions together: composition. If \( f(x) \) and \( g(x) \) are functions, we can form two functions by composition: \( f(g(x)) \) and \( g(f(x)) \). For example, if \( f(x) = \sqrt{x} \) and \( g(x) = x^2 + 5 \), then \( f(g(x)) = \sqrt{x^2 + 5} \) and \( g(f(x)) = (\sqrt{x})^2 + 5 = x + 5 \). Here is a companion to theorem 2.3.6 for composition.

**Theorem 2.3.9** Suppose that \( \lim_{x \to a} f(x) = L \) and \( \lim_{x \to a} g(x) = M \). Then 
\[
\lim_{x \to a} f(g(x)) = f(L).
\]

Note the special form of the condition on \( f \) : it is not enough to know that \( \lim_{x \to a} f(x) = M \), though it is a bit tricky to see why. Many of the most familiar functions do have this property, and this theorem can therefore be applied. For example:

**Theorem 2.3.10** Suppose that \( a \) is a positive integer. Then
\[
\lim_{x \to a} \sqrt[3]{x} = \sqrt[3]{a}.
\]

42 Chapter 2 Instantaneous Rate of Change: The Derivative

A handful of such theorems give us the tools to compute many limits without explicitly working with the definition of limit.

**Theorem 2.3.6** Suppose that \( \lim_{x \to a} f(x) = L \) and \( \lim_{x \to a} g(x) = M \) and \( k \) is some constant. Then
\[
\begin{align*}
\lim_{x \to a} f(x) + k &= \lim_{x \to a} f(x) + \lim_{x \to a} k = L + k \\
\lim_{x \to a} f(x)g(x) &= \lim_{x \to a} f(x) \cdot \lim_{x \to a} g(x) = LM \\
\lim_{x \to a} f(x) - g(x) &= \lim_{x \to a} f(x) - \lim_{x \to a} g(x) = L - M \\
\lim_{x \to a} f(x)^{g(x)} &= \lim_{x \to a} f(x)^{\lim_{x \to a} g(x)} = f(L)^M.
\end{align*}
\]

Roughly speaking, these rules say that to compute the limit of an algebraic expression, it is enough to compute the limits of the “innermost bits” and then combine those limits. This often means that it is possible to simply plug in a value for the variable, since \( \lim_{x \to a} f(x) = L \).

**Example 2.3.7** Compute \( \lim_{x \to 2} \frac{x^2 - 3x + 5}{x - 2} \).

If we apply the theorem in all its gory detail, we get
\[
\lim_{x \to 2} \frac{x^2 - 3x + 5}{x - 2} = \lim_{x \to 2} (x^2 - 3x + 5) - \lim_{x \to 2} (x - 2) = (2^2 - 3 \cdot 2 + 5) - (2 - 2) = 1.
\]

It is worth commenting on the trivial limit \( \lim \). From one point of view this might seem meaningless, as the number 5 can’t “approach” any value, since it is simply a fixed value.
\[ \lim_{x \to 0} 1 = 1. \] The limit of \( f(x) \) must be equal to both the left and right limits; since they are different, the limit \( \lim_{x \to 0} f(x) \) does not exist.

### Exercises 2.3.

Compute the limits. If a limit does not exist, explain why:

1. \( \lim_{x \to -1} (x^3 + x - 12) \)  
2. \( \lim_{x \to 2} (x^2 + x - 12) \)  
3. \( \lim_{x \to 3} x^3 + x - 12 \)  
4. \( \lim_{x \to 2} (x^2 + x - 12) \)  
5. \( \lim_{x \to 2} \sqrt{x^2 + 3} \)  
6. \( \lim_{x \to 2} \sqrt{7 + 2 - \sqrt{2}} \)  
7. \( \lim_{x \to 2} \sqrt{x^2 + 3} \)  
8. \( \lim_{x \to 3} \sqrt{x^2 + 3} - 5x \)  
9. \( \lim_{x \to 3} 4x - 5x^2 \)  
10. \( \lim_{x \to 2} \sqrt{x^2 - 1} \)  
11. \( \lim_{x \to 3} \sqrt{x^2 - 1} \)  
12. \( \lim_{x \to 1} \sqrt{x^2 + 3} \)  
13. \( \lim_{x \to 1} \frac{x^2 - x}{x - 1} \)  
14. \( \lim_{x \to 1} \frac{x^2 - 3x + 2}{x - 1} \)  
15. \( \lim_{x \to 1} \frac{x^2 - 3x}{x - 1} \)  
16. \( \lim_{x \to 1} \frac{x^2 - x}{x - 1} \)  

Give an epsilon-delta proof, similar to example 2.3.3, of the fact that \( \lim_{x \to 2} (2x - 5) = 3 \).

### 2.4 The Derivative Function

We know that \( f' \) carries important information about the original function \( f \). In one example we saw that \( f'(x) \) tells us how steep the graph of \( f(x) \) is; in another we saw that \( f'(x) \) tells us the velocity of an object if \( f(x) \) tells us the position of the object at time \( x \). As we said earlier, this same mathematical idea is useful whenever \( f(x) \) represents some changing quantity and we want to know something about how it changes, or roughly, the “rate” at which it changes. Most functions encountered in practice are built up from a small collection of “primitive” functions in a few simple ways, for example, by adding or multiplying functions together to get new, more complicated functions. To make good use of the information provided by \( f'(x) \) we need to be able to compute it for a variety of such functions.

We will begin to use different notations for the derivative of a function. While initially confusing, each is often useful so it is worth maintaining multiple versions of the same thing.

Consider again the function \( f(x) = \sqrt{25 - x^2} \). We have computed the derivative \( f'(x) = -\frac{x}{\sqrt{25 - x^2}} \), and have already noted that if we use the alternate notation \( y = \sqrt{25 - x^2} \) then we might write \( y' = -\frac{x}{\sqrt{25 - x^2}} \). Another notation is quite different, and in time it will become clear why it is often a useful one. Recall that to compute the derivative of \( f \) we computed

\[
\lim_{\Delta x \to 0} \frac{\sqrt{25 - (x + \Delta x)^2} - 24}{\Delta x}
\]

The denominator here measures a distance in the \( x \) direction, sometimes called the “run”, and the numerator measures a distance in the \( y \) direction, sometimes called the “rise,” and “rise over run” is the slope of a line. Recall that sometimes such a numerator is abbreviated \( \Delta y \), exchanging brevity for a more detailed expression. So in general, a derivative is given by

\[
y' = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}
\]

To recall the form of the limit, we sometimes say instead that

\[
\frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}
\]

In other words, \( dy/dx \) is another notation for the derivative, and it reminds us that it is related to an actual slope between two points. This notation is called Leibniz notation, after Gottfried Leibniz, who developed the fundamentals of calculus independently, at about the same time that Isaac Newton did. Again, since we often use \( f \) and \( f(x) \) to mean the original function, we sometimes use \( df/dx \) and \( df(x)/dx \) to refer to the derivative. If

### 2.4.1 Definition

The derivative of a function \( f \), denoted \( f' \), is

\[
f'(x) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}
\]

### 46 Chapter 2 Instantaneous Rate of Change: The Derivative

18. Evaluate the expressions by reference to this graph:

19. Use a calculator to estimate \( \lim_{x \to 0} \frac{\sin \pi x}{x} \).

20. Use a calculator to estimate \( \lim_{x \to 0} \frac{\sin(3x)}{x} \).

### 2.4.2 Example

Find the derivative of \( y = f(t) = t^2 \).

We compute

\[
y' = \lim_{\Delta t \to 0} \frac{\Delta y}{\Delta t} = \lim_{\Delta t \to 0} \frac{(t + \Delta t)^2 - t^2}{\Delta t} = \lim_{\Delta t \to 0} \frac{t^2 + 2t\Delta t + \Delta t^2 - t^2}{\Delta t} = \lim_{\Delta t \to 0} \left( \frac{2t\Delta t + \Delta t^2}{\Delta t} \right) = \lim_{\Delta t \to 0} (2t + \Delta t) = 2t
\]

Remember that \( \Delta t \) is a single quantity, not a “\( \Delta t \) times a “\( t \)”, and so \( \Delta t^2 \) is \( (\Delta t)^2 \) not \( \Delta t(\Delta t) \).

### 2.4.3 Example

Find the derivative of \( y = f(x) = 1/x \).

The computation:

\[
y' = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{\frac{1}{x + \Delta x} - \frac{1}{x}}{\Delta x} = \lim_{\Delta x \to 0} \frac{1}{x(x + \Delta x)} = \frac{1}{x^2}
\]

Note: If you happen to know some “derivative formulas” from an earlier course, for the time being you should pretend that you do not know them. In examples like the ones above and the exercises below, you are required to know how to find the derivative.
formula starting from basic principles. We will later develop some formulas so that we do not always need to do such computations, but we will continue to need to know how to do the more involved computations.

Sometimes one encounters a point in the domain of a function \( y = f(x) \) where there is no derivative, because there is no tangent line. In order for the notion of the tangent line at a point to make sense, the curve must be "smooth" at that point. This means that if you imagine a particle traveling at some steady speed along the curve, then the particle does not experience an abrupt change of direction. There are two types of situations you should be aware of—corners and cusps—where there's a sudden change of direction and hence no derivative.

**EXAMPLE 2.4.4** Discuss the derivative of the absolute value function \( y = f(x) = |x| \).

If \( x \) is positive, then this is the function \( y = x \), whose derivative is the constant 1. (Recall that when \( y = f(x) = mx + b \), the derivative is the slope \( m \).) If \( x \) is negative, then we're dealing with the function \( y = -x \), whose derivative is the constant \(-1\). If \( x = 0 \), then the function has a corner, i.e., there is no tangent line. A tangent line would have to point in the direction of the curve—but there are two directions of the curve that come together at the origin. We can summarize this as

\[
y' = \begin{cases} 
1 & \text{if } x > 0, \\
-1 & \text{if } x < 0, \\
\text{undefined} & \text{if } x = 0.
\end{cases}
\]

**EXAMPLE 2.4.5** Discuss the derivative of the function \( y = x^{2/3} \) shown in figure 2.4.1. We will later see how to compute this derivative; for now we use the fact that \( y' = (2/3)x^{-1/3} \). Usually this looks much like the absolute value function, but it technically has a cusp, not a corner. The absolute value function has no tangent line at 0 because there are (at least) two obvious contenders—the tangent line of the left side of the curve and the tangent line of the right side. The function \( y = x^{2/3} \) does not have a tangent line at 0, but unlike the absolute value function it can be said to have a single direction: as we approach 0 from either side the tangent line becomes closer and closer to a vertical line; the curve is vertical at 0. But as before, if you imagine traveling along the curve, an abrupt change in direction is required at 0; a full 180 degree turn.

In practice we won't worry much about the distinction between these examples; in both cases the function has a "sharp point" where there is no tangent line and no derivative.

### 2.5 Adjectives For Functions

As we have defined it in Section 1.3, a function is a very general object. At this point, it is useful to introduce a collection of adjectives to describe certain kinds of functions; these adjectives name useful properties that functions may have. Consider the graphs of the functions in Figure 2.5.1. It would clearly be useful to have words to help us describe the distinct features of each of them. We will point out and define a few adjectives (these are many more for the functions pictured here. For the sake of the discussion, we will assume that the graphs do not exhibit any unusual behavior off-stage (i.e., outside the view of the graphs).

**Functions.** Each graph in Figure 2.5.1 certainly represents a function—since each passes the vertical line test. In other words, as you sweep a vertical line across the graph of each function, the line never intersects the graph more than once. If it did, then the graph would not represent a function.

**Bounded.** The graph in (c) appears to approach zero as \( x \) goes to both positive and negative infinity. It also never exceeds the value 1 or drops below the value 0. Because the

**Continuity.** The graphs shown in (b) and (c) both represent continuous functions. Geometrically, this is because there are no jumps in the graphs. That is, if you pick a point on the graph and approach it from the left and right, the values of the function
DEFINITION 2.5.2 Continuous at a Point  A function $f$ is continuous at a point $a$ if \( \lim_{x \to a} f(x) = f(a) \).

DEFINITION 2.5.3 Continuous  A function $f$ is continuous if it is continuous at every point in its domain.

DEFINITION 2.5.4 Differentiable at a Point  A function $f$ is differentiable at a point $a$ if \( f'(a) \) exists.

DEFINITION 2.5.5 Differentiable  A function $f$ is differentiable if it is differentiable at every point (excluding endpoints and isolated points in the domain of $f$) in the domain of $f$.

We now have a collection of adjectives to describe the very rich and complex set of objects known as functions. We close with a useful theorem about continuous functions:

THEOREM 2.5.6 Intermediate Value Theorem  If $f$ is continuous on the interval $[a, b]$ and $d$ is between $f(a)$ and $f(b)$, then there is a number $c$ in $[a, b]$ such that $f(c) = d$.

This is most frequently used when $d = 0$.

EXAMPLE 2.5.7  Explain why the function $f(x) = x^3 + 3x^2 + x - 2$ has a root between 0 and 1.

By theorem 2.3.6, $f$ is continuous. Since $f(0) = -2$ and $f(1) = 3$, and 0 is between $-2$ and 3, there is a $c \in [0, 1]$ such that $f(c) = 0$.

This example also points the way to a simple method for approximating roots.

EXAMPLE 2.5.8  Approximate the root of the previous example to one decimal place.

If we compute $f(0.1)$, $f(0.2)$, and so on, we find that $f(0.6) < 0$ and $f(0.7) > 0$, so by the Intermediate Value Theorem, $f$ has a root between 0.6 and 0.7. Repeating the process with $f(0.61)$, $f(0.62)$, and so on, we find that $f(0.61) < 0$ and $f(0.62) > 0$, so $f$ has a root between 0.61 and 0.62, and the root is 0.6 rounded to one decimal place.

Exercises 2.5.

1. Along the lines of Figure 2.5.1, for each part below sketch the graph of a function that is:
   a. bounded, but not continuous.
   b. differentiable and unbounded.
   c. continuous at $x = 0$, not continuous at $x = 1$, and bounded.
   d. differentiable everywhere except at $x = -1$, continuous, and unbounded.

2. Is $f(x) = \sin(x)$ a bounded function? If so, find the smallest $M$.

3. Is $g(x) = 1/(1 + x^2)$ a bounded function? If so, find the smallest $M$.

4. Is $h(x) = 2\ln(x)$ a bounded function? If so, find the smallest $M$.

5. Consider the function
   \[ h(x) = \begin{cases} 
   2x - 3, & \text{if } x < 1 \\
   0, & \text{if } x \geq 1. 
   \end{cases} \]
   Show that it is continuous at the point $x = 0$. Is $h$ a continuous function?

6. Approximate a root of $f(x) = x^3 - 4x^2 + 2x + 2$ to one decimal place.

7. Approximate a root of $f(x) = x^3 - 3x^2 + 1$ to one decimal place.