2

Instantaneous Rate of Change: The Derivative

2.1 The slope of a function

Suppose that $y = f(x)$ is a function of $x$. It is often necessary to know how sensitive the value of $y$ is to small changes in $x$.

**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 $y = mx + A$, the slope $m = \Delta y/\Delta x$ 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 $\Delta x = 7 - 7 = 0.1$ is the change in $x$, and

$$\Delta y = (f(7 + \Delta x)) - f(7) = \sqrt{625 - (7 + 0.1)^2} - \sqrt{625 - 7^2} \approx -0.294.$$ 

Thus, $\Delta y/\Delta x \approx -0.294 / 0.1 = -2.94$. This means that $y$ changes by less than one third the change in $x$, so apparently $y$ is not very sensitive to changes in $x$ at $x = 7$.

We say “apparently” here because we don’t really know what happens between 7 and 7.1. Perhaps $y$ changes dramatically as $x$ runs through the values from 7 to 7.1, but at 7.1 $y$ just happens to be close to its value at 7. This is not in fact the case for this particular function, but we don’t yet know why.

**2.1 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

$$\frac{\sqrt{625 - (7 + \Delta x)^2} - 24}{\Delta x} = \frac{\sqrt{625 - (7 + \Delta x)^2} - \sqrt{625 - 7^2}}{\Delta x} = \frac{\Delta y}{\Delta x}.$$ 

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 zero? The numerator clearly gets very close to zero, so apparently

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \frac{\sqrt{625 - (7 + \Delta x)^2} - \sqrt{625 - 7^2}}{\Delta x} = \frac{\Delta y}{\Delta x} \approx -\frac{2x}{2\sqrt{625 - x^2}}.$$ 

Replacing $x$ by $7$ gives $-7/24$, as before, and now we can easily do the computation for 12 or any other value of $x$ between $-25$ and 25.

So now we have a simple, simple formula, $-x/\sqrt{625 - x^2}$, that tells us the slope of the tangent line for any value of $x$. This slope, in turn, tells us how sensitive the value of $y$ is to changes in the value of $x$.

What do we call such a formula? That is, a function with one variable, so that substituting an “input” value for the variable produces a new “output” value? This is a function. Starting with one function, $\sqrt{625 - x^2}$, we have derived, by means of some slightly nasty algebra, a new function, $-x/\sqrt{625 - x^2}$, that gives us important information about the original function. This new function in fact is called the derivative of the original function. If the original is referred to as $f$ or $y$ then the derivative is often written $f'(x)$ or $y'$ and pronounced “$f$ prime” or “$y$ prime”, so in this case we might write $f'(x) = -x/\sqrt{625 - x^2}$.

At a particular point, say $x = 7$, we say that $f'(7) = -7/24$ or “$f$ prime of $7$ is $-7/24$” or “the derivative of $f$ at 7 is $-7/24$.”

To summarize, we compute the derivative of $f(x)$ by forming the difference quotient $f(x + \Delta x) - f(x)$

$$\frac{f(x + \Delta x) - f(x)}{\Delta x}$$

which is the slope of a line, then we figure out what happens when $\Delta x$ gets very close to zero. This slope is not very sensitive to changes in $x$.

For example, if $x$ changes only from 7 to 7.01, then the difference quotient (slope of the chord) is approximately equal to $(25.97061 - 24)/0.01 \approx -0.294$. This is slightly less steep than the chord from $(7, 24)$ to $(7, 24.5)$.

As the second value $7 + \Delta x$ moves in towards 7, the chord joining $(7, f(7))$ to $(7 + \Delta x, f(7 + \Delta x))$ shifts slightly. As indicated in figure 2.1.1, as $\Delta x$ gets smaller and smaller, the chord joining $(7, 24)$ to $(7 + \Delta x, f(7 + \Delta x))$ gets closer and closer to the tangent line.

This slope, in turn, tells us how sensitive the value of $y$ is to changes in $x$. For example, if $x$ changes from 7 to 7.01, then the difference quotient (slope of the chord) is approximately equal to $(25.97061 - 24)/0.01 \approx -0.294$. This is slightly less steep than the chord from $(7, 24)$ to $(7, 24.5)$.
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 $-24/7$. We have shown that the tangent line is \(-24/7\)x + 24 = 0, or that it is \(-24/7\)x + 24 = 0. 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 \(x\) 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 \(f\) at the slope of the tangent line at \(x\) is the same everywhere. One curve that always has the same slope is a line; it seems odd to talk about the tangent line to a line, but if it makes sense at all, the tangent must be the line itself. It is not hard to see that the derivative of \(f(x) = mx + b\) is \(f'(x) = m\); see exercise 6.

Exercises 2.1.

1. Draw the graph of the function \(y = f(x) = \sqrt{625 - x^2}\) between \(x = 0\) and \(x = 13\). Find the slope \((\Delta y)/\Delta x\) of the chord between the points of the circle lying over \((a)\) \(x = 12\) and \(x = 13\), \((b)\) \(x = 12\) and \(x = 12.01\), \((c)\) \(x = 12\) and \(x = 12.001\). Now use algebra to find a simple formula for the slope of the chord between \((1, f(1))\) and \((3, f(3))\). Determine what happens when \(\Delta x\) approaches 0. In your graph of \(y = \sqrt{125 - x^2}\), draw the straight line through the point \((3, 13/3)\) whose slope is this limiting value of the difference quotient as \(\Delta x\) approaches 0.

2. Find an algebraic expression for the difference quotient \(f(1 + \Delta x) - f(1)/\Delta x\) when \(f(x) = x^2 - 3x + 4\). Simplify the expression as much as possible. Then determine what happens when \(\Delta x\) approaches 0. That value is \(f'(1)\).

3. Draw the graph of the function \(y = f(x) = (x^3) / (x^2 + 1)\) for \(0 \leq x \leq 2\). Draw a graph of the upper semicircle, and draw the tangent line at each of these four points.

4. Find an algebraic expression for the difference quotient \(f(x + \Delta x) - f(x)/\Delta x\) when \(f(x) = x^3\). Simplify the expression as much as possible. Then determine what happens when \(\Delta x\) approaches 0. That value is \(f'(x)\).

5. Find the slope of the chord \((3 + \Delta x, f(3 + \Delta x))\) of the function \(f(x) = x^2\) between \((a)\) \(x = 1\) and \(x = 1.001\), \((b)\) \(x = 1\) and \(x = 1.0001\), \((c)\) \(x = 1\) and \(x = 1.0000001\). Then use algebra to find a simple formula for the slope of the chord between \(1\) and \(1 + \Delta x\). (Use the expansion \((a + b)^2 = a^2 + 2ab + b^2\).) Determine what happens as \(\Delta x\) approaches 0, and in your graph of \(y = x^2\) draw the straight line through the point \((1, 1)\) whose slope is equal to the value you just found.

6. Find an algebraic expression for the difference quotient \(f(x + 2\Delta x) - f(x)/\Delta x\) when \(f(x) = mx + b\). Simplify the expression as much as possible. Then determine what happens as \(\Delta x\) approaches 0. That value is \(f'(x)\).

2.2 An example

A bit of algebra:

\[
\frac{h(2) - h(2 + \Delta t)}{\Delta t} = \frac{90.4 - (100 - 4.9\Delta t)^2}{\Delta t}
\]

When \(\Delta t\) is very small, this is very close to 19.6, and indeed it seems clear that as \(\Delta t\) goes to zero, the average speed goes to 19.6, so the exact speed at \(t = 2\) is 19.6 meters per second. This calculation should look very familiar. In the language of the previous section, we might have started with \(f(x) = 100 - 4.9x^2\) and asked for the slope of the tangent line at \(x = 2\). We would have answered that question by computing

\[
\frac{f(x + \Delta x) - f(x)}{\Delta x} = \frac{19.6\Delta t - 4.9\Delta x^2}{\Delta t} = -19.6 - 4.9\Delta x
\]

The algebra is the same, except that following the pattern of the previous section the subtraction would be reversed, and we would say that the slope of the tangent line is \(-19.6\). Indeed, in hindsight, perhaps we should have subtracted the other way even for the dropping ball. At \(t = 2\) the height was 80.0, and \(\Delta t\) goes to 55.9. The usual way to compute a “distance traveled” is to subtract the earlier position from the later one, or 55.9 - 80.4 = -24.5. This tells us that the distance traveled was 24.5 meters, and the negative sign tells us that the height went down during the second. If we continue the original calculation we then have \(-19.6\) meters per second as the exact speed at \(t = 2\). If we interpret the negative sign as meaning that the motion is downward, which seems reasonable, then in fact this is the same answer as before, but with even more information, since the numerical answer contains the direction of motion as well as the speed. Thus, the speed of the ball is the value of the derivative of a certain function, namely, of the function that gives the position of the ball. (More properly, this is the velocity of the ball; velocity is signed speed, that is, speed with a direction indicated by the sign.)

The upshot is that this problem, finding the speed of the ball, is exactly the same problem mathematically as finding the slope of a curve. This may already be enough evidence to convince you that whenever one says “the height of the curve,” or the height of a ball or the size of the economy or the distance of a space probe from earth or the population of the world) the rate at which the quantity changing can, in principle, be computed in exactly the same way, by finding a derivative.

Exercises 2.2.

1. An object is traveling in a straight line so that its position (that is, distance from some fixed point) is given by this table:

<table>
<thead>
<tr>
<th>time (seconds)</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>distance (meters)</td>
<td>0</td>
<td>10</td>
<td>25</td>
<td>60</td>
</tr>
</tbody>
</table>

Find the average speed of the object during the following time intervals: \([0, 1]\), \([0, 2]\), \([1, 2]\), \([2, 3]\). If you had to guess the speed at \(t = 2\) just on the basis of these, what would you guess?

2. Let \(y = f(t) = 2t^3\), 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 3 sec, \(b\) 2 sec and 2.1 sec, \(c\) 2 sec and 2.01 sec, \(d\) 2 sec and 2.001 sec. Then use algebra to find a simple formula for the average speed between time two and time two + \(\Delta t\). If you substitute \(\Delta t = 0.1, 0.01, 0.001\) in this formula you should again get the answers to parts \((a-d)\) above.

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 \(y = 2t^3\) draw the straight line through the point \((2, 16)\) whose slope is the instantaneous velocity you just computed; it should of course be the tangent line. \(t\)

3. An object is dropped from an 80-meter high window. Its height y above the ground at time \(t\) seconds is given by the formula \(y = f(t) = 80 - 4.9t^2\). (Here we are neglecting air resistance; the graph of this function was shown in figure 1.0.1.) Find the average velocity of the falling object between \((a)\) 1 sec and 1.1 sec, \((b)\) 1 sec and 1.01 sec, \((c)\) 1 sec and 1.001 sec. Now use algebra to find a simple formula for the average velocity of the falling object between 1 sec and 1 + \(\Delta t\) sec. Determine what happens to this average velocity as \(\Delta t\) approaches 0. That is the instantaneous velocity at time \(t = 1\) second; it will be negative, because the object is falling.

2.3 Limits

In the previous two sections we computed some quantities of interest (slope, velocity) by seeing that some expression “goes to” or “approaches” or “gets really close to” a particular value. In the examples we saw, this idea may have been clear enough, but it is too fuzzy to rely on in more difficult circumstances. In this section we will see how to make the idea more precise.

There is an important feature of the examples we have seen. Consider again the formula

\[-19.6\Delta t - 4.9\Delta t^2\]

We wanted to know what happens to this fraction as \(\Delta t\) goes to zero. Because we were able to simplify the fraction, it was easy to see the answer, but it was not quite as simple
as “substituting zero for $\Delta x$,” as that would give
\[
\begin{align*}
19.6 - 0.4 - 4.9 - 0.0 \\
0
\end{align*}
\]
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 $\sin(x)/x$ 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:
\[
\begin{align*}
\sqrt{2.00001} &= 1.414213566, \\
\sqrt{2.0000002} &= 1.414213566, \\
\sqrt{2.000000002} &= 1.414213566.
\end{align*}
\]
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, $\sqrt{2}$ is not even yet 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 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
\[
\begin{align*}
|19.6 - 0.4 - 9.4| &= 19.6 - 0.4 - 9.4.
\end{align*}
\]
These two quantities are equal as long as $\Delta x$ is not zero, if $\Delta x$ is zero, the left hand quantity is meaningless, while the right hand one is $-19.6$. Can we say more than we did before about why the right hand “approaches” $-19.6$ in the desired sense? Can we really make it “as close as we want” to $-19.6$? Let’s try a test case. We want to make $19.6 - 4.9\Delta x$ as close as we can to 0.

**DEFINITION 2.3.2 Limit** Suppose $f$ is a function. We say that $\lim_{x \to a} f(x) = L$ if for every $\epsilon > 0$ there is a $\delta > 0$ so that whenever $0 < |x - a| < \delta$,
\[
|f(x) - L| < \epsilon.
\]
The $\epsilon$ 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 $|f(x) - L| < \epsilon$ 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 $a$ for $x$ 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
\[
\begin{align*}
-19.6 - 4.9\Delta x.
\end{align*}
\]
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 is literally a fixed quantity, as we focused on the time $t$.) 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 not something we “need” to 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 $|x + 4 - 6| < \epsilon$, or $|x - 2| < \epsilon$. Under what circumstance? 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 $0 < |x - 2| < \delta$ implies $|x - 2| < \epsilon$? Of course: no matter what $\epsilon$ is, $\delta = \epsilon$ works.

So it turns out to be very easy to prove something “obvious,” which is nice. It doesn’t take long before things get trickier, however.

**EXAMPLE 2.3.4** It seems clear that $\lim_{x \to 2} x^2 = 4$. Let’s try to prove it. We will want to be able to show that $x^2 - 4 < \epsilon$ whenever $0 < |x - 2| < \delta$, by choosing $\delta$ carefully. Is there any connection between $|x - 2|$ and $x^2 - 4$? Yes, and it’s not hard to spot, but it is not so simple as the previous example. We can write $x^2 - 4 = (x + 2)(x - 2)$ or $x^2 - 4 = (x + 2)\cdot(x - 2)$, namely $|x - 2|$. We know $x$ is close to 2, $|x - 2|$ certainly can’t be too big, but we need to somehow be precise about it. Let’s recall the “game” version of what is going on here. You get to pick an $\epsilon$ and I have to prove that $|x - 2|$ makes things work out. Presumably it is the really tiny values of $\epsilon$ I need to worry about, but I have to be prepared for anything, even an apparently “bad” move like $\epsilon = 1000$. I expect that $|x - 2|$ is going to be small, and that the corresponding $\delta$ will be small, or nearly less than 1. If $\delta = 1$ then $|x - 2| < 5$ when $|x - 2| < \delta$ (because if $x$ is within 1 of 2, then $x$ is between 1 and 3 and $x + 2$ is between 3 and 5). Then I’d be trying to show that $|x + 2|/|x - 2| < 5$ and $|x - 2| < \epsilon$. So now how can I pick $\delta$ so that $|x - 2| < \epsilon$ implies $|x + 2|/|x - 2| < 5$? This is easy: use $\delta = \epsilon/5$, so $|x - 2| < 5(\epsilon/5) = \epsilon$. But if the $\epsilon$ we choose is not small? You can choose $\epsilon = 1000$, should I pick $\delta = 2000$? No, to keep things “same” I will never pick a $\delta$ bigger than 1. Here’s the final “game strategy.” When you pick a value for $\epsilon$ I will pick $\delta = \epsilon/5$ or $\delta = 1$, whichever is smaller. Now when $|x - 2| < \delta$, I know $|x + 2| > 5$ and $|x - 2| < 5$. Thus $|x + 2|/|x - 2| < 5(\epsilon/5) = \epsilon$.

This has been a long discussion, but most of it was explanation and scratch work. If this were written down as a proof, it would be quite short, like this:

Proof that $\lim_{x \to 2} x^2 = 4$. Given any $\epsilon$, pick $\delta = \epsilon/5$ or $\delta = 1$, whichever is smaller. Then when $|x - 2| < \delta$, $|x| < 5$ and $|x - 2| < \epsilon$. Hence $|x^2 - 4| = |x + 2||x - 2| < 5\epsilon/5 = \epsilon$.

It probably seems obvious that $\lim_{x \to 2} x^2$, and it is worth examining more closely why it seems obvious. If we write $x^2 = x \cdot x$, and ask what happens when $x$ approaches 2, we might say something like “Well, the first $x$ approaches 2, and the second $x$ approaches 2, so the product must approach 2.” This is in fact pretty much right on the money, except for that word “must.” It is really true if $x$ approaches a and $y$ approaches b then $xy$ approaches ab. It is, however, not really obvious, since $x$ and $y$ might be quite complicated. The good news is that we can see that this is true once and for all, and then
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 \( \epsilon > 0 \) we need to find a \( \delta > 0 \) so that \( 0 < |x - a| < \delta \) implies \( |f(x)g(x) - LM| < \epsilon \). 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:

\[
|f(x)g(x) - LM| = |(f(x) - L)(g(x) - M) + f(x)g(x) - Lg(x) + Lg(x) - LM| \\
\leq |(f(x) - L)(g(x) - M)| + |f(x)g(x) - Lg(x)| + |Lg(x) - LM| \\
\leq |f(x) - L||g(x) - M| + |f(x) - L||g(x)| + |L||g(x) - M|.
\]

This is all straightforward except perhaps for the "\( \epsilon \)". 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 \( \lim f(x) = L \) there is a \( \delta_1 \) so that \( 0 < |x - a| < \delta_1 \) implies \( |f(x) - L| < \epsilon/|2M| \). This means that \( 0 < |x - a| < \delta_1 \) implies \( |f(x)| < L + \epsilon/|2M| \). You can see where this is going: if we can make \( |f(x)|g(x) - Lg(x) < \epsilon/|2M| \) then \( |Lg(x) - LM| < \epsilon/|2M| \), so \( |f(x)g(x) - LM| < \epsilon \). Thus, if we can make \( |g(x)| > 0 \) (by choosing \( x \) to be large enough) we get \( |f(x)|g(x) - Lg(x) < \epsilon/|2M| \), which is true for all \( \epsilon > 0 \).

Next, we have another little trick, just like the one we used in analyzing \( x^2 \). We can find a \( \delta_2 \) so that \( |x - a| < \delta_2 \) implies \( |f(x) - L| < \frac{\epsilon}{2\sqrt{2}} \), meaning that \( L - \frac{\epsilon}{2\sqrt{2}} < f(x) < L + \frac{\epsilon}{2\sqrt{2}} \). This means that \( f(x) < L + \frac{\epsilon}{2\sqrt{2}} \). Now we're ready to put everything together. Let \( \delta = \min\{\delta_1, \delta_2\} \). Then \( |x - a| < \delta \) implies \( f(x) < L + \frac{\epsilon}{2\sqrt{2}} \) and \( |g(x) - M| < \epsilon/|2M| \). Thus

\[
|f(x)g(x) - LM| \leq |f(x)| |g(x) - M| + |f(x) - L||g(x)| + |L||g(x) - M| \\
\leq \left|L + \frac{\epsilon}{2\sqrt{2}}\right| \frac{\epsilon}{|2M|} + \frac{\epsilon}{2\sqrt{2}} \frac{\epsilon}{|2M|} + \frac{\epsilon}{|2M|} \frac{\epsilon}{|2M|} \\
\leq \frac{\epsilon}{2\sqrt{2}} \frac{\epsilon}{|2M|} + \frac{\epsilon}{2\sqrt{2}} \frac{\epsilon}{|2M|} + \frac{\epsilon}{|2M|} \frac{\epsilon}{|2M|} \\
= \frac{\epsilon}{2} + \frac{\epsilon}{2} \frac{\epsilon}{|2M|} + \frac{\epsilon}{|2M|} \frac{\epsilon}{|2M|}.
\]

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

### 2.3 Limits

number. But \( 5 \) can, and should, be interpreted here as the function that has value \( 5 \) everywhere, \( f(x) = 5 \), with a 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 arithmetical 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: \( (f \circ g)(x) \) and \( (g \circ f)(x) \). For example, if \( f(x) = \sqrt{x} \) and \( g(x) = x^2 + 5 \), then \( (f \circ g)(x) = \sqrt{x^2 + 5} \) and \( (g \circ f)(x) = (\sqrt{x})^2 + 5 = x + 5 \). Here is a companion to theorem 2.3.6 for composition:

**Theorem 2.3.9** Suppose that \( \lim_{x \to a} g(x) = L \) and \( \lim_{x \to a} f(x) = f(L) \). 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 \( n \) is a positive integer. Then

\[
\lim_{x \to a} \sqrt[n]{x} = \sqrt[n]{a},
\]

provided that \( a \) is positive if \( n \) is even.

### 4.4 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

\[
\lim_{x \to a} kf(x) = k \lim_{x \to a} f(x) = kL,
\]

\[
\lim_{x \to a} (f(x) + g(x)) = \lim_{x \to a} f(x) + \lim_{x \to a} g(x) = L + M,
\]

\[
\lim_{x \to a} (f(x) - g(x)) = \lim_{x \to a} f(x) - \lim_{x \to a} g(x) = L - M.
\]

If \( M \neq 0 \) then

\[
\lim_{x \to a} \frac{f(x)}{g(x)} = \frac{\lim_{x \to a} f(x)}{\lim_{x \to a} g(x)} = \frac{L}{M}.
\]

Roughly speaking, these rules say that to compute the limit of an algebraic expression, it is enough to compute the limits of the "smallest parts" and then combine those limits. This often means that it is possible to simply plug in a value for the variable, since \( \lim x \to a \).

**Example 2.3.7** Compute \( \lim_{x \to 1} \frac{x^2 - 3x + 5}{x - 2} \). If we apply the theorem in all its glory detail, we get

\[
\lim_{x \to 1} \frac{x^2 - 3x + 5}{x - 2} = \lim_{x \to 1} \frac{x^2 - 3x + 5}{x - 2} = \lim_{x \to 1} \frac{x^2 - 3}{x - 2} = \frac{1^2 - 3}{1 - 2} = -1
\]

It is worth commenting on the trivial limit \( \lim x \to a \). From one point of view this might seem meaningless, as the number 5 can’t “approach” any value, since it is simply a fixed
2.3 Limits

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 one 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 \( f'(x) = -x/\sqrt{25-x^2} \), and have already noted that if we use the alternate notation \( y = \sqrt{25 - x^2} \) then we might write \( y' = -x/\sqrt{25 - x^2} \). 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

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

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

\[
\lim_{x \to 0} \frac{f(x) - f(0)}{x - 0} = 1.
\]

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.

2.4 The Derivative Function

We have seen how to create, or derive, a new function \( f'(x) \) from a function \( f(x) \), summarized in the paragraph containing equation 2.1.1. Now that we have the concept of limits, we can make this more precise.

**Definition 2.4.1** The derivative of a function \( f \), denoted \( f' \), is

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

50 Chapter 2 Instantaneous Rate of Change: The Derivative

The function \( f(x) \) is written out in full we often write the last of these something like this

\[
f'(x) = \frac{df}{dx} \quad \text{with the function written to the side, instead of trying to fit it into the numerator.}
\]

**Example 2.4.2** Find the derivative of \( y = f(t) = t^2 \).

We compute

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

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

**Example 2.4.3** Find the derivative of \( y = f(z) = 1/z \).

The computation:

\[
y' = \lim_{\Delta z \to 0} \frac{f(z + \Delta z) - f(z)}{\Delta z} = \lim_{\Delta z \to 0} \frac{1/(z + \Delta z) - 1/z}{\Delta z}
\]

\[
= \lim_{\Delta z \to 0} \frac{1}{z(z + \Delta z)} \frac{z - (z + \Delta z)}{\Delta z} = \lim_{\Delta z \to 0} \frac{-\Delta z}{z(z + \Delta z)} = \lim_{\Delta z \to 0} \frac{-1}{z(z + \Delta z)} = -\frac{1}{z^2}.
\]

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 \( y' = (2/3)x^{-1/3} \). Visually 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 (there 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
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 $f$.

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 $f = x^3 + 3x^2 + x - 2$ 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 $f = x^3 - 4x^2 + 2x + 1$ to one decimal place.
7. Approximate a root of $f = x^3 + x^2 - 5x + 1$ to one decimal place.