Sometimes we can approximate complicated functions with simpler ones that give the accuracy we want for specific applications and are easier to work with. The approximating functions discussed in this section are called linearizations, and they are based on tangent lines. Other approximating functions, such as polynomials, are discussed in Chapter 11.
We introduce new variables \( dx \) and \( dy \), called \textit{differentials}, and define them in a way that makes Leibniz’s notation for the derivative \( dy/dx \) a true ratio. We use \( dy \) to estimate error in measurement and sensitivity of a function to change. Application of these ideas then provides for a precise proof of the Chain Rule (Section 3.5).

**Linearization**

As you can see in Figure 3.46, the tangent to the curve \( y = x^2 \) lies close to the curve near the point of tangency. For a brief interval to either side, the \( y \)-values along the tangent line give good approximations to the \( y \)-values on the curve. We observe this phenomenon by zooming in on the two graphs at the point of tangency or by looking at tables of values for the difference between \( f(x) \) and its tangent line near the \( x \)-coordinate of the point of tangency. Locally, every differentiable curve behaves like a straight line.

For as long as this line remains close to the graph of \( f \), \( L(x) \) gives a good approximation to \( f(x) \).
EXAMPLE 1 Finding a Linearization

Find the linearization of \( f(x) = \sqrt{1 + x} \) at \( x = 0 \) (Figure 3.48).

**Solution** Since

\[
f'(x) = \frac{1}{2} (1 + x)^{-1/2},
\]

we have \( f(0) = 1 \) and \( f'(0) = 1/2 \), giving the linearization

\[
L(x) = f(a) + f'(a)(x - a) = 1 + \frac{1}{2} (x - 0) = 1 + \frac{x}{2}.
\]

See Figure 3.49.

Look at how accurate the approximation \( \sqrt{1 + x} \approx 1 + (x/2) \) from Example 1 is for values of \( x \) near 0.

As we move away from zero, we lose accuracy. For example, for \( x = 2 \), the linearization gives 2 as the approximation for \( \sqrt{3} \), which is not even accurate to one decimal place.

Do not be misled by the preceding calculations into thinking that whatever we do with a linearization is better done with a calculator. In practice, we would never use a linearization to find a particular square root. The utility of a linearization is its ability to replace a complicated formula by a simpler one over an entire interval of values. If we have to work with \( \sqrt{1 + x} \) for \( x \) close to 0 and can tolerate the small amount of error involved, we can...
work with instead. Of course, we then need to know how much error there is. We have more to say on the estimation of error in Chapter 11.

A linear approximation normally loses accuracy away from its center. As Figure 3.48 suggests, the approximation will probably be too crude to be useful near There, we need the linearization at

EXAMPLE 2 Finding a Linearization at Another Point

Find the linearization of at

Solution

We evaluate the equation defining With we have

At the linearization in Example 2 gives

which differs from the true value by less than one one-thousandth. The linearization in Example 1 gives

a result that is off by more than 25%.

EXAMPLE 3 Finding a Linearization for the Cosine Function

Find the linearization of at (Figure 3.50).

Solution

Since \( f(\pi/2) = \cos(\pi/2) = 0 \), \( f'(x) = -\sin x \), and \( f'(\pi/2) = -\sin(\pi/2) = -1 \), we have

\[
L(x) = f(a) + f'(a)(x - a) \\
= 0 + (-1)(x - \pi/2) \\
= -x + \pi/2.
\]
An important linear approximation for roots and powers is
\[(1 + x)^k \approx 1 + kx \quad (x \text{ near 0; any number } k)\]
(Exercise 15). This approximation, good for values of \(x\) sufficiently close to zero, has broad application. For example, when \(x\) is small,
\[
\sqrt{1 + x} \approx 1 + \frac{1}{2} x
\]
\[
\frac{1}{1 - x} = (1 - x)^{-1} \approx 1 + (1)(-x) = 1 + x
\]
\[
\sqrt{1 + 5x^4} = (1 + 5x^4)^{1/3} \approx 1 + \frac{1}{3} \left(5x^4\right) = 1 + \frac{5}{3}x^4
\]
\[
\frac{1}{\sqrt{1 - x^2}} = (1 - x^2)^{-1/2} \approx 1 + \left(-\frac{1}{2}\right)(-x^2) = 1 + \frac{1}{2}x^2
\]

**Differentials**

We sometimes use the Leibniz notation \(dy/dx\) to represent the derivative of \(y\) with respect to \(x\). Contrary to its appearance, it is not a ratio. We now introduce two new variables \(dx\) and \(dy\) with the property that if their ratio exists, it will be equal to the derivative.

**DEFINITION** Differential

Let \(y = f(x)\) be a differentiable function. The **differential** \(dx\) is an independent variable. The **differential** \(dy\) is
\[
dy = f'(x) \, dx.
\]

Unlike the independent variable \(dx\), the variable \(dy\) is always a dependent variable. It depends on both \(x\) and \(dx\). If \(dx\) is given a specific value and \(x\) is a particular number in the domain of the function \(f\), then the numerical value of \(dy\) is determined.

**EXAMPLE 4** Finding the Differential \(dy\)

(a) Find \(dy\) if \(y = x^5 + 37x\).

(b) Find the value of \(dy\) when \(x = 1\) and \(dx = 0.2\).

**Solution**

(a) \(dy = (5x^4 + 37) \, dx\)

(b) Substituting \(x = 1\) and \(dx = 0.2\) in the expression for \(dy\), we have
\[
dy = (5 \cdot 1^4 + 37)0.2 = 8.4.
\]

The geometric meaning of differentials is shown in Figure 3.51. Let \(x = a\) and set \(dx = \Delta x\). The corresponding change in \(y = f(x)\) is
\[
\Delta y = f(a + dx) - f(a).
\]
The corresponding change in the tangent line $L$ is

$$
\Delta L = L(a + dx) - L(a) = f(a) + f'(a)(a + dx) - a - f(a) = f'(a)dx.
$$

That is, the change in the linearization of $f$ is precisely the value of the differential $dy$ when $x = a$ changes by an amount $dx = \Delta x$. Therefore, $dy$ represents the amount the tangent line rises or falls when $x$ changes by an amount $dx = \Delta x$.

If $dx \neq 0$, then the quotient of the differential $dy$ by the differential $dx$ is equal to the derivative $f'(x)$ because

$$
dy / dx = \frac{f'(x) dx}{dx} = f'(x) = \frac{dy}{dx}.
$$

We sometimes write

$$
df = f'(x) dx
$$

in place of $dy = f'(x) dx$, calling $df$ the **differential of $f$**. For instance, if $f(x) = 3x^2 - 6$, then

$$
df = d(3x^2 - 6) = 6x \, dx.
$$

Every differentiation formula like

$$
\frac{d(u + v)}{dx} = \frac{du}{dx} + \frac{dv}{dx}
$$

or

$$
\frac{d(sin u)}{dx} = \cos u \frac{du}{dx}
$$

has a corresponding differential form like

$$
(\sin u) = \cos u \, du
$$

or

$$
d(u + v) = du + dv
$$

or

$$
d(sin u) = \cos u \, du
.$$
EXAMPLE 5  Finding Differentials of Functions

(a) \( d(\tan 2x) = \sec^2(2x) \, d(2x) = 2 \sec^2 2x \, dx \)

(b) \( d\left(\frac{x}{x + 1}\right) = \frac{(x + 1) \, dx - x \, d(x + 1)}{(x + 1)^2} = \frac{x \, dx + dx - x \, dx}{(x + 1)^2} = \frac{dx}{(x + 1)^2} \)

Estimating with Differentials

Suppose we know the value of a differentiable function \( f(x) \) at a point \( a \) and want to predict how much this value will change if we move to a nearby point \( a + dx \). If \( dx \) is small, then we can see from Figure 3.51 that \( \Delta y \) is approximately equal to the differential \( dy \).

Since the differential approximation gives

\[ f(a + dx) = f(a) + \Delta y, \]

the differential approximation gives

\[ f(a + dx) \approx f(a) + dy \]

where \( dx = \Delta x \). Thus the approximation \( \Delta y \approx dy \) can be used to calculate \( f(a + dx) \) when \( f(a) \) is known and \( dx \) is small.

EXAMPLE 6  Estimating with Differentials

The radius \( r \) of a circle increases from \( a = 10 \) m to 10.1 m (Figure 3.52). Use \( dA \) to estimate the increase in the circle’s area \( A \). Estimate the area of the enlarged circle and compare your estimate to the true area.

Solution  Since \( A = \pi r^2 \), the estimated increase is

\[ dA = A'(a) \, dr = 2\pi a \, dr = 2\pi(10)(0.1) = 2\pi \, m^2. \]

Thus,

\[ A(10 + 0.1) \approx A(10) + 2\pi \]

\[ = \pi(10)^2 + 2\pi = 102\pi. \]

The area of a circle of radius 10.1 m is approximately 102\( \pi \) m\(^2\).

The true area is

\[ A(10.1) = \pi(10.1)^2 \]

\[ = 102.01\pi \, m^2. \]

The error in our estimate is 0.01\( \pi \) m\(^2\), which is the difference \( \Delta A - dA \).

Error in Differential Approximation

Let \( f(x) \) be differentiable at \( x = a \) and suppose that \( dx = \Delta x \) is an increment of \( x \). We have two ways to describe the change in \( f \) as \( x \) changes from \( a \) to \( a + \Delta x \):

The true change: \( \Delta f = f(a + \Delta x) - f(a) \)

The differential estimate: \( df = f'(a) \Delta x \).

How well does \( df \) approximate \( \Delta f \)?
We measure the approximation error by subtracting \( df \) from \( \Delta f \):

\[
\text{Approximation error} = \Delta f - df = \Delta f - f'(a) \Delta x
\]

\[
= \frac{f(a + \Delta x) - f(a) - f'(a) \Delta x}{\Delta x}
\]

\[
= \left( \frac{f(a + \Delta x) - f(a)}{\Delta x} - f'(a) \right) \cdot \Delta x
\]

Call this part \( \epsilon \)

\[
= \epsilon \cdot \Delta x.
\]

As \( \Delta x \to 0 \), the difference quotient

\[
\frac{f(a + \Delta x) - f(a)}{\Delta x}
\]

approaches \( f'(a) \) (remember the definition of \( f'(a) \)), so the quantity in parentheses becomes a very small number (which is why we called it \( \epsilon \)). In fact, \( \epsilon \to 0 \) as \( \Delta x \to 0 \). When \( \Delta x \) is small, the approximation error \( \epsilon \cdot \Delta x \) is smaller still.

\[
\Delta f = f'(a) \Delta x + \epsilon \Delta x
\]

true change estimated change error

Although we do not know exactly how small the error is and will not be able to make much progress on this front until Chapter 11, there is something worth noting here, namely the form taken by the equation.

**Change in \( y = f(x) \) near \( x = a \)**

If \( y = f(x) \) is differentiable at \( x = a \) and \( x \) changes from \( a \) to \( a + \Delta x \), the change \( \Delta y \) in \( f \) is given by an equation of the form

\[
\Delta y = f'(a) \Delta x + \epsilon \Delta x
\]

in which \( \epsilon \to 0 \) as \( \Delta x \to 0 \).

In Example 6 we found that

\[
\Delta A = \pi(10.1)^2 - \pi(10)^2 = (102.01 - 100)\pi = (2\pi + 0.01\pi) \text{ m}^2
\]

so the approximation error is \( \Delta A - dA = \epsilon \Delta r = 0.01\pi \) and \( \epsilon = 0.01\pi/\Delta r = 0.01\pi/0.1 = 0.1\pi \text{ m} \).

Equation (1) enables us to bring the proof of the Chain Rule to a successful conclusion.

**Proof of the Chain Rule**

Our goal is to show that if \( f(u) \) is a differentiable function of \( u \) and \( u = g(x) \) is a differentiable function of \( x \), then the composite \( y = f(g(x)) \) is a differentiable function of \( x \).
More precisely, if \( g \) is differentiable at \( x_0 \) and \( f \) is differentiable at \( g(x_0) \), then the composite is differentiable at \( x_0 \) and

\[
\frac{dy}{dx} \bigg|_{x=x_0} = f'(g(x_0)) \cdot g'(x_0).
\]

Let \( \Delta x \) be an increment in \( x \) and let \( \Delta u \) and \( \Delta y \) be the corresponding increments in \( u \) and \( y \). Applying Equation (1) we have,

\[
\Delta u = g'(x_0) \Delta x + \epsilon_1 \Delta x = (g'(x_0) + \epsilon_1) \Delta x,
\]

where \( \epsilon_1 \to 0 \) as \( \Delta x \to 0 \). Similarly,

\[
\Delta y = f'(u_0) \Delta u + \epsilon_2 \Delta u = (f'(u_0) + \epsilon_2) \Delta u,
\]

where \( \epsilon_2 \to 0 \) as \( \Delta u \to 0 \). Notice also that \( \Delta u \to 0 \) as \( \Delta x \to 0 \). Combining the equations for \( \Delta u \) and \( \Delta y \) gives

\[
\Delta y = (f'(u_0) + \epsilon_2)(g'(x_0) + \epsilon_1) \Delta x,
\]

so

\[
\frac{\Delta y}{\Delta x} = f'(u_0)g'(x_0) + \epsilon_2 g'(x_0) + f'(u_0)\epsilon_1 + \epsilon_2 \epsilon_1.
\]

Since \( \epsilon_1 \) and \( \epsilon_2 \) go to zero as \( \Delta x \) goes to zero, three of the four terms on the right vanish in the limit, leaving

\[
\frac{dy}{dx} \bigg|_{x=x_0} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = f'(u_0)g'(x_0) = f'(g(x_0)) \cdot g'(x_0).
\]

This concludes the proof. \( \blacksquare \)

**Sensitivity to Change**

The equation \( df = f'(x) \, dx \) tells how sensitive the output of \( f \) is to a change in input at different values of \( x \). The larger the value of \( f' \) at \( x \), the greater the effect of a given change \( dx \). As we move from \( a \) to a nearby point \( a + dx \), we can describe the change in \( f \) in three ways:

<table>
<thead>
<tr>
<th>True</th>
<th>Estimated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absolute change</td>
<td>( \Delta f = f(a + dx) - f(a) )</td>
</tr>
<tr>
<td>Relative change</td>
<td>( \frac{\Delta f}{f(a)} )</td>
</tr>
<tr>
<td>Percentage change</td>
<td>( \frac{\Delta f}{f(a)} \times 100 )</td>
</tr>
</tbody>
</table>

**EXAMPLE 7** Finding the Depth of a Well

You want to calculate the depth of a well from the equation \( s = 16t^2 \) by timing how long it takes a heavy stone you drop to splash into the water below. How sensitive will your calculations be to a 0.1-sec error in measuring the time?

**Solution** The size of \( ds \) in the equation

\[
ds = 32t \, dt
\]
depends on how big \( t \) is. If \( t = 2 \) sec, the change caused by \( \Delta t = 0.1 \) is about
\[
\Delta s = 32(2)(0.1) = 6.4 \text{ ft}.
\]
Three seconds later at \( t = 5 \) sec, the change caused by the same \( \Delta t \) is
\[
\Delta s = 32(5)(0.1) = 16 \text{ ft}.
\]
The estimated depth of the well differs from its true depth by a greater distance the longer
the time it takes the stone to splash into the water below, for a given error in measuring the
time.

**EXAMPLE 8  Unclogging Arteries**

In the late 1830s, French physiologist Jean Poiseuille ("pwa-ZOY") discovered the for-

mula we use today to predict how much the radius of a partially clogged artery has to
be expanded to restore normal flow. His formula,
\[
V = kr^4,
\]
says that the volume \( V \) of fluid flowing through a small pipe or tube in a unit of time at a
fixed pressure is a constant times the fourth power of the tube’s radius \( r \). How will a 10% 
increase in \( r \) affect \( V \)?

**Solution**  The differentials of \( r \) and \( V \) are related by the equation
\[
dV = \frac{dV}{dr} dr = 4kr^3 dr.
\]
The relative change in \( V \) is
\[
\frac{dV}{V} = \frac{4kr^3 dr}{kr^4} = 4 \frac{dr}{r}.
\]
The relative change in \( V \) is 4 times the relative change in \( r \), so a 10% increase in \( r \) will pro-
duce a 40% increase in the flow.

**EXAMPLE 9  Converting Mass to Energy**

Newton’s second law,
\[
F = \frac{d}{dt} (mv) = m \frac{dv}{dt} = ma,
\]
is stated with the assumption that mass is constant, but we know this is not strictly true be-
cause the mass of a body increases with velocity. In Einstein’s corrected formula, mass has
the value
\[
m = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}},
\]
where the “rest mass” \( m_0 \) represents the mass of a body that is not moving and \( c \) is the
speed of light, which is about 300,000 km/sec. Use the approximation
\[
\frac{1}{\sqrt{1 - x^2}} \approx 1 + \frac{1}{2} x^2 \quad (2)
\]
to estimate the increase \( \Delta m \) in mass resulting from the added velocity \( v \).
Solution When $v$ is very small compared with $c$, $v^2/c^2$ is close to zero and it is safe to
use the approximation

$$
\frac{1}{\sqrt{1 - v^2/c^2}} \approx 1 + \frac{1}{2} \left( \frac{v^2}{c^2} \right)
$$

Eq. (2) with $x = v/c$

to obtain

$$
m = \frac{m_0}{\sqrt{1 - v^2/c^2}} \approx m_0 \left[ 1 + \frac{1}{2} \left( \frac{v^2}{c^2} \right) \right] = m_0 + \frac{1}{2} m_0 v^2 \left( \frac{1}{c^2} \right),
$$
or

$$
m \approx m_0 + \frac{1}{2} m_0 v^2 \left( \frac{1}{c^2} \right).
$$

Equation (3) expresses the increase in mass that results from the added velocity $v$.

Energy Interpretation
In Newtonian physics, $(1/2)m_0 v^2$ is the kinetic energy (KE) of the body, and if we rewrite
Equation (3) in the form

$$
(m - m_0)c^2 \approx \frac{1}{2} m_0 v^2,
$$

we see that

$$
(m - m_0)c^2 \approx \frac{1}{2} m_0 v^2 = \frac{1}{2} m_0 v^2 - \frac{1}{2} m_0(0)^2 = \Delta(KE),
$$
or

$$
(\Delta m)c^2 \approx \Delta(KE).
$$

So the change in kinetic energy $\Delta(KE)$ in going from velocity 0 to velocity $v$ is approximately equal to $(\Delta m)c^2$, the change in mass times the square of the speed of light. Using $c \approx 3 \times 10^8$ m/sec, we see that a small change in mass can create a large change in energy.