The Derivative

1 Tangent Lines and the Derivative

Let's start by describing what we will call the *tangent line to a curve.* If we have a function $f(x)$ and a point x_0 in its domain, then the **tangent line** of f at x_0 is the line passing through the point $(x_0, f(x_0))$ with the property that in a small enough neighborhood around $(x_0, f(x_0))$, the function and the the line are indistinguishable. That is, the graph of the function will look like the tangent line if you zoom in to a small enough interval around the point. The slope of the tangent line to f at $(x_0, f(x_0))$ is also called the **slope of the function** $f(x)$ at x_0.

It is important to realize that there are points on graphs of certain functions for which no tangent line exists. For example, the graph of the function $f(x) = |x|$ in any open interval around $x = 0$ will always look like a 'V' and so it could not have a tangent line at $x = 0$.

We need a way to compute the slope of the tangent line to a function at a given point. Remember that in order to compute the slope of a line, we need to know two distinct points which lie on the line. We will start by computing the slope of a *secant line* (a line going through two points on the graph of $f(x)$). Suppose x_0 is a point in the domain D of $f(x)$ and h is a non-zero real number (we want h to be close to zero, though it can be either positive or negative). We will approximate the tangent line to $f(x)$ at $(x_0, f(x_0))$ using the secant line passing through $(x_0, f(x_0))$ and $(x_0 + h, f(x_0 + h))$. The slope of this secant line is

$$m(x_0, h) = \frac{f(x_0 + h) - f(x_0)}{(x_0 + h) - x_0} = \frac{f(x_0 + h) - f(x_0)}{h}.$$

The notation $m(x_0, h)$ means that the slope of the secant line depends on both x_0 and h. Furthermore, if we fix x_0, then $m(x_0, h)$ is a function of h. The limit concept provides us with a tool which allows us to define the slope of the tangent line to a function at a point.

Definition 1.1 *Let $f(x)$ be a function and let x_0 be a point in its domain. The **slope of the tangent line to $f(x)$ at the point** $(x_0, f(x_0))$ is the real number given by*

$$m(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

whenever this limit exists. The limit does not exist if and only if there is not tangent line to that function at $(x_0, f(x_0))$.

Note that we get the second expression for $m(x_0)$ when we take $h = x - x_0$ in the first expression and we interpret h to be a 'signed distance' from x to x_0 (meaning h takes on positive and negative values depending on whether x lies to the right or to the left of x_0).

Given a function f, we will derive a function $f'(x)$ from $f(x)$ which takes each point in the domain of $f(x)$ for which $f(x)$ has a tangent line to the slope of that tangent line.

Definition 1.2 *Let $f(x)$ be a function with domain D and let D' be the set of all points in D at which $f(x)$ has a tangent line. The* ***derivative of*** *f is the function $f'(x)$ with domain D' given by*

$$f'(x) = m(x) = \lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h} = \lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0}$$

which takes a point x_0 in D' to the slope of the tangent line to $f(x)$ at $(x_0, f(x_0))$. We call $f'(x)$ the ***derivative of*** *f* ***at*** *x and we say that f is* ***differentiable on*** *D'.*

Let's see if we can compute the slope of the tangent line of constant, linear, and quadratic functions at any point on them. To do this, consider the function $f(x) = ax^2 + bx + c$, where a, b, and c are any real numbers. Fix a real number x_0. We compute

$$\begin{aligned} f(x_0+h) &= a(x_0+h)^2 + b(x_0+h) + c \\ &= ax_0^2 + 2ahx_0 + ah^2 + bx_0 + bh + c \\ &= (ax_0^2 + bx_0 + c) + 2ahx_0 + ah^2 + bh \\ &= f(x_0) + 2ahx_0 + ah^2 + bh. \end{aligned}$$

Consequently,

$$\begin{aligned} f'(x) &= \lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h} \\ &= \lim_{h\to 0} \frac{2ahx_0 + ah^2 + bh}{h} \\ &= \lim_{h\to 0} \frac{h(2ax_0 + ah + b)}{h} \\ &= \lim_{h\to 0} (2ax_0 + ah + b) = 2ax_0 + a\cdot 0 + b = 2ax_0 + b \end{aligned}$$

by continuity of the polynomial $ah + (2ax_0 + b)$.

Note that if $a = b = 0$, then $f(x) = c$ (i.e., f is a constant function). In this case, $f'(x_0) = 0$; that is, *the slope of a constant function at any point is* 0. This makes sense since the graph of $f(x)$ is a horizontal line (so the slope is 0) and the tangent line should correspond exactly to the graph of the function. If $a = 0$, then we have the linear function $f(x) = bx + c$ whose graph is a line with slope b. As we hope, $f'(x) = b$.

The computation of the derivative of $f(x) = ax^2 + bx + c$ shows that $f(x)$ is differentiable at every real number, and the slope of the tangent line to $f(x)$ at any point x_0 is $f'(x_0) = 2ax_0 + b$.

One of the simplest examples of a function which is not differentiable at every point of its domain is the absolute value function defined by

$$g(x) = |x| = \begin{cases} x & \text{if } x \geq 0 \\ -x & \text{if } x < 0 \end{cases}.$$

We claim that g is differentiable at every real number except $x = 0$. If $x > 0$, then $g(x) = x$ and so $g'(x) = 1$. If $x < 0$, then $g(x) = -x$ and therefore $g'(x) = -1$. Thus, $g(x)$ is differentiable at any positive or negative point. To show that $g(x)$ is not differentiable at 0, consider the limit:

$$\lim_{x\to 0} \frac{g(x)-g(0)}{x-0} = \lim_{x\to 0} \frac{|x|-|0|}{x} = \lim_{x\to 0} \frac{|x|}{x} = \lim_{x\to 0} \begin{cases} 1 & \text{if } x > 0 \\ -1 & \text{if } x < 0 \end{cases}$$

which does not exist. Recall that the graph of $g(x) = |x|$ looks like a 'V' with its sharp vertex at the point $(0, 0)$. In general, a function is not differentiable at 'sharp' points like this on its graph.

1.1 Notation

We will introduce three different notations associated with the derivative. If $f(x)$ is a function, then we write $f'(x)$ for its derivative. If we have an *equation* $y = f(x)$, then we use the notation

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

which is read "the derivative of y with respect to x." Note that y is the dependent variable and x is the independent variable. Sometimes we will use x as the independent variable (representing horizontal displacement) and t as the independent variable (representing time) and so we would write $\frac{dx}{dt}$ in this case.

A third notation for the derivative comes when we think of the operation of taking the derivative as a BIG function from the set of differentiable functions into the set of all functions; that is, this big function takes a differentiable function, say f, to its derivative f' (recall that f' is also a function). The notation we use to emphasize this idea is

$$D_x(f) = f' \quad \text{or} \quad (D_x(f))(x) = f'(x).$$

1.2 Rates of Change

The slope of a line is a number which communicates to us the rate at which a line is rising or falling as we move in the positive direction. If f is a function, then its derivative, $f'(x)$, tells us what the slope of the tangent line to f at x is. We therefore might guess that there is a strong relation between $f'(x)$ and the rate of change of f. Most useful to us in this regard are functions which measure a quantity of something such as length, population, money, area, volume, etc. with respect to time. In this case, $f'(t)$ will tell us the rate at which the quantity is changing at any given time. We will explore this idea in what follows.

Suppose Q is some quantity which varies with time t so that $Q = f(t)$ for some function f. Then the change in Q from time t to time $t + \Delta t$ is the number $\Delta Q = f(t + \Delta t) - f(t)$. Hence, the **average rate of change of** Q from time t to time $t + \Delta t$ is the quotient

$$\frac{\Delta Q}{\Delta t} = \frac{f(t + \Delta t) - f(t)}{\Delta t}.$$

The **instantaneous rate of change of** Q at time t is the derivative

$$\frac{dQ}{dt} = \lim_{\Delta t \to 0} \frac{\Delta Q}{\Delta t} = \lim_{\Delta t \to 0} \frac{f(t + \Delta t) - f(t)}{\Delta t}.$$

For example, suppose $x(t)$ is a function which gives the horizontal displacement of a particle at time t. You can think of the particle moving along the x-axis where $x(t)$ is the x-coordinate of the particle at time t. In this case, the quantity is length and the derivative $\frac{dx}{dt}$ represents the *velocity* of the particle at time t. Recall that velocity is a vector quantity (so it has both a magnitude and a direction); in this case its direction is given by a plus or minus sign: if it is positive, it is moving to the right and if it is negative, it is moving to the left.

Similarly, suppose that $y(t)$ is a function which gives the vertical displacement of a particle at time t. Its derivative $\frac{dy}{dt}$ gives the velocity of the particle at time t, where a positive sign means it is moving upward and a negative sign means it is moving downward. An important special case of this vertical motion is the function

$$y(t) = -16t^2 + v_0 t + y_0$$

which gives the vertical displacement of an object with initial vertical velocity v_0 (given in ft/sec) and initial height y_0 (given in ft) under the influence of the downward acceleration of gravity $a = -32$ ft/sec^2.

If Q is a quantity that changes with respect to time t, the sign of its derivative determines whether the quantity is increasing or decreasing. In particular, we have

$$\frac{dQ}{dt} > 0 \quad \text{means} \quad Q \text{ is increasing}$$
$$\frac{dQ}{dt} < 0 \quad \text{means} \quad Q \text{ is decreasing.}$$

2 Rules of Differentiation

We will now use the definition of the derivative to prove some nice rules for differentiating functions that will make finding the derivative much easier than using the definition directly. I emphasize here that the definition is the foundation upon which we will build these rules. Each rule that we prove is a result of the definition of the derivative, and you can see that if you look carefully at the proofs.

The first rule is one that we have already proven, but we restate it here for reference.

Theorem 2.1 *If $c \in \mathbb{R}$ is a constant and $f(x) = c$ for all $x \in \mathbb{R}$ (so f is a* ***constant function****), then f is differentiable, and*

$$f'(x) = D_x(c) = 0 \quad \textit{for all } x \in \mathbb{R}.$$

The next theorem follows easily from the limit laws.

Theorem 2.2 *If $f, g : D \to \mathbb{R}$ are differentiable on the domain D and a is a constant, then the functions $f + g : D \to \mathbb{R}$ and $af : D \to \mathbb{R}$ are differentiable on D and their derivatives are respectively*

$$D_x(f+g) = D_x(f) + D_x(g) \quad \textit{and} \quad D_x(af) = aD_x(f).$$

Proof. We will compute the derivative of $f + g$ and af using the definition. First,

$$\begin{aligned}
(D_x(f+g))(x) &= \lim_{h\to 0} \frac{(f(x+h)+g(x+h)) - (f(x)+g(x))}{h} \\
&= \lim_{h\to 0} \frac{f(x+h)-f(x)+g(x+h)-g(x)}{h} \\
&= \lim_{h\to 0} \left(\frac{f(x+h)-f(x)}{h} + \frac{g(x+h)-g(x)}{h} \right) \\
&= \lim_{h\to 0} \frac{f(x+h)-f(x)}{h} + \lim_{h\to 0} \frac{g(x+h)-g(x)}{h} \\
&= (D_x(f))(x) + (D_x(g))(x).
\end{aligned}$$

For the second function,

$$\begin{aligned}
(D_x(af))(x) = \lim_{h\to 0} \frac{af(x+h)-af(x)}{h} &= \lim_{h\to 0} a\frac{f(x+h)-f(x)}{h} \\
&= a\lim_{h\to 0} \frac{f(x+h)-f(x)}{h} = a(D_x(f))(x).
\end{aligned}$$

These computations give the formulas for the derivative and also show that $f + g$ and af are differentiable whenever f and g are; i.e., on D. □

2.1 Product and Quotient Rules

Let's pause here a moment and discuss certain characteristics of a function. First, a function is either defined or it is not defined at a point. In pre-calculus, one finds problems concerned with finding where a function is defined: the set of all such points is called the domain of the function. Earlier in this course, we defined what we mean by a function being continuous at a point and had several examples of functions which are defined at a point but not continuous there. In the last section, the derivative was defined and we said what we mean by the function being differentiable at a point. The set of all such points is the domain of the derivative of the function. This next theorem says that differentiability is a stronger condition than continuity; that is, differentiability implies continuity.

Theorem 2.3 *If $f(x)$ is differentiable at a point x_0 in its domain, then $f(x)$ is continuous at x_0.*

Proof. Let x_0 be a point where $f(x)$ is differentiable, so that $f'(x_0)$ exists. Using the limit laws, we compute

$$\begin{aligned}
\lim_{x\to x_0} f(x) - f(x_0) &= \lim_{x\to x_0} f(x) - \lim_{x\to x_0} f(x_0) \\
&= \lim_{x\to x_0} (f(x) - f(x_0)) \\
&= \lim_{x\to x_0} \left((x - x_0)\frac{f(x) - f(x_0)}{x - x_0} \right) \\
&= \lim_{x\to x_0} (x - x_0) \cdot \lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0} \\
&= (\lim_{x\to x_0} x - \lim_{x\to x_0} x_0) f'(x_0) \\
&= (\lim_{x\to x_0} x - x_0) f'(x_0) \\
&= (x_0 - x_0) f'(x_0) = 0.
\end{aligned}$$

That is, $\lim_{x\to x_0} f(x) = f(x_0)$ which means that $f(x)$ is continuous at x_0. □

We have already seen that functions may be continuous but not differentiable. For instance, we have seen that $f(x) = |x|$ is continuous on $\mathbb{R}$, but it is not differentiable at $x = 0$. We thus have a hierarchy of functions. At the lowest level we have functions which are defined on a domain D. At the second level, we have those functions which are continuous (and therefore defined) on D. Finally, we have those functions which are differentiable (and therefore continuous) on D.

We are now ready to discuss the derivative of products and quotients of functions. The rules in these cases are a little more complicated than you might expect.

Theorem 2.4 (Product Rule for Derivatives) *If $f(x)$ and $g(x)$ are differentiable functions on a common domain D, then the function $(fg)(x) := f(x)g(x)$ is differentiable on D and its derivative is*

$$(D_x(fg)(x)) = f'(x)g(x) + f(x)g'(x) = (f'g + fg')(x).$$

Proof. Let x_0 be any point in D. We compute the derivative by adding in and subtracting out a carefully chosen term in the definition of the derivative:

$$(D_x(fg))(x_0) = \lim_{x\to x_0} \frac{f(x)g(x) - f(x_0)g(x_0)}{x - x_0}$$

$$
\begin{aligned}
&= \lim_{x\to x_0} \frac{f(x)g(x) - f(x_0)g(x) + f(x_0)g(x) + f(x_0)g(x_0)}{x - x_0} \\
&= \lim_{x\to x_0} \left(\frac{f(x) - f(x_0)}{x - x_0} g(x) + f(x_0) \frac{g(x) - g(x_0)}{x - x_0} \right) \\
&= \lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0} \cdot \lim_{x\to x_0} g(x) + \lim_{x\to x_0} f(x_0) \cdot \lim_{x\to x_0} \frac{g(x) - g(x_0)}{x - x_0} \\
&= f'(x_0)g(x_0) + f(x_0)g'(x_0)
\end{aligned}
$$

where we have used the fact that $\lim_{x\to x_0} g(x) = g(x_0)$ since g is differentiable and hence continuous. Note that all of the limits exist by the differentiability of f and g and therefore fg is differentiable on D. □

Theorem 2.5 (Quotient Rule for Derivatives) *If $f(x)$ and $g(x)$ are differentiable functions on a common domain D and $g(x) \neq 0$ for any x in D, then the function $(f/g)(x) := f(x)/g(x)$ is differentiable on D and its derivative is*

$$
(D_x(f/g)(x)) = \frac{f'(x)g(x) - f(x)g'(x)}{(g(x))^2} = \left(\frac{f'g - fg'}{g^2} \right)(x).
$$

Proof. Let x_0 be any point in D. Again we compute the derivative using the definition and adding in and subtracting off a carefully chosen term. We find

$$
\begin{aligned}
(D_x(f/g))(x_0) &= \lim_{x\to x_0} \frac{\frac{f(x)}{g(x)} - \frac{f(x_0)}{g(x_0)}}{x - x_0} \\
&= \lim_{x\to x_0} \frac{f(x)g(x_0) - f(x_0)g(x)}{g(x)g(x_0)(x - x_0)} \\
&= \lim_{x\to x_0} \frac{f(x)g(x_0) - f(x_0)g(x_0) + f(x_0)g(x_0) - f(x_0)g(x)}{g(x)g(x_0)(x - x_0)} \\
&= \lim_{x\to x_0} \left[\frac{1}{g(x)g(x_0)} \left(\frac{f(x) - f(x_0)}{x - x_0} g(x_0) - f(x_0) \frac{g(x) - g(x_0)}{x - x_0} \right) \right] \\
&= \frac{1}{\lim_{x\to x_0} g(x) \cdot \lim_{x\to x_0} g(x_0)} \\
&\qquad \left(\lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0} \cdot \lim_{x\to x_0} g(x_0) - \lim_{x\to x_0} f(x_0) \cdot \lim_{x\to x_0} \frac{g(x) - g(x_0)}{x - x_0} \right) \\
&= \frac{1}{(g(x_0))^2} [f'(x_0)g(x_0) - f(x_0)g'(x_0)]
\end{aligned}
$$

where again we have used the fact that $\lim_{x\to x_0} g(x) = g(x_0)$ since g is differentiable and hence continuous. All of the limits exist by the differentiability of f and g and the fact that $g(x) \neq 0$ for all x in D and therefore f/g is differentiable on D. □

2.2 The Power Rule

We are now ready to give a relatively easy way to compute the derivative of power functions like $f(x) = x^5$.

Theorem 2.6 (Power Rule for Derivatives) *If n is any integer and $f(x) = x^n$, then f is differentiable on $\mathbb{R}$ and its derivative is*

$$
f'(x) = nx^{n-1}.
$$

Proof. Let x_0 be any real number. We have three cases. The first case is that $n = 0$. This is easy because f is a constant function in this case and it is easy to check that the formula holds in this case; I will leave it to the reader to check.

The second case is that n is a positive integer. Convince yourself that the following equation holds for all x:

$$x^n - x_0^n = (x - x_0)(x^{n-1} + x^{n-2}x_0 + x^{n-3}x_0^2 + \cdots + x^2x_0^{n-3} + xx_0^{n-2} + x_0^{n-1}).$$

If $x \neq x_0$, then this yields

$$\frac{x^n - x_0^n}{x - x_0} = x^{n-1} + x^{n-2}x_0 + x^{n-3}x_0^2 + \cdots + x^2x_0^{n-3} + xx_0^{n-2} + x_0^{n-1}.$$

By the product law and identity function law of limits, for each $k = 1, 2, \ldots, n-1, n$ we have

$$\lim_{x \to x_0} (x^{n-k}x_0^{k-1}) = x_0^{n-k}x_0^{k-1} = x_0^{n-1}.$$

Thus if we take the limit of both sides of our equation and use the limit laws, we obtain

$$f'(x) = \lim_{x \to x_0} \frac{x^n - x_0^n}{x - x_0} = \lim_{x \to x_0} (x^{n-1} + x^{n-2}x_0 + \cdots + xx_0^{n-2} + x_0^{n-1}) = nx_0^{n-1}$$

since there are n terms in the sum on the right side.

Now suppose n is a negative integer. Then $-n$ is a positive integer and $f(x) = \frac{1}{x^{-n}}$. Hence, we can apply the quotient law for derivatives and the power rule we just proved for positive integers to compute

$$f'(x) = \frac{0 \cdot x^{-n} - 1 \cdot (-n)x^{-n-1}}{x^{-2n}} = \frac{nx^{-n-1}}{x^{-2n}} = nx^{2n-n-1} = nx^{n-1}$$

as claimed. □

Using the Power Rule and Theorem 2.2, we now know that polynomials are differentiable on $\mathbb{R}$ and we also can quickly compute the derivative of any polynomial. In fact, if $a_n, a_{n-1}, \ldots, a_2, a_1, a_0$ are the coefficients of an nth degree polynomial, then

$$\begin{aligned} &D_x(a_nx^n + a_{n-1}x^{n-1} + \cdots + a_2x^2 + a_1x + a_0) \\ &\quad = a_nD_x(x^n) + a_{n-1}D_x(x^{n-1}) + \cdots + a_2D_x(x^2) + a_1D_x(x) + D_x(a_0) \\ &\quad = na_nx^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + 2a_2x + a_1. \end{aligned}$$

2.3 The Chain Rule

We now wish to find a way to differentiate a function which is the composition of functions. For example, we would like to be able to differentiate a function like $f(x) = (x-5)^{23}$. This rule, called the *chain rule* is one of the most powerful of the differentiation rules! We will state it without proof since the proof is much harder.

Theorem 2.7 (Chain Rule) *If $f(x)$ and $g(x)$ are differentiable functions on their respective domains D_1 and D_2 and if the range of $f(x)$ is contained in the domain D_2 of $g(x)$, then the composite function $(g \circ f)(x) := g(f(x))$ is differentiable on D_1 and its derivative is*

$$(D_x(g \circ f))(x) = g'(f(x)) \cdot f'(x).$$

One way to remember this rule is that the *derivative of $g \circ f$ is the derivative of the outer function g times the derivative of the inner function f*. This statement is a little imprecise because we actually take the derivative of the outer function with respect to the inner function $f(x)$ and then multiply by the derivative of the inner function with respect to x. It is very useful to introduce an intermediate variable as follows. Let $u = f(x)$. Then $g(f(x)) = g(u)$ and the chain rule says

$$D_x(g(u)) = g'(u)\frac{du}{dx} = g'(u)f'(x) = g'(f(x)) \cdot f'(x).$$

Using the chain rule to differentiate equations via an intermediate variable is especially easy to remember. If we again take $u = f(x)$ and also let $y = g(u)$, then we write the chain rule as

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}.$$

This notation, due to Leibnitz, looks like a fraction and if we view it as such we should be convinced that the equation is true. A strong word of warning is due here however: the expression $\frac{dy}{dx}$ is a notational facilitator that means 'the derivative of y with respect to x' and is NOT a fraction! Even though we will frequently do things with it that one may do with a fraction, do not for one instant think that it is a fraction!

Let's do one example showing how one might use the chain rule in applications. This is called a *related rates problem* because we will express one rate we wish to know in terms of another rate that is related in some way.

Example. Suppose that you pump air into a spherical balloon at the constant rate of 200π cm^3/s. At what rate is the radius increasing when the radius is 5 cm? To solve this problem, think about what we want to know in mathematical language. The derivative of a function measuring the quantity of something at any time t tells you the rate at which the quantity is changing at a particular instant with respect to time. Hence, if r is the radius of the balloon, we want to know $\frac{dr}{dt}$ at the instant when the the balloon has radius 5 cm. What do we know? We know the rate at which the volume V of the balloon is changing at any time. We state this as $\frac{dV}{dt} = 200\pi$ cm^3/s. The only thing we need to know now is some equation relating the volume of a sphere to its radius. A geometry formula helps us out there: $V = \frac{4}{3}\pi r^3$. Let's use this equation to find a relationship between $\frac{dV}{dt}$ and $\frac{dr}{dt}$. If we differentiate the equation with respect to time using the chain rule, we get

$$\frac{dV}{dt} = \frac{dV}{dr} \cdot \frac{dr}{dt} = 4\pi r^2 \frac{dr}{dt}$$

(remember, we are differentiating with respect to t, not with respect to r). Hence, at the instant when $r = 5$ cm, we have

$$200\pi \text{ cm}^3/\text{s} = 4\pi(5^2 \text{ cm}^2)\frac{dr}{dt} \quad \Rightarrow \quad \frac{dr}{dt} = 2 \text{ cm/s}.$$

2.4 The Generalized Power Rule

We have not yet given a rule to compute the derivative of a root function like $f(x) = \sqrt{x}$. Recall that we write a root as an exponent:

$$\sqrt[n]{x^m} = x^{\frac{m}{n}}.$$

This exponent is a *rational number*; i.e., it is a fraction with an integer in the numerator and an integer in the denominator. We call such an exponent a **rational**

exponent. The power rule we have so far is only valid for integers and so it does not apply in general to this case. What is nice, however, is that the differentiation rule we find for this case is just a generalization of the power rule which includes rational exponents!

Theorem 2.8 (Generalized Power Rule) *Let $r = \frac{p}{q}$ be a rational number. If $f(x) = x^r$, then $f(x)$ is differentiable on $(-\infty, 0) \cup (0, \infty)$ if q is odd and on $(0, \infty)$ if q is even. Furthermore, its derivative is*

$$f'(x) = rx^{r-1}.$$

We will not give the proof of this theorem here. As a quick example, we compute the derivative of $f(x) = \sqrt{x} = x^{1/2}$ to be

$$f'(x) = \frac{1}{2}x^{-1/2} = \frac{1}{2\sqrt{x}}.$$