MATH 137: Calculus 1 for Honours Mathematics

Fall 2025 Edition

We would like to acknowledge contributions from:

Erik Hintz (Lead Author, Chapters 0, 1 & 2)

Sachin Kotecha (Lead Author, Chapters 3 & 4)

Eric Bembenek (Reviewer and Problems Author)

Jordan Hamilton (Reviewer and Project Lead)

Zack Cramer (Reviewer and Problems Author)

Matthew Scott (Problems Author)

Chelsea Uggenti (Reviewer)

Contents

Chapter 0

Pre-Calculus Review

0.1 Real-Valued Functions

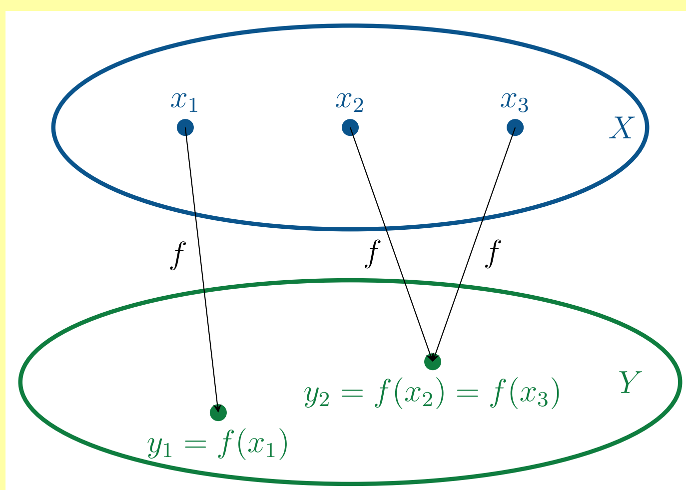
Functions are one of the most important objects used in mathematics and can be used to describe a variety of real-world phenomena. For example, the distance driven by a car at a constant velocity v is given by $x(t) = v \cdot t$ where t denotes the time. Similarly, the total money accrued in a savings account with an interest rate $0 < r < 1$ is given by $M(n) = P \cdot (1 + r)^n$ where P in the initial principal value and n denotes the number of years. The area of a square is a function of the side length, $A(s) = s^2$, and so on.

Given the wide range of examples, we define a function in the following way.

Definition 0.1.1
Function

Let X and Y be sets. A function f is a mapping that assigns to each $x \in X$ exactly one $y = f(x) \in Y$. We use the notation

$$f : X \to Y, \quad x \mapsto f(x).$$



REMARK

- Note that definition 0.1.1 does not specify what exactly the sets X and Y ought to look like. Common sets that may appear in MATH 137 and 138 are
  - Natural Numbers: $\mathbb{N} = \{1, 2, 3, \dots\}$ or sometimes $\mathbb{N} = \{0, 1, 2, \dots\}$ (Whether we include 0 in $\mathbb{N}$ will be clear in the context of any assessment.)
  - Integers: $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$
  - Rational Numbers: $\mathbb{Q} = \{\frac{p}{q} : p, q \in \mathbb{Z}, q \neq 0\}$
  - Real Numbers: $\mathbb{R}$ which is the set rational numbers plus the set of irrational numbers! (Irrational numbers are numbers that cannot be expressed as $\frac{p}{q}$ where p and q are integers. This includes numbers like $\pi, e, \sqrt{2}$, etc.)
- In higher level courses, you'll see all kinds of other functions as well, for instance, where X and Y are sets of vectors, or even sets of functions!

A function is *well defined* when any input, x, results in at most one unique output, y.

An easy way to verify if a given mapping f is a well-defined function, we can apply the vertical line test.

Theorem 0.1.2 **(Vertical Line Test)**

If any vertical line intersects the graph of a mapping in all places at most once, then the given mapping is a well-defined function. We say this function passes the vertical line test.

Proof: A mapping $f : X \to Y$ is a function if and only if for every $x \in X$ there is exactly one $y \in Y$ such that $f(x) = y$. This is the case if and only if the graph of f crosses the vertical line at x exactly once (namely, at the point $(x, f(x))$). □

The set of values X that we are allowed to plug into a function f as well as the set of outputs Y after applying the mapping f get special names!

Definition 0.1.3
Domain of a Function

Let $f : X \to Y$ be a function. We call the set of numbers for which the function f is well defined the **domain** of f. More formally,

$$D(f) = \{x : f(x) \text{ is well defined}\}.$$

Definition 0.1.4
Range of a Function

The **range** of a function $f : X \to Y$ is the set

$$R(f) = \{f(x) : x \in X\}.$$

In simpler words, the domain of a function is the set of numbers that we can input into f and get a number back, and the range of a function is the set of numbers that f attains. We will later learn methods to compute the range of complicated functions, but for now let us look at a few review examples.

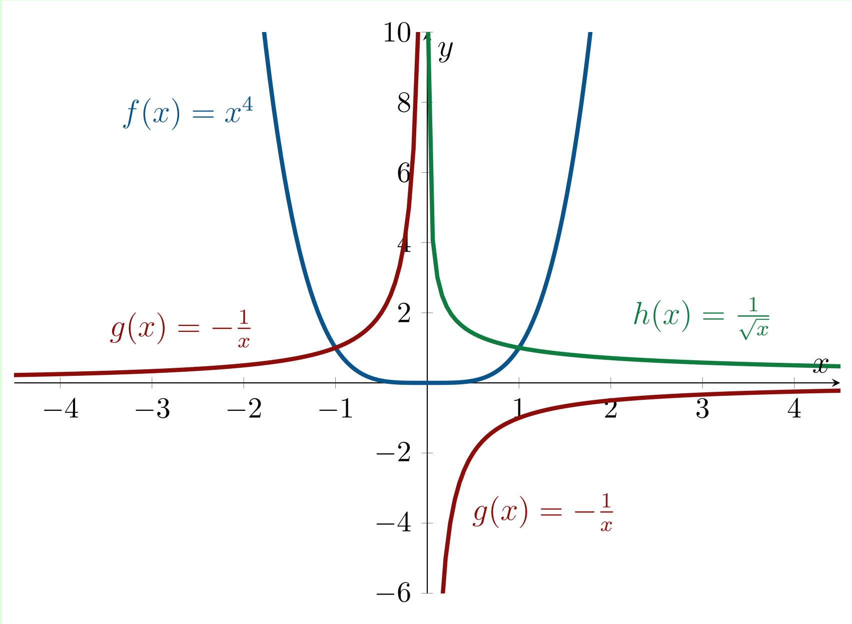
Example 0.1.5

For each of the functions below, give its domain and range.

1. $f(x) = x^4$,
2. $g(x) = -\frac{1}{x}$,
3. $h(x) = \frac{1}{\sqrt{x}}$.

Solution:

1. The domain of $f(x) = x^4$ is $\mathbb{R}$, as the power x^4 is always defined. The range is $[0, \infty)$.
2. The domain of $g(x) = -\frac{1}{x}$ is $D = \mathbb{R} \setminus \{0\}$, i.e., all real numbers except for $x = 0$, and the range is also $\mathbb{R} \setminus \{0\}$.
3. Finally, the domain of $h(x) = \frac{1}{\sqrt{x}}$ is $(0, \infty)$, because we 1) cannot divide by 0 and 2) the square root is only defined for non-negative numbers. Since $\sqrt{x} > 0$ for all $x > 0$, we have that the range is $(0, \infty)$.



Definition 0.1.6

Odd and Even Function

A function f is called **even** if $f(-x) = f(x)$ for all $x \in D$.

A function g is called **odd** if $g(-x) = -g(x)$ for all $x \in D$.

Example 0.1.7

Any power function of the form $f(x) = x^{2k}$ for some $k \in \mathbb{N}$ is even, as

$$f(-x) = (-x)^{2k} = (-1)^{2k}x^{2k} = x^{2k} = f(x).$$

Any odd power function of the form $g(x) = x^{2k-1}$ for some $k \in \mathbb{N}$ is odd, as

$$g(-x) = (-x)^{2k-1} = (-1)^{2k-1}x^{2k-1} = -x^{2k-1} = -g(x).$$

We are often required to solve equations, say $x^2 = 2$ or $x^2 - 2 = 0$, and of course much more complex ones. The solution of an equation can often be expressed as the root of a function, which we define as follows.

Definition 0.1.8
Root of a function

Suppose f is a function, and suppose there is $r \in D$ so that $f(r) = 0$. Then we call r a **root** of f.

A given function may have no root, one root, or multiple roots.

We can build more complicated functions by summing, multiplying, or dividing other functions. For instance, the function $(f + g)$ is given by $f(x) + g(x)$ (e.g., if $f(x) = x^2$ and $g(x) = x$, then $(f + g)(x) = f(x) + g(x) = x^2 + x$). We can also apply translations: if $f(x) = x^2$ is an upward-opening parabola with a vertex at $(0, 0)$ then $g(x) = (x - 1)^2$ is an upward-opening parabola with a vertex at $(1, 0)$. Similarly, we can translate functions vertically as well (e.g., $f(x) = x^2$ to $g(x) = x^2 + 1$).

Another method to create new functions is to "plug them into each other", which we define as follows.

Definition 0.1.9
Function Composition

Let $f : X \to Y$ and $g : Y \to Z$ be functions. The composition of f and g, denoted by $g \circ f$, is the function

$$g \circ f : X \to Z \quad \text{with} \quad (g \circ f)(x) = g(f(x)).$$

The situation in Definition 0.1.9 is illustrated in Figure 0.1.1.

The order of the composition matters: $(g \circ f)(x)$ is in general not $(f \circ g)(x)$; the two functions may not even have the same domain.

Example 0.1.10

Let $f(x) = x^2$ and $g(x) = x + 1$. Write down the functions $g \circ f$ and $f \circ g$ along with their domains and ranges.

Solution:

First,

$$(g \circ f)(x) = g(f(x)) = g(x^2) = x^2 + 1$$

has domain $\mathbb{R}$ and range $[1, \infty)$.

Next,

$$(f \circ g)(x) = f(g(x)) = f(x + 1) = (x + 1)^2$$

has domain $\mathbb{R}$ and range $[0, \infty)$.

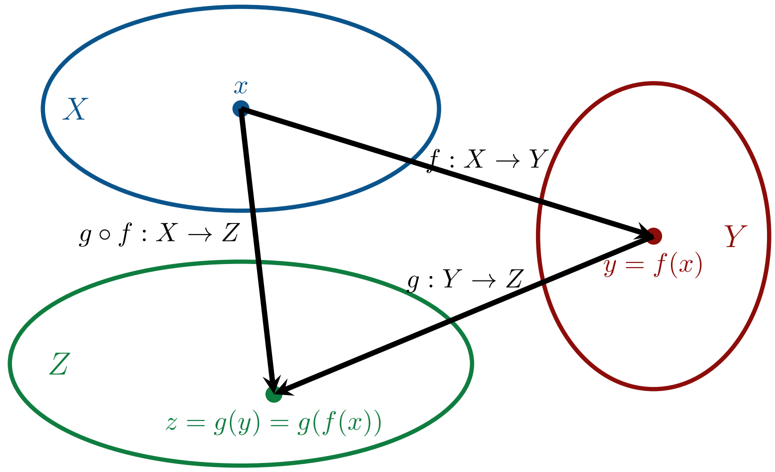


Figure 0.1.1: The figure displays the function f represented by an arrow from $x \in X$ to $y \in Y$, the function g represented by an arrow from $y \in Y$ to $z \in Z$, and and finally the composition $g \circ f$ represented by an arrow directly from $x \in X$ to $z \in Z$.

Some caution is needed when determining the domain and range of the composition $(g \circ f)(x) = g(f(x))$. Notice that we are first evaluating $f(x)$, meaning that we need x to be in the domain of f. Then, we substitute $y = f(x)$ into g to compute $g(y) = g(f(x))$. For this to be well defined, we additionally require that $y = f(x)$ to be in the domain of g. That is, the domain of $g(f(x))$ are those x in the domain of f which satisfy that $y = f(x)$ is in the domain of g. In particular, the domain of $g(f(x))$ is a subset of the domain of f.

Example 0.1.11 Let $f(x) = \sqrt{x-9}$ and $g(x) = x^2 + 9$. Determine the domain and range of the functions f, g, and $g \circ f$.

Solution:

The domain of f is $[9, \infty)$ with range $[0, \infty)$, while the domain of g is $\mathbb{R}$ with range $[9, \infty)$.

For the function $g(f(x))$, the domain are those $x \in [9, \infty)$ for which $f(x)$ is in the domain of g, that is, for which $f(x) \in \mathbb{R}$. This holds for all $x \in [9, \infty)$, so the range of $g(f(x))$ is $[9, \infty)$.

For the range of $g(f(x))$, we first note that the function f maps the domain $[9, \infty)$ to the set $[0, \infty)$. The function g then maps the set $[0, \infty)$ to $[9, \infty)$, which is the range of $g(f(x))$.

Care must be taking when just looking at a simplified version of the composition, for instance after writing it out. Here, we find

$$g(f(x)) = g(\sqrt{x-9}) = \left(\sqrt{x-9}\right)^2 + 9 = x,$$

from which we might have guessed a domain and range $\mathbb{R}$. But for the composition $g(f(x))$ to be well defined, we need $f(x)$ well defined, hence, $x \geq 9$.

Next, we define the inverse function as the function f^{-1} that "undoes" what f did.

Definition 0.1.12
Inverse Function

Let $f : X \to Y$ be a function with domain X and range Y. Then f is **invertible** if there exists a function $f^{-1} : Y \to X$ so that

$$f^{-1}(f(x)) = x \quad \text{for all } x \in X \qquad \text{and} \qquad f(f^{-1}(x)) = x \quad \text{for all } x \in Y.$$

If the inverse exists, we can find it as follows:

1. Write down the equation $f(x) = y$.
2. Solve the equation for x.
3. Finally, swap x with y to get the inverse function $y = f^{-1}(x)$

If the inverse does not exist, we would be unable to complete Step 2.

Geometrically, the inverse function is the mirror image at the diagonal $y = x$, as will be illustrated in the next example.

Example 0.1.13

Let $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = 2x + 4$. Compute the inverse, $f^{-1}(x)$, for $x \in \mathbb{R}$.

Solution:

Let $y \in \mathbb{R}$. We solve the equation $f(x) = y$ for x:

$$f(x) = y \Leftrightarrow 2x + 4 = y \Leftrightarrow x = \frac{y-4}{2}.$$

After swapping the variables x and y, we find that

$$f^{-1}(x) = \frac{x-4}{2}$$

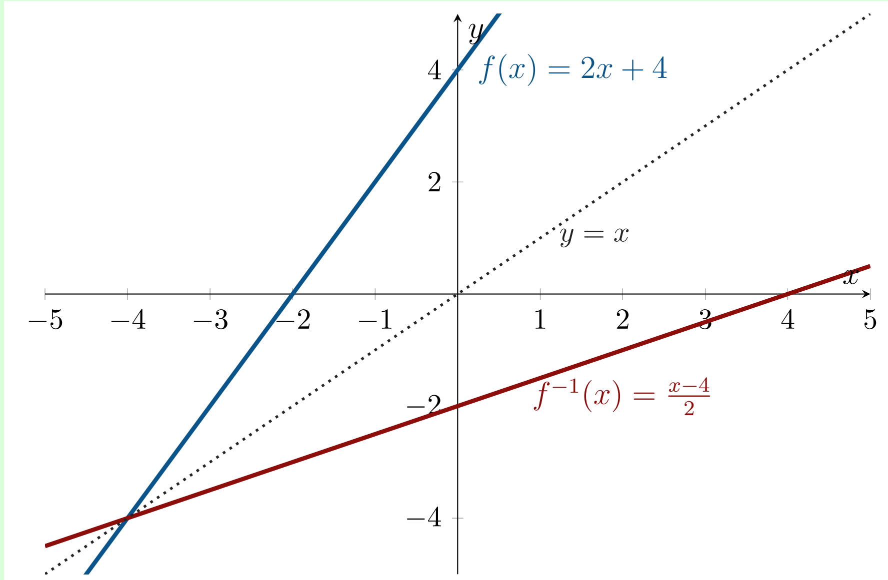
is the inverse function of f. We can verify this manually by checking the definition: First,

$$f(f^{-1}(x)) = f\left(\frac{x-4}{2}\right) = 2 \cdot \frac{x-4}{2} + 4 = x$$

for all $x \in \mathbb{R}$. Second,

$$f^{-1}(f(x)) = f^{-1}(2x+4) = \frac{2x+4-4}{2} = x$$

for all $x \in \mathbb{R}$.



Not every function has an inverse everywhere!

Example 0.1.14 For each of the following functions, determine the inverse function, if it exists.

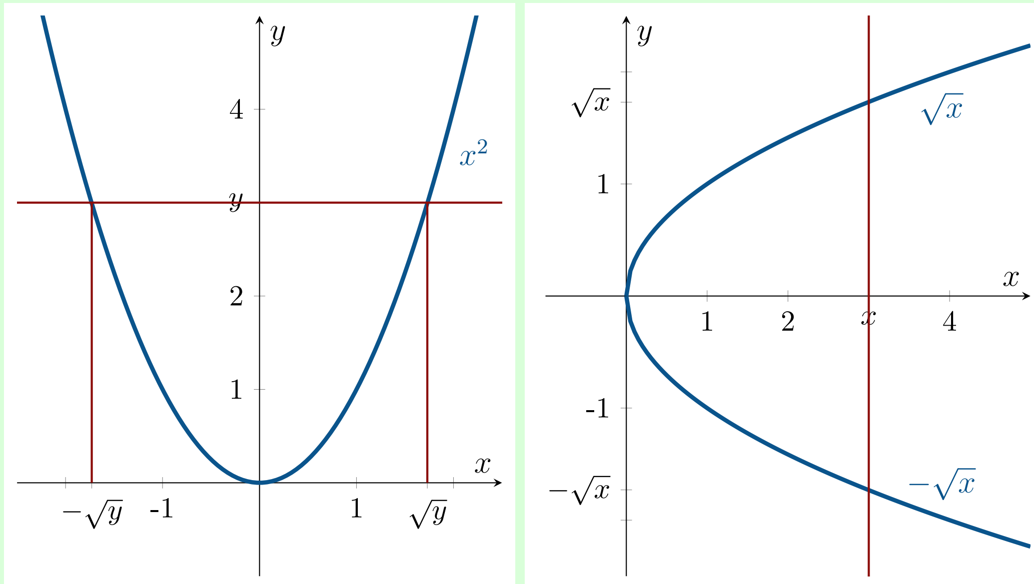
$$f_1 : \mathbb{R} \to [0, \infty), \quad f_1(x) = x^2;$$
$$f_2 : [0, \infty) \to [0, \infty), \quad f_2(x) = x^2.$$

Solution:

Let $f_1 : \mathbb{R} \to [0, \infty)$ with $f_1(x) = x^2$. The range of $f_1(x)$ is $[0, \infty)$. Let's try solving $f_1(x) = y$ for x, where $y \geq 0$ is given:

$$f_1(x) = y \Leftrightarrow x^2 = y \Leftrightarrow x = \pm\sqrt{y}.$$

But we cannot define something like $\pm\sqrt{x}$ to be a function! A function must map to each x in the domain exactly to one y in the range. As such, f_1 does not have an inverse function over its full domain $\mathbb{R}$. This is illustrated in the figure below: The left shows the graph of the function x^2, which crosses the horizontal line at y twice, so there is no unique solution x satisfying $f_1(x) = y$. The right shows the functions $\pm\sqrt{x}$, and we can see that it fails the vertical line test (in that there is no unique y such that $y = f(x)$). Furthermore, we can see on the right that the inverse cannot be defined for $x < 0$, as the square root of a negative number is undefined, or in other words, because the image of the function x^2 is the non-negative real line.



But what about restricting the domain? That's exactly what the function f_2 is doing. Suppose we restrict the domain to the non-negative real line, that is, consider $f_2 : [0, \infty) \to [0, \infty)$ with $f(x) = x^2$. Proceeding just as above, we find

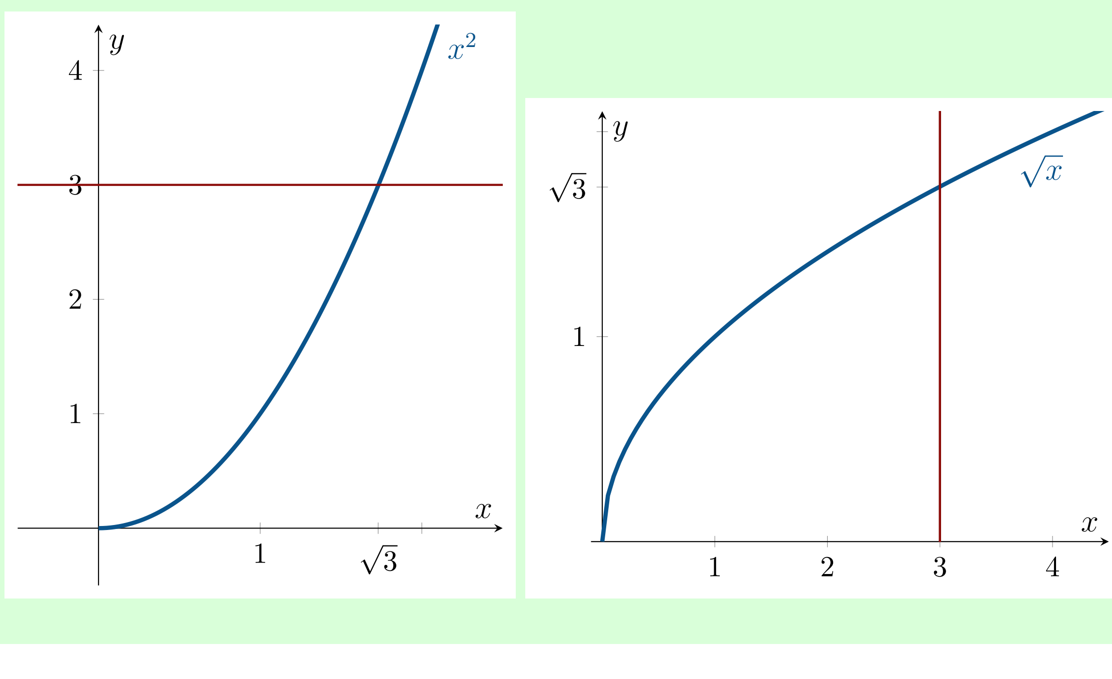
$$f_2(x) = y \Leftrightarrow x = \pm\sqrt{y}.$$

But (unless $y = 0$), we have $x = -\sqrt{y} < 0$ is not in the domain $[0, \infty)$, so we can ignore the negative solution of this equation and define (after swapping x and y)

$$f_2^{-1}(x) = \sqrt{x}$$

to be the inverse of $f_2(x) = x^2$ on $[0, \infty)$.

This is illustrated in the figure below: The graph of the function crosses the horizontal line at y exactly once, so there is unique solution $x = f_2^{-1}(y)$. Furthermore, the graph of the function $\sqrt{x}$ on the right passes the vertical line test.



The previous example illustrated different situations that can arise when solving $f(x) = y$ for a function $f : X \to Y$ for some $x \in X$ and $y \in Y$.

- There is no solution $x \in X$ satisfying $f(x) = y$. In this case, $f^{-1}(y)$ does not exist.
- There are at least two solutions, say $x_1, x_2 \in X$, satisfying $f(x_1) = f(x_2) = y$. In this case, $f^{-1}(y)$ does not exist.
- There is exactly one solution $x \in X$ satisfying $f(x) = y$. In this case, $f^{-1}(y)$ exists and $x = f^{-1}(y)$.

But how can we easily determine whether a function is invertible? Luckily, we have an easy visual test that can be applied!

Theorem 0.1.15 **(Horizontal Line Test)**

If the horizontal line intersects the graph of a function in all places at exactly one point, then the given function is invertible. We say this function passes the horizontal line test.

Next, let us review some basic elementary functions along with some of their properties.

0.2 Polynomials

A common class of functions are polynomials:

Definition 0.2.1
Polynomial Function

A **polynomial** is a function $p : \mathbb{R} \to \mathbb{R}$ of the form

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

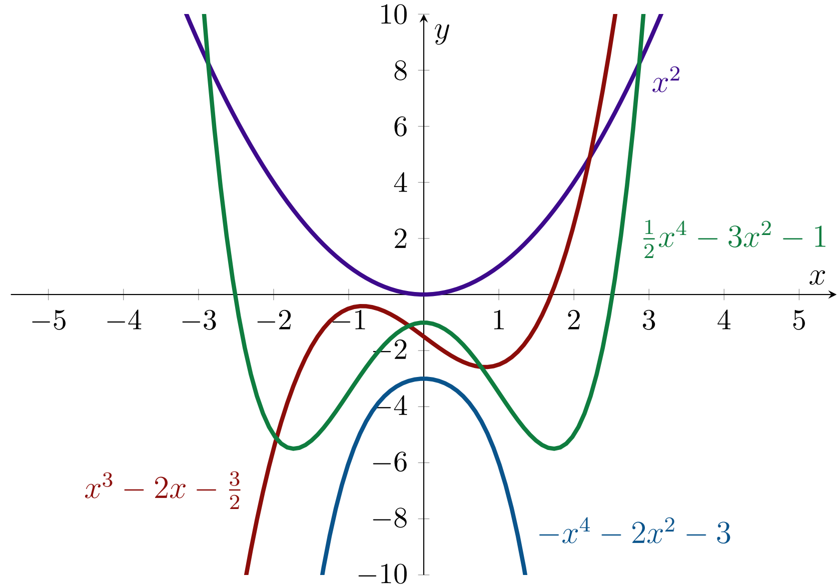
where $a_0, a_1, \ldots, a_n \in \mathbb{R}$ and $n = 0, 1, 2, \ldots$).

The domain of any polynomial is the entire real line, and the range depends on the coefficients.

Definition 0.2.2
Degree of Polynomial

Let $p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$ be a polynomial. The largest n such that $a_n \neq 0$ is called the **degree** of the polynomial and is denoted $\deg(p) = n$.

For instance, $p(x) = 5x^5 + 3x^2 + 1$ has degree $\deg(p) = 5$. Various polynomials are depicted below. It is important that the powers of x are natural numbers (including 0): The functions $f(x) = x^{0.5} + 2x^2$ and $g(x) = \frac{1}{x}$ are **not** polynomials.



The existence, and formulas, for the root of polynomials depends on the degree n. A polynomial of degree n can have at most n real roots.

- The linear function $p(x) = a_1 x + a_0$ has one root $x_1 = -\frac{a_0}{a_1}$ if $a_1 \neq 0$, otherwise no roots exist.
- The quadratic polynomial $p(x) = a_2 x^2 + a_1 x + a_0$ has
  - No real root if $a_1^2 - 4a_2 a_0 < 0$ (we call these "irreducible quadratic polynomials");
  - One root if $a_1^2 - 4a_2 a_0 = 0$, given by
  $$x_1 = -\frac{a_1}{2a_2};$$
  - Two roots if $a_1^2 - 4a_2 a_0 > 0$, given by
  $$x_{1,2} = \frac{-a_1 \pm \sqrt{a_1^2 - 4a_2 a_0}}{2a_2}.$$
- For polynomials of degree 3 and 4, there are (complicated) closed formulas similarly to the quadratic case.

However, for polynomials of degree larger than 4, there are no closed formulas! If we can guess a root, we can then use polynomial long division to reduce the problem to finding the root of a polynomial with smaller degree.

For example, suppose that $p(x)$ is a polynomial with $\deg(p) = n$, and r_1 is a root of $p(x)$. Then there exists a polynomial $q(x)$ with degree $\deg(q) = n - 1$ such that

$$p(x) = (x - r_1)q(x).$$

Similarly, if r_2 is a root of $q(x)$, then we can write

$$p(x) = (x - r_1)(x - r_2)s(x)$$

where $s(x)$ is a polynomial with $\deg(s) \leq n-2$. We can repeat this process, so that if $p(x)$ is a polynomial with $\deg(p) = n$ with roots, $r_1, r_2, \ldots, r_m$ (where $m \leq n$), then we can write

$$p(x) = (x - r_1)(x - r_2) \cdot \cdots \cdot (x - r_m)q(x)$$

where $q(x)$ is a polynomial of $\deg(q) = n - m$. If $p(x)$ does not have any roots, we get the trivial decomposition $p(x) = q(x)$.

A LOOK AHEAD

In practice, it is often difficult to find the exact roots of functions. For example, the roots of the function $f(x) = e^x - x - 2$ cannot be written in a "nice" way. However, a plot of the function shows roots near -1.84 and 1.15.

To better approximate roots of functions, we will learn about the Bisection Method and Newton's Method in Sections 2.10 and 3.6!

Example 0.2.3

For the polynomial $p(x) = 5x^5 + 10x^4 - 10x^3 - 40x^2 - 115x + 150$, one gets

$$p(x) = 5x^5 + 10x^4 - 10x^3 - 40x^2 - 115x + 150 = 5(x^2 + 2x + 5)(x - 1)(x + 3)(x - 2),$$

from which we can see it has roots $\{-3, 1, 2\}$. Since $2^2 - 4 \cdot 1 \cdot 5 < 0$, the quadratic polynomial $x^2 + 2x + 5$ has no roots.

In the previous example, how did we go from $p(x) = 5x^5 + 10x^4 - 10x^3 - 40x^2 - 115x + 150$ to $p(x) = 5(x^2 + 2x + 5)(x - 1)(x + 3)(x - 2)$? Verifying this equation is correct is easy (we just multiply and simplify the factored expression), but without knowledge of the roots, it is hard to find the factored form. Note that by taking a close look at p, we might have guessed the root $r_1 = 1$. So we know there is a polynomial q with degree 4 so that $5x^5 + 10x^4 - 10x^3 - 40x^2 - 115x + 150 = (x - 1)q(x)$. But how can we find q?

More generally, suppose you have a polynomial $p(x)$ with $\deg(p) = n$, and you are able to guess a root of p, say $x = r$. How can $q(x)$ be found such that $p(x) = (x - r)q(x)$? One approach is to use polynomial long division, which we discuss next!

Polynomial Long Division

Polynomial long division is an algorithm that works similarly to long division of real numbers. It starts with two polynomials $a(x)$ (dividend) and $b(x)$ (divisor). Then there are polynomials $Q(x)$ (quotient) and $R(x)$ (remainder) with

$$a(x) = b(x) \cdot Q(x) + R(x),$$

or equivalently

$$\frac{a(x)}{b(x)} = Q(x) + \frac{R(x)}{b(x)},$$

where either $R(x) = 0$ or $\deg(R) < \deg(b)$.

Example 0.2.4

1. Let $a(x) = (x-2)(x+1) = x^2 - x - 2$ and $b(x) = (x+1)$. Then
$$\frac{a(x)}{b(x)} = \frac{(x-2)(x+1)}{(x+1)} = x - 2 + 0,$$
so we find
$$Q(x) = (x-2), \quad R(x) = 0.$$

2. Let $a(x) = (x-2)(x+1) = x^2 - x - 2 = (x-1)x - 2$ and $b(x) = x$. Then
$$\frac{a(x)}{b(x)} = \frac{(x-1)x-2}{x} = x - 1 - \frac{2}{x},$$
so we find
$$Q(x) = (x-1), \quad R(x) = -2$$

In these two examples, it was easy to find $Q(x)$ and $R(x)$. But what about more complex polynomials? If we can't already "see" the factorization, we can use polynomial long division. We will illustrate the process through an example.

Example 0.2.5 Consider
$$\frac{x^3 - 12x^2 + 38x - 17}{x - 7}$$

1. Divide the leading term of the dividend x^3 by the leading term of the divisor x and put the result $\frac{x^3}{x} = x^2$ on top of the table.
$$\begin{array}{r|llll} & x^2 & & & \\ \hline (x-7) & x^3 & -12x^2 & +38x & -17 \end{array}$$
Then, multiply this result x^2 with the divisor $(x-7)$ and put the product $x^3 - 7x^2$ below the dividend.
$$\begin{array}{r|llll} & x^2 & & & \\ \hline (x-7) & x^3 & -12x^2 & +38x & -17 \\ & x^3 & -7x^2 & & \end{array}$$

2. Next, subtract the dividend $x^3 - 12x^2 + 38x - 17$ from the obtained product $x^3 - 7x^2$ and write the difference $-5x^2 + 38x - 17$ at the bottom.
$$\begin{array}{r|llll} & x^2 & & & \\ \hline (x-7) & x^3 & -12x^2 & +38x & -17 \\ & x^3 & -7x^2 & & \\ \hline & & -5x^2 & +38x & -17 \end{array}$$

3. We repeat Steps 1. and 2. with the new dividend $-5x^2+38x-17$. That is, we divide its leading term $-5x^2$ by the leading term of the divisor x, and put the result $\frac{-5x^2}{x} = -5x$ on top of the table. Then, we subtract the product of the result and the divisor $(-5x)\cdot(x-7) = -5x^2+35x$ from the new dividend $-5x^2+38x-17$ and note the result $-5x^2+38x-17-(-5x^2+35x) = 3x-17$ below.

$$\begin{array}{c|cccc}
 & x^2 & -5x & & \\
\hline
(x-7) & x^3 & -12x^2 & +38x & -17 \\
 & x^3 & -7x^2 & & \\
\hline
 & & -5x^2 & +38x & -17 \\
 & & -5x^2 & +35x & \\
\hline
 & & & 3x & -17
\end{array}$$

4. Again, we repeat Steps 1. and 2. with the new dividend $3x-17$. Its leading term $3x$ divided by the leading term of the divisor x is 3, which we put on top of the table. Then, subtract $3\cdot(x-7) = 3x-21$ from $3x-17$ to find the remainder $3x-17-3x-21 = 4$, which is noted at the bottom of the table.

$$\begin{array}{c|cccc}
 & x^2 & -5x & +3 & \\
\hline
(x-7) & x^3 & -12x^2 & +38x & -17 \\
 & x^3 & -7x^2 & & \\
\hline
 & & -5x^2 & +38x & -17 \\
 & & -5x^2 & +35x & \\
\hline
 & & & 3x & -17 \\
 & & & 3x & -21 \\
\hline
 & & & & +4
\end{array}$$

5. Since the new dividend is a constant, its degree is 0, and this is smaller than the degree of the divisor (the divisor has degree 1). Hence, we stop and we can read off $Q(x)$ on the top and $R(x)$ at the bottom of the table, which is repeated here:

$$\begin{array}{c|cccc}
 & x^2 & -5x & +3 & \\
\hline
(x-7) & x^3 & -12x^2 & +38x & -17 \\
 & x^3 & -7x^2 & & \\
\hline
 & & -5x^2 & +38x & -17 \\
 & & -5x^2 & +35x & \\
\hline
 & & & 3x & -17 \\
 & & & 3x & -21 \\
\hline
 & & & & +4
\end{array}$$

That is,

$$Q(x) = x^2-5x+3, \quad R(x) = 4$$

so that

$$\frac{x^3-12x^2+38x-17}{x-7} = x^2-5x+3+\frac{4}{x-7}.$$

When dividing a polynomial by a linear term with a leading 1, such as $x-r$, one can also use synthetic division, which achieves the same result as polynomial long division with fewer computations. We illustrate synthetic division using the example just shown.

Example 0.2.6 Consider again

$$\frac{x^3 - 12x^2 + 38x - 17}{x - 7}.$$

We can obtain the same result as previously as follows:

1. Initialize a table the coefficients of $a(x) = 1 \cdot x^3 - 12 \cdot x^2 + 38 \cdot x - 17$ on the top, followed by two empty rows. Put the root of the denominator $b(x) = x - 7$ (here 7) to the left of the middle row.

$$\begin{array}{c|cccc} & 1 & -12 & 38 & -17 \\ 7 & & & & \\ \hline & & & & \end{array}$$

2. The first element in the last row (the result row) is the first element in the first row, here 1.

$$\begin{array}{c|cccc} & 1 & -12 & 38 & -17 \\ 7 & & & & \\ \hline & 1 & & & \end{array}$$

3. Multiply the most recent element in the result row (here 1) with the number to the left of the scheme (here 7) and put the result (here $1 \cdot 7 = 7$) in the middle row of the next column.

$$\begin{array}{c|cccc} & 1 & -12 & 38 & -17 \\ 7 & & 7 & & \\ \hline & 1 & & & \end{array}$$

4. Add the elements in the next column and put the result (here $-12 + 7 = -5$) in the result row.

$$\begin{array}{c|cccc} & 1 & -12 & 38 & -17 \\ 7 & & 7 & & \\ \hline & 1 & -5 & & \end{array}$$

5. We repeat Steps 3. and 4. until we reach the end of the table. That is, we multiply the most recent element in the result row (which is -5) with the number to the left of the scheme (here 7) and put the result (here $-5 \cdot 7 = -35$) in the next column.

$$\begin{array}{c|cccc} & 1 & -12 & 38 & -17 \\ 7 & & 7 & -35 & \\ \hline & 1 & -5 & & \end{array}$$

   We add up the numbers in the next column (here $38 + -35 = 3$) to obtain the next element in the result row.

$$\begin{array}{c|cccc} & 1 & -12 & 38 & -17 \\ 7 & & 7 & -35 & \\ \hline & 1 & -5 & 3 & \end{array}$$

6. We repeat Steps 3. and 4. with the new result: Multiply the most recent result (here 3) with the number to the left of the scheme (here 7) to obtain the next number in the middle row, given by 21. Add up the column to obtain the next element in the result row as $-17 + 21 = 4$.

$$\begin{array}{c|cccc} & 1 & -12 & 38 & -17 \\ 7 & & 7 & -35 & 21 \\ \hline & 1 & -5 & 3 & 4 \end{array}$$

7. We can read off the result of

$$\frac{a(x)}{b(x)} = Q(x) + \frac{R(x)}{b(x)}$$

in the result row: The last number correspond to $R(x)$, while the remaining numbers are the coefficients of $Q(x)$. Here, we find

$$R(x) = 4$$

and

$$Q(x) = 1 \cdot x^2 - 5 \cdot x + 3$$

so that

$$\frac{x^3 - 12x^2 + 38x - 17}{x - 7} = x^2 - 5x + 3 + \frac{4}{x - 7}.$$

REMARK

When using either polynomial long division or synthetic division, we must ensure that we leave a space or we put a $0 \cdot x^k$ in places where the polynomial is missing a term. For instance, the division of $x^2 - 4$ by $x + 2$ in the synthetic division table needs to be initialized as

$$\begin{array}{c|ccc} & 1 & 0 & -4 \\ -2 & & & \\ \hline & & & \end{array}$$

This will correctly lead to

$$\begin{array}{c|ccc} & 1 & 0 & -4 \\ -2 & & -2 & 4 \\ \hline & 1 & -2 & 0 \end{array}$$

From this, we can read that

$$\frac{x^2 - 4}{x + 2} = 1 \cdot x - 2 + 0$$

As mentioned earlier, one application of polynomial long division is to find a factorization of it. Recall if $p(x)$ is a polynomial with root r, then $p(x) = (x - r)q(x)$, and we now know exactly how to find $q(x)$.

Example 0.2.7 Let $p(x) = x^3 + 8x^2 - 31x + 22$. Compute a fully factored form of $p(x)$.

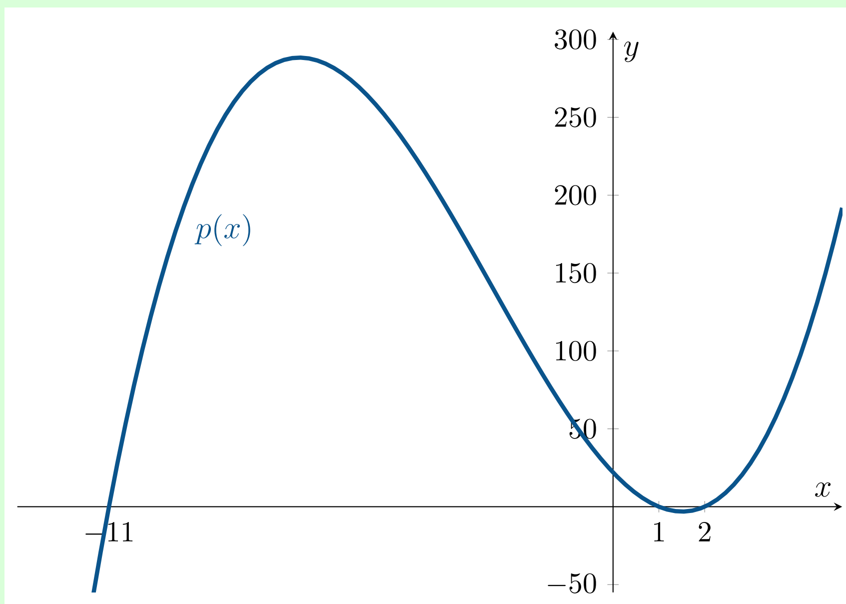
Solution:

First, we try to guess one of the roots of $p(x)$. After some trial and error, we find that $p(1) = 0$, i.e., $r_1 = 1$ is a root of $p(x)$. Polynomial long division or synthetic division of $(x^3 + 8x^2 - 31x + 22)$ divided by $(x - 1)$ gives $x^2 + 9x - 22$, hence,

$$p(x) = (x - 1) \cdot (x^2 + 9x - 22).$$

Next, we need to find the roots of $x^2 + 9x - 22$. Being a quadratic function, we can use the quadratic formula, and find that it has roots $r_2 = 2$ and $r_3 = -11$. Hence,

$$p(x) = (x - 1) \cdot (x - 2) \cdot (x + 11).$$



0.3 Exponential and Logarithm Functions

The (natural) exponential function is given by

$$f(x) = e^x$$

with domain $\mathbb{R}$ and range $(0, \infty)$ where $e \approx 2.71$ is Euler's constant.

The exponential function e^x has domain $(0, \infty)$ and is invertible, with inverse

$$f^{-1}(x) = \ln(x)$$

called the natural logarithm. The domain of $\ln(x)$ is $(0, \infty)$ and the range is $(-\infty, \infty)$ (in other words, $\mathbb{R}$).

We summarize important exponentiation and logarithm rules in the following lemma.

Lemma 0.3.1 The following rules hold for $x, y \in \mathbb{R}$,

$$\begin{aligned} e^x \cdot e^y &= e^{x+y}, \\ (e^x)^y &= e^{xy}, \\ \ln(e^x) &= x, \end{aligned}$$

while for $x, y > 0$ and $a \in \mathbb{R}$,

$$\begin{aligned} \ln(x^a) &= a \ln(x), \\ \ln(x \cdot y) &= \ln(x) + \ln(y), \\ \ln(x/y) &= \ln(x) - \ln(y), \\ e^{\ln(x)} &= x. \end{aligned}$$

Instead of the exponential function with Euler's constant e as the base, we could also consider, for $a > 0$, the function

$$g(x) = a^x, \quad x \in \mathbb{R},$$

with range $(0, \infty)$ and inverse

$$g^{-1}(x) = \log_a(x), \quad x > 0,$$

with range $\mathbb{R}$, called the logarithm with base a.

We can relate the function a^x to the natural exponential function as follows:

$$a^x = e^{\ln(a^x)} = e^{x \ln(a)}. \tag{1}$$

This connection also allows us to express the logarithm with any base $a > 0$ in terms of the natural logarithm.

Lemma 0.3.2 For any $a, x > 0$,

$$\log_a(x) = \frac{\ln(x)}{\ln(a)}.$$

Proof: By definition, $\log_a(x)$ is the inverse of the function $f(x) = a^x$. We find the inverse by solving $f(x) = y$ for x, where $y > 0$ is fixed. Using Equation (1) in the first step, and applying the natural logarithm in the second step, we obtain

$$a^x = y \Leftrightarrow e^{x \ln(a)} = y \Leftrightarrow x \ln(a) = \ln(y) \Leftrightarrow x = \frac{\ln(y)}{\ln(a)}.$$

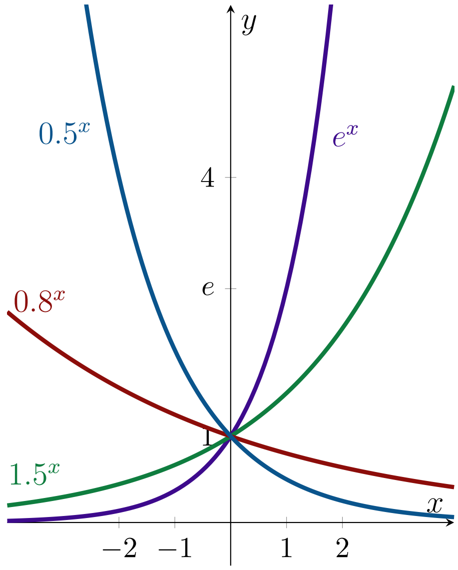
Renaming y to x, we conclude

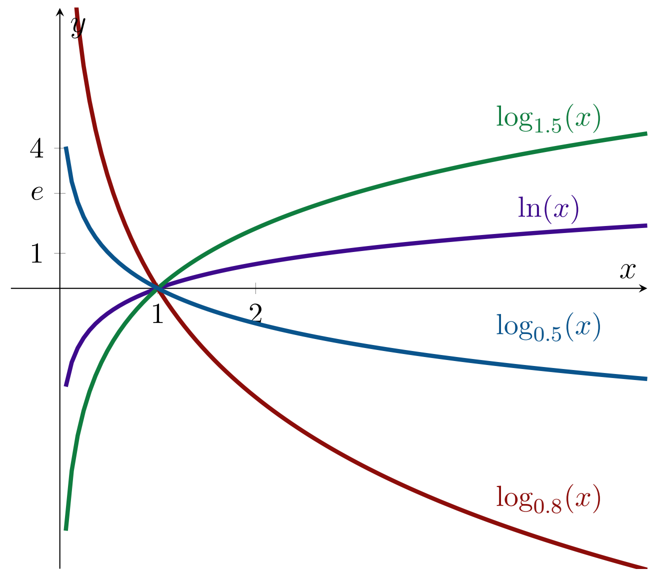
$$\log_a(x) = f^{-1}(x) = \frac{\ln(x)}{\ln(a)},$$

as desired. $\square$

This lemma implies that you don't actually need a $\log_a(x)$ button on your calculator, the $\ln(x)$ button is all you need!

The figures below show the exponential and logarithm functions for various bases.





0.4 Power Functions

We have already encountered functions of the form $f(x) = x$, $f(x) = x^3 + 2x^2$, called polynomials. Polynomials are always defined for all real numbers.

Consider finding the problem of finding the positive solution to

$$f(x) = x^2 = y$$

for some $y \geq 0$, which is the inverse function at y. We know that the solution is $x = +\sqrt{y} = y^{\frac{1}{2}}$, since

$$f(\sqrt{y}) = (\sqrt{y})^2 = \left(y^{\frac{1}{2}}\right)^2 = y.$$

More generally, if we seek the solution of

$$f(x) = x^q = y,$$

we find

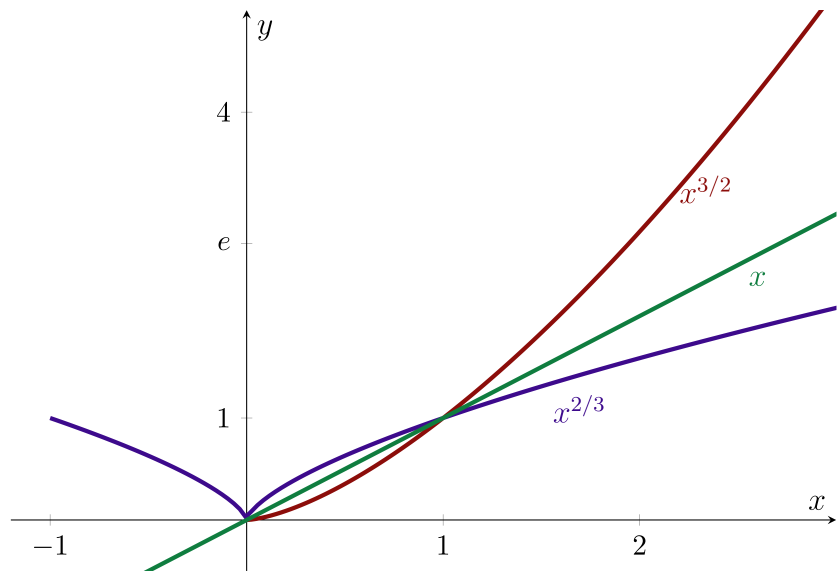
$$y = x^{\frac{1}{q}} = x^p,$$

where $p = \frac{1}{q}$.

As such, we often use the power function with general exponent $p \in \mathbb{R}$, given by

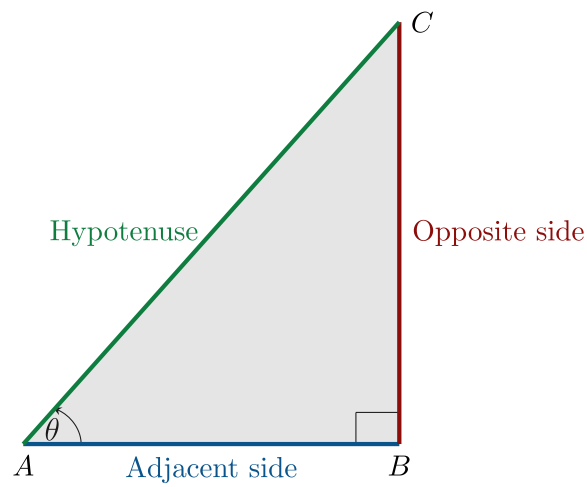
$$f(x) = x^p, \quad x > 0,$$

which is the inverse of the function $x^{1/p}$. The domain and range of the power function depends on p. For instance, if $p = 1/2$, then the domain is $[0, \infty)$ (the square root of a negative number is not defined over the real numbers) while for $p = 1$, the domain is $\mathbb{R}$. The figure below shows the functions $f(x) = x$, $g(x) = x^{2/3} = \sqrt[3]{x^2}$ and $h(x) = x^{3/2} = \sqrt{x^3}$.



0.5 Trigonometric Functions

The trigonometric functions can be defined in a right-angle triangle as shown below. Note the right angle at the corner B, and the angle θ at the corner A.



The sine, cosine and tangent of the angle θ are defined as ratios of the sides of the right angled triangle:

Definition 0.5.1
$\sin(\theta), \cos(\theta), \tan(\theta)$

We define $\sin(\theta)$ as the ratio between the length of the opposite side and the hypotenuse, i.e.,

$$\sin(\theta) = \frac{\text{length opposite side}}{\text{length hypotenuse}},$$

while we define the cosine of the angle θ as the ratio between the length of the adjacent side and the hypothenuse:

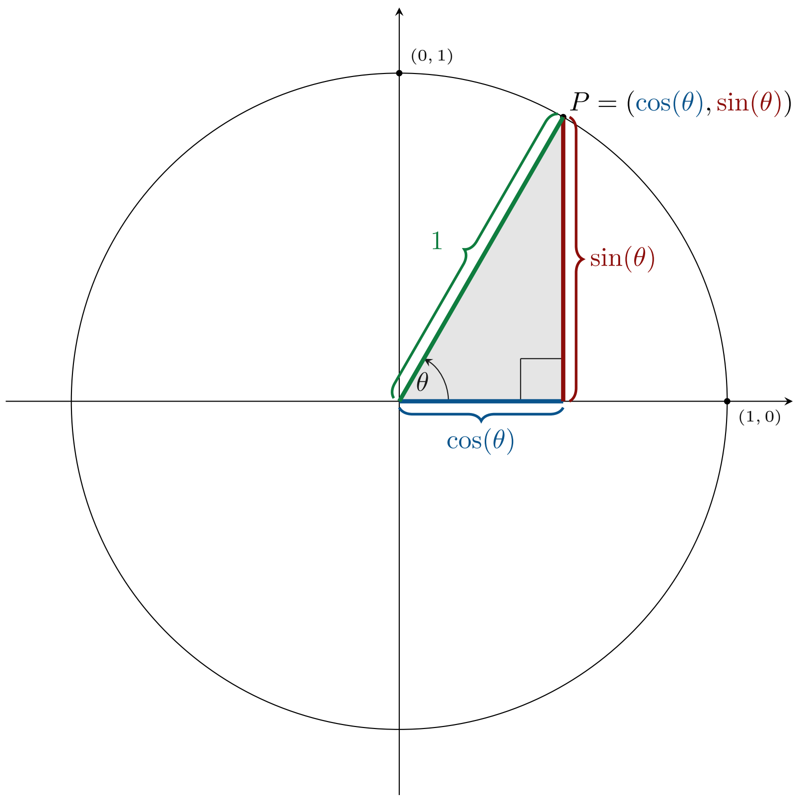
$$\cos(\theta) = \frac{\text{length adjacent side}}{\text{length hypotenuse}}.$$

Finally,

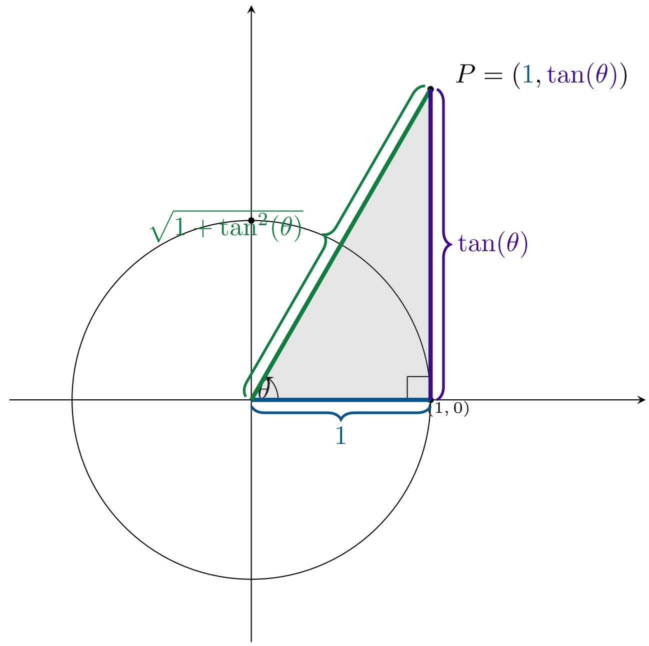
$$\tan(\theta) = \frac{\text{length opposite side}}{\text{length adjacent side}} = \frac{\sin(\theta)}{\cos(\theta)},$$

is the tangent of θ, which is only defined when $\cos(\theta) \neq 0$.

The trigonometric functions are best visualized in a unit circle with radius 1 like the one shown below. Note the grey triangle. Because the radius is unity, the hypotenuse has length 1. As such, $\sin(\theta)$ is exactly the length of the opposite side of θ (divided by 1), while the $\cos(\theta)$ is exactly the length of the adjacent side of θ. As such, the point P on the unit circle which forms angle θ with the x axis has coordinates $(\cos(\theta), \sin(\theta))$.



For the tangent of an angle, we consider the (large) triangle with points $(0, 0)$, $(0, 1)$ and angle θ, as depicted below. Since in this case the adjacent side has length 1, we have $\tan(\theta) = \frac{\text{length opposite side}}{\text{length adjacent side}} = \text{length opposite side}$. This means that the other point of the big triangle is $(1, \tan(\theta))$, and by the Pythagorean theorem, the length of the adjacent side must be $\sqrt{1 + \tan^2(x)}$.



By varying the angle θ, we see that $\sin(\theta)$ and $\cos(\theta)$ are always between $[-1, 1]$, and behave periodically, while the tangent $\tan(\theta)$ can approach $\pm\infty$. This happens whenever $\cos(\theta) = 0$, or geometrically, when the outer large triangle becomes degenerate, i.e., when two of its angles approach the right angle. The curves of $\sin(\theta)$, $\cos(\theta)$ and $\tan(\theta)$ are shown below.

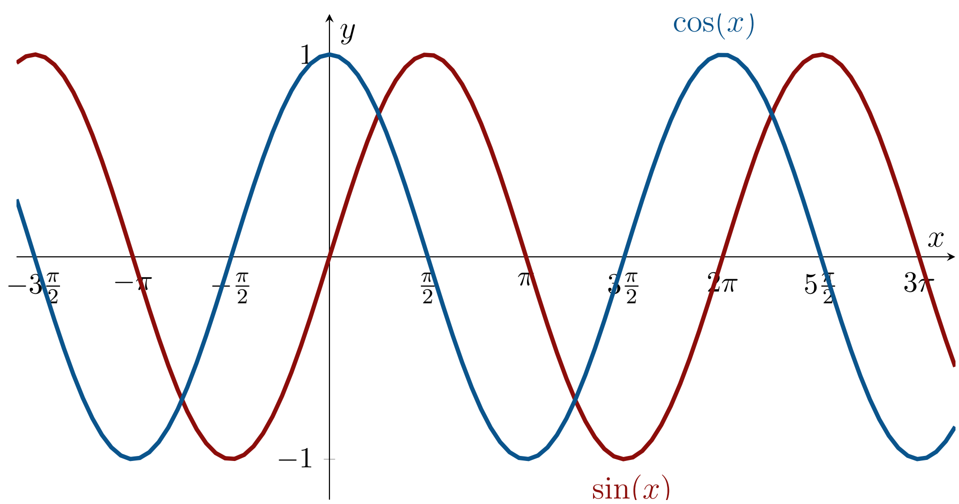


Figure 0.5.2: Graph of the functions $\sin(x)$ and $\cos(x)$.

Due to the periodic nature of these functions, their inverses are not defined on $\mathbb{R}$ (instead, we must restrict the domain to intervals where the function is monotonic). We will postpone this discussion to a later point, and only note that $\sin(x)$ and $\cos(x)$ have domain $\mathbb{R}$ and range $[-1, 1]$, while $\tan(x)$ has domain $\mathbb{R} \setminus \{\ldots, -3\pi/2, -\pi/2, \pi/2, 3\pi/2, \ldots\}$ (i.e., all real numbers where $\cos(x) \neq 0$) and range $\mathbb{R}$.

Finally, we summarize some handy formulas involving the trigonometric functions in the following theorem:

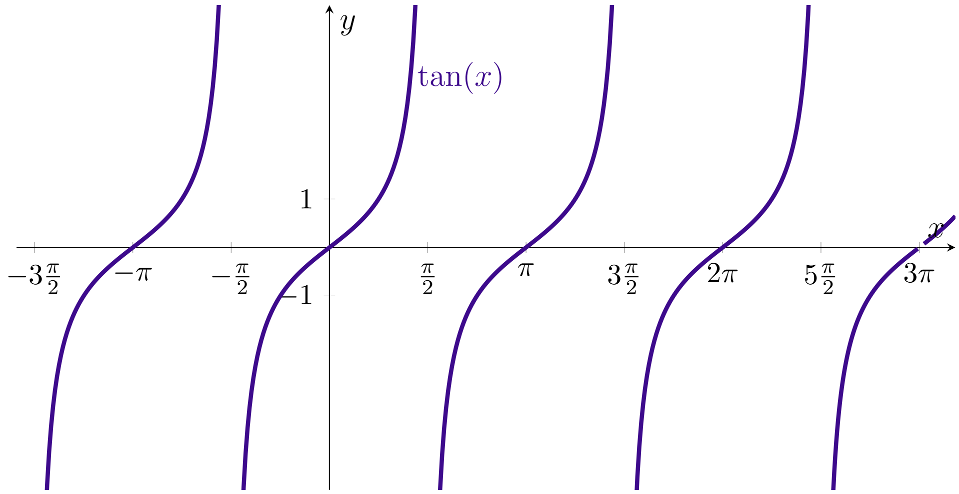


Figure 0.5.3: Graph of the functions $\sin(x)$ and $\cos(x)$.

Theorem 0.5.2 (Properties of Sine and Cosine)

1. **Pythagorean Identity.** Since in a right-angled triangle the squared hypotenuse is the sum of the squares of the adjacent and opposite sides, we have

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

2. **Periodicity.** Since sine and cosine repeat their behaviour, we have

$$\cos(\theta \pm 2\pi) = \cos(\theta),$$
$$\sin(\theta \pm 2\pi) = \sin(\theta).$$

3. **Symmetry.** For any angle $\theta \in \mathbb{R}$, it holds that

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

As such, $\cos(\theta)$ is an even function while $\sin(\theta)$ is an odd function.

4. **Sums and Differences.** The following identities can be useful when evaluating sums (or differences) of angles. Here $\alpha, \beta \in \mathbb{R}$, and we have

$$\cos(\alpha + \beta) = \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta),$$
$$\cos(\alpha - \beta) = \cos(\alpha)\cos(\beta) + \sin(\alpha)\sin(\beta),$$
$$\sin(\alpha + \beta) = \sin(\alpha)\cos(\beta) + \cos(\alpha)\sin(\beta),$$
$$\sin(\alpha - \beta) = \sin(\alpha)\cos(\beta) - \cos(\alpha)\sin(\beta).$$

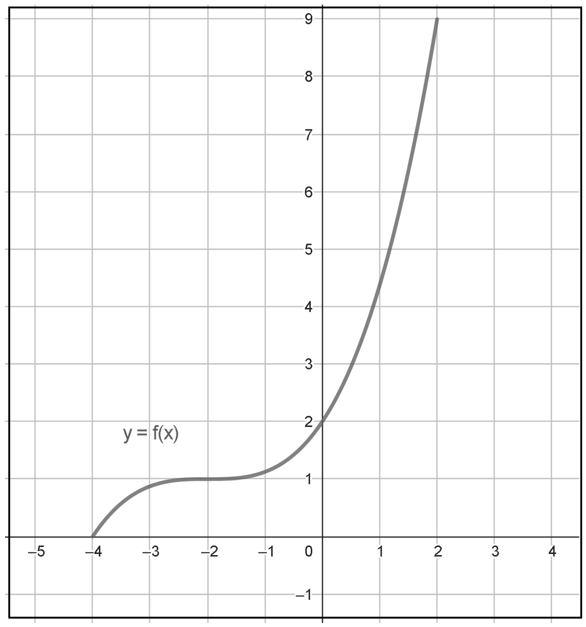
Chapter 0 Problems

0.1 Determine whether $f(x) = \dfrac{x^3}{(x^2+1)^2}$ is even, odd, or neither.

0.2 (a) If $f(x)$ is even and $g(x)$ is odd, show that $h(x) = f(x)g(x)$ is odd.

(b) If $f(x)g(x)$ is odd, is it necessarily true that one of $f(x)$ or $g(x)$ is even while the other is odd? Explain.

0.3 Consider the function f whose graph is shown below.



(a) Explain how you know that f is invertible.

(b) What is $f^{-1}(1)$?

(c) What is $f^{-1}(f^{-1}(2))$?

0.4 Fully factor $x^3 - 6x^2 - 25x - 18$ given that -1 is a root.

0.5 Use polynomial long division to write

$$\frac{2x^3 - x^2 - 4}{x-3} = q(x) + \frac{r(x)}{x-3},$$

where $q(x)$ is a polynomial and $\deg(r(x)) < 1$ (i.e., $r(x)$ is constant).

0.6 Throughout this question, you will make use of the sum and difference identities for sine and cosine.

(a) For all real numbers θ, prove that

$$\sin(2\theta) = 2\sin(\theta)\cos(\theta).$$

(b) For all real numbers θ, prove that

$$\cos(2\theta) = \begin{cases} \cos^2(\theta) - \sin^2(\theta) \\ 2\cos^2(\theta) - 1 \\ 1 - 2\sin^2(\theta) \end{cases}.$$

That is, you must prove each expression on the right-hand side is equal to $\cos(2\theta)$.

(c) Compute the exact value of $\sin\left(\frac{\pi}{12}\right)$ and $\cos\left(\frac{\pi}{12}\right)$. It may help to recall that $\sin\left(\frac{\pi}{3}\right) = \frac{\sqrt{3}}{2}$, $\cos\left(\frac{\pi}{3}\right) = \frac{1}{2}$, and $\sin\left(\frac{\pi}{4}\right) = \cos\left(\frac{\pi}{4}\right) = \frac{\sqrt{2}}{2}$.

0.7 Find all x that satisfy the given equation.

(a) $2^x + 2^{x+1} = 48$

(b) $\log_5(x+4) + \log_5(x) = 1$

(c) $5 = \ln(3 + 2\ln x)$

0.8 If $\ln(a) = 2$, $\ln(b) = 3$, and $\ln\left(\frac{ac}{b^2}\right) = -2$, what is the value of c?

Chapter 1

Sequence Limits

We start our calculus journey by reviewing the absolute value function, which is our tool to measure distances between numbers. Indeed, many problems in mathematics involve approximations, and to measure the quality of an approximation, we must quantify the distance between the approximation and the actual value.

Next, we will define what a sequence of real numbers is, and what it means for such sequence to converge, that is, get closer and closer to a limit.

Sequences are an elementary object in calculus and will appear throughout the course, hence, it is important that the reader becomes comfortable working with them.

By the end of this chapter, you will be able to

- define a sequence and what it means for a sequence to converge;
- compute the limit of a variety of sequences;
- prove statements via mathematical induction.

1.1 Absolute Values

1.1.1 The Absolute Value Function and Its Properties

Absolute values play a crucial role in calculus, as we can use them to measure distances between real numbers, and hence quantify how close they are. What is an absolute value? We often think of it as an operation that removes negative signs:

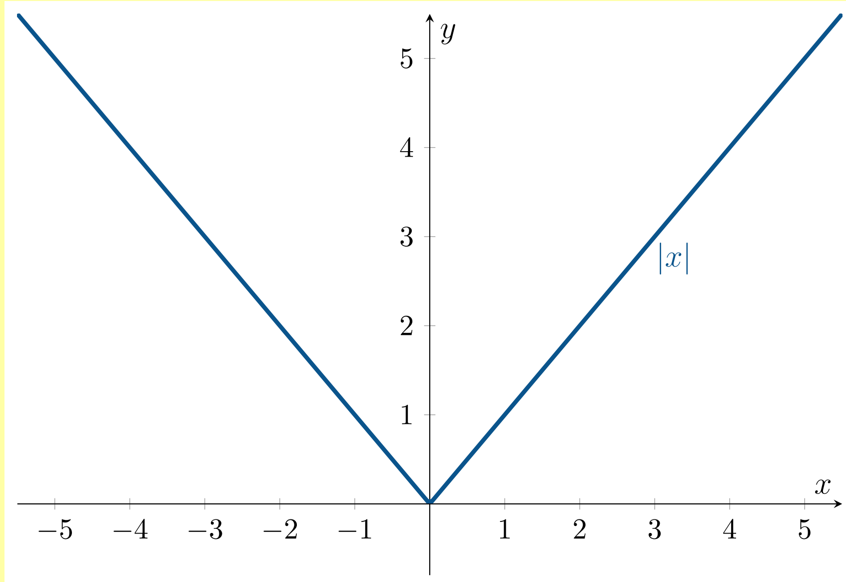
$$|-2| = 2, \quad |-\sqrt{2}| = \sqrt{2}, \quad |0| = 0, \text{ etc.}$$

But is $|-x| = x$ true for all $x \in \mathbb{R}$? Clearly not, for $x = -3$, we have $|-(-3)| = 3 \neq -3$. Let's more formally define what we mean by $|x|$ in the following.

Definition 1.1.1
Absolute Value

For $x \in \mathbb{R}$, we define the **absolute value of x** by

$$|x| = \begin{cases} x & \text{if } x \geqslant 0, \\ -x & \text{if } x < 0. \end{cases}$$



From the definition, we can clearly see that $|-x| = |x|$, and that $|x| \neq x$ for $x < 0$. We also note that the absolute value function is multiplicative, as we show in the following lemma.

Lemma 1.1.2 For any $x, c \in \mathbb{R}$, it holds that

$$|cx| = |c| \cdot |x|.$$

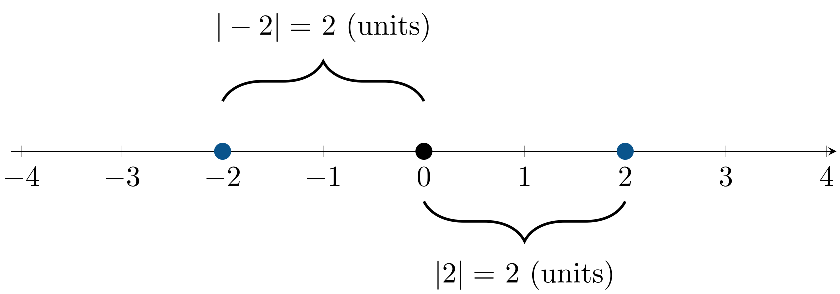
Proof: If $c, x > 0$, $cx > 0$ and $|cx| = cx$. Furthermore, in this case, $|x| = x$, $|c| = c$, so

$$|cx| = cx = |c| \cdot |x|.$$

Similarly, if $c, x < 0$, then $|x| = -x$, $|c| = -c$, and $|cx| = cx = (-c) \cdot (-x) = |c| \cdot |x|$.

The other two cases follow analogously. □

The geometric interpretation of the absolute value function is that it computed the distance from $x \in \mathbb{R}$ to 0. Of course, a distance cannot be negative, which is why the distance from -2 to 0 is $|-2| = -(-2) = 2$, and this is of course the same as the distance from 2 to 0, given by $|2| = 2$.



Generalizing this idea, the absolute value function can also be used to compute the distance between two real numbers x and y, which we address in the following remark.

REMARK

For numbers $a, b \in \mathbb{R}$, the distance between a and b is $|a-b| = |b-a|$. Geometrically, the symmetry $|a-b| = |b-a|$ means that the distance between a and b is the same as the distance between b and a, and being an absolute value, this distance is never negative. This is illustrated in the figure below for $a = 3$ and $b = 7$.

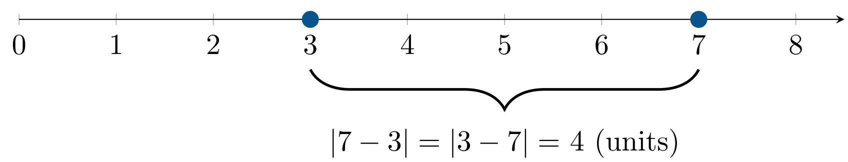


Figure 1.1.1: The numbers $a = 3$, $b - 7$ and distance $|a-b| = |b-a| = 4$ on the real number line.

Let's look at a number of examples. The distance between the numbers 2 and 3 is 1. Similarly, the distance between -2 and 0 is 2. The distance between 3 and 7 is 4, as illustrated in Figure 1.1.1. The distance between -2 and 3 is 5, as we would need 2 steps from -2 to 0 and another 3 steps from 0 to 3. In terms of the absolute value, we can write for these examples

$$
\begin{aligned}
&|2-3| = |3-2| = 1, \\
&|-2-0| = |0-(-2)| = 2, \\
&|3-7| = |7-3| = 4, \\
&|-2-3| = |3-(-2)| = 5
\end{aligned}
$$

1.1.2 Inequalities involving absolute values

A common task in calculus is approximation (for example, in MATH 137, we will develop methods to approximate the roots of a function). When we approximate a solution, however, we generally introduce error. Thus, we want to quantify when an approximation is close enough to the true, exact value. To do this, we need to look at distances (in the form of absolute values) and then solve an inequality.

One of the most fundamental inequalities is the triangle inequality, which, in simple terms, states that the shortest distance between two points is a straight line.

We will state, and prove, the triangle inequality in one dimension, though we will see in later courses that it holds in higher dimensions. The name "triangle inequality" comes from the fact that the sum of any two sides of a triangle is always at least as large as the remaining side. This is illustrated in Figure 1.1.2 below: The distance $|X-Y|$ between X and Y is never larger than the sum of the distances $|X-Z|$ and $|Z-Y|$, that is, $|X-Y| \leq |X-Z| + |Z-Y|$.

That being said, let us now state and prove the Triangle Inequality on the real line.

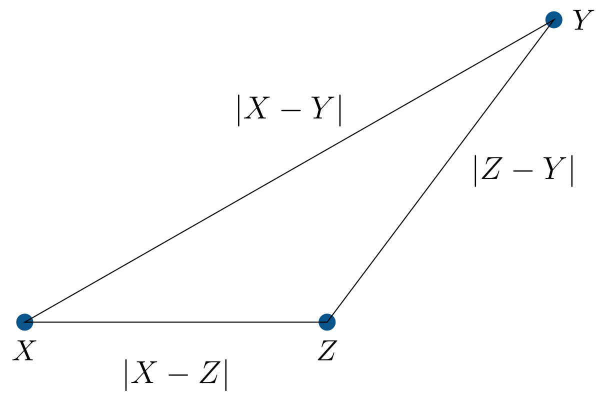


Figure 1.1.2: Illustration of the Triangle Inequality in 2 dimensions.

Theorem 1.1.3 **(Triangle Inequality I)**

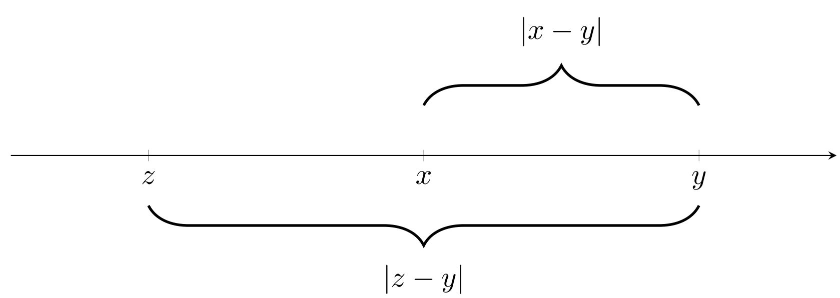
For all $x, y, z \in \mathbb{R}$, it holds that

$$|x-y| \leq |x-z| + |z-y| \,. \tag{1.1}$$

Proof: Since $|x-y| = |y-x|$, we can assume without loss of generality that $x \leq y$. Then, there are three cases for where z can be:

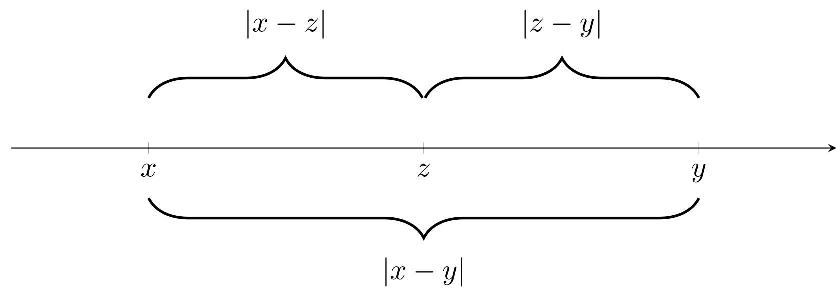
Case 1: $z < x$.

From the figure below, we can clearly see that $|x-y| \leq |z-y|$, so the inequality $|x-y| \leqslant |x-z| + |z-y|$ holds true.



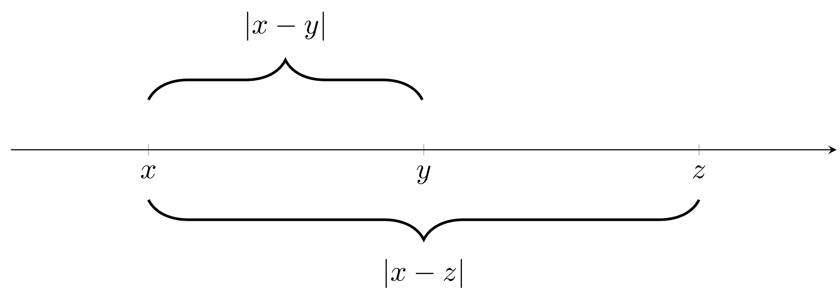
Case 2: $x \leq z \leq y$.

Again, from the illustration below, we can see that we have equality of the distance between x and y, and the sum of the distances between x and z and z and y. Hence, $|x-y| = |x-z| + |z-y|$, which also implies that $|x-y| \leq |x-z| + |z-y|$, as desired.



Case 3: $y < z$.

In this case, it is again easy to see from the picture that $|x-y| < |x-z|$. Hence, $|x-y| \leq |x-z| + |z-y|$ holds true.



$\square$

We can apply Theorem 1.1.3 to prove a useful variant of the Triangle Inequality, as shown in the following Theorem.

Theorem 1.1.4 **(Triangle Inequality II)**

For all $a, b \in \mathbb{R}$,

$$|a+b| \leq |a| + |b|.$$

Proof: Using Triangle Triangle Inequality I in (1.1) with $x = a$, $y = -b$, and $z = 0$ gives

$$\begin{aligned} |x-y| &\leq |x-z| + |z-y| \\ \Rightarrow |a-(-b)| &\leq |a-0| + |0-(-b)| \\ \Rightarrow |a+b| &\leq |a| + |b| \end{aligned}$$

$\square$

When assessing approximations, we will often have to consider inequalities of the form

$$|x-a| < \delta,$$

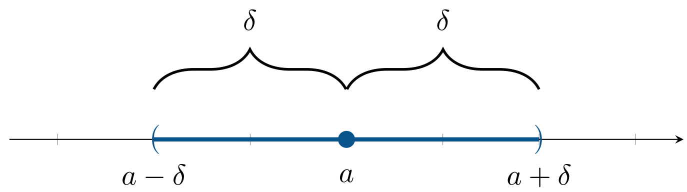


Figure 1.1.3: The solutions of the inequality $|x-a|<\delta$ are all x in the open interval $(a-\delta, a+\delta)$.

where $a \in \mathbb{R}$ and $\delta > 0$. Geometrically, the solutions of this inequality are all real numbers x that have distance from a less than δ, see Figure 1.1.3. It depicts the set of all $x \in \mathbb{R}$ that are less than δ-units away from a.

Of course, points could lie to the left or right of a, which is why showing $|x-a| < \delta$ is equivalent to showing $-\delta < x-a < \delta$. Thus, mathematically,

$$\begin{aligned} |x-a| < \delta &\Leftrightarrow -\delta < x-a < \delta \\ &\Leftrightarrow a-\delta < x < a+\delta. \end{aligned}$$

As such, the inequality $|x-a| < \delta$ describes the interval $(a-\delta, a+\delta)$. Note that the endpoints are not included since this inequality is strict. In contrast, the inequality

$$|x-a| \leq \delta,$$

is equivalent to $-\delta \leq x-a \leq \delta$, so it describes the interval $[a-\delta, a+\delta]$ whose endpoints are included.

Next, let us examine the inequality

$$0 < |x-a| < \delta.$$

Here, the distance cannot be zero, so $x \neq a$. The inequality $0 < |x-a| < \delta$ therefore describes the set $(a-\delta, a+\delta) \setminus \{a\}$, or equivalently

$$(a-\delta, a) \cup (a, a+\delta).$$

Example 1.1.5 Find all solutions of the inequality

$$2 \leq |x-4| < 4.$$

Solution:

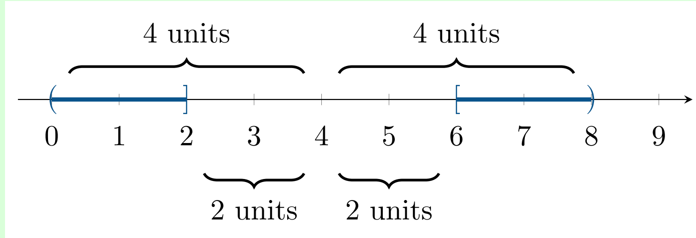
Geometrically, we are looking for all numbers that have distance of at least 2 units and less than 4 unit from 4, as illustrated in the figure below.

Indeed, $2 \leq |x-4| < 4$ means $2 \leq |x-4|$ and $|x-4| < 4$ need to hold simultaneously.

From $2 \leq |x-4|$, we know that either $2 \leq x-4$ or $x-4 \leq -2$, which is equivalent to $x \geq 6$ or $x \leq 2$.

At the same time, however, we need $|x-4| < 4$, which is the case if and only if $-4 < x-4 < 4$ or equivalently $0 < x < 8$.

Overall, we need ($x \geq 6$ or $x \leq 2$) and $0 < x < 8$, so $0 < x \leq 2$ or $6 \leq x < 8$, as depicted below.



We close this section with another example, obtaining the solution of which requires examining several cases.

Example 1.1.6

Find all $x \in \mathbb{R}$ that satisfy
$$|x-1| + |x+2| \geq 4.$$

Solution:

Depending on the value of x, the left hand side changes due to the presence of the absolute value. We distinguish several cases.

Case 1: If $x \geq 1$, then we can omit the absolute value signs, since $x-1 \geq 0$ and $x+2 \geq 0$. Hence, the inequality becomes $x-1+x+2 \geqslant 4$, or $x \geq 3/2$.

Case 2: If $-2 \leqslant x < 1$, then $x-1 < 0$, so by definition of the absolute value, we get $|x-1| = -(x-1) = 1-x$. In this case, we still have $x+2 \geq 0$, so $|x+2| = x+2$. The inequality thus becomes $1-x+x+2 \geqslant 4$ or $3 \geq 4$ which false for any x. Thus, the inequality does not have any solution x with $-2 \leqslant x < 1$.

Case 3: If $x < -2$, $|x-1| = 1-x$ and $|x+2| = -x-2$, so the inequality becomes $1-x+(-x-2) \geqslant 4$, which holds if and only if $x \leq -5/2$.

Overall, the inequality holds for $x \geq 3/2$ or $x \leq -5/2$, or equivalently for
$$x \in (-\infty, -5/2] \cup [3/2, \infty).$$

Section 1.1 Problems

1.1.1. Solve the following equations.

(a) $|x-2| = 5$

(b) $|x-4| = |3x+2|$

1.1.2. For real numbers a and b, write the expression
$$\frac{1}{2}(a+b+|a-b|)$$
as a piecewise function. You should consider 2 cases. What is this expression essentially calculating?

What about the expression

$$\frac{1}{2}(a+b-|a-b|)?$$

1.1.3. Let a and b be real numbers such that $|a^2b-ab^2|=7$, $|a-b|=4$, and $|a|-|b|=3$. Determine $|a|$.

1.1.4. Solve the following inequalities.

(a) $1<|x+1|$

(b) $|x+3|+|1-2x|\leq 5$

(c) $|x-4||x+2|>7$

1.1.5. Let x and y be real numbers.

(a) Prove that $|x|-|y|\leq|x-y|$.

(b) Prove that $||x|-|y||\leq|x-y|$.

Hint: Consider using the triangle inequality.

1.2 Sequences and their Limits

A central object of interest in calculus is a sequence, which is essentially an ordered list of numbers. We can use sequences to approximate continuous processes through discrete data points, which naturally leads to the following definition.

Definition 1.2.1
Infinite Sequence

An **infinite sequence** of real numbers is a list of numbers in a definite order

$$a_1, a_2, a_3, \ldots, a_n, \ldots$$

where $a_i \in \mathbb{R}$ for $i \in \mathbb{N}$. We use the following notations:

$$\{a_1, a_2, a_3, \ldots, a_n \ldots\} \quad \text{or} \quad \{a_n\}_{n=1}^{\infty} \quad \text{or} \quad \{a_n\}.$$

In this course, we understand $\mathbb{N}$ as the set of the natural numbers, $\mathbb{N} = \{1, 2, \ldots\}$ and note that in other courses or texts, the convention might be to include 0 in this set. We also note that a sequence can start at any integer value n, though in most cases in MATH 137, the sequences will start at $n = 1$.

Example 1.2.2

Suppose $0 < r < 1$ is some interest rate (for example, $r = 5\% = 0.05$), and you are investing \$10 000 at that interest rate for exactly one year.

The amount after one year depends on how the interest is compounded. If interest is paid yearly, then after one year we would receive

$$B_1 = 10\,000 \cdot (1 + r) \text{ dollars :}$$

the amount, $10,000$, we invested, plus the interest, $r \cdot 10\ 000$.

Suppose now instead of paying the interest after one year, you get the interest amount in two equally spaced installments. After 6 months, you'd have a total of $(1 + r/2) \cdot 10\,000$ dollars, which is then re-invest for another 6 months. After one year, you'd receive

$$B_2 = 10\,000 \cdot \left(1 + \frac{r}{2}\right)\left(1 + \frac{r}{2}\right) = 10\,000 \cdot \left(1 + \frac{r}{2}\right)^2 \text{ dollars.}$$

Now, instead of semi-annual paid interest, consider daily paid interest. In a similar fashion, we'd get

$$B_{365} = 10\ 000 \cdot \left(1 + \frac{r}{365}\right)\left(1 + \frac{r}{365}\right) \cdot \cdots \cdot \left(1 + \frac{r}{365}\right) = 10\ 000 \left(1 + \frac{r}{365}\right)^{365} \text{ dollars.}$$

Overall, if we divide the year into n subintervals and get paid interest after each of them, after one year, we'd get

$$B_n = 10\ 000 \left(1 + \frac{r}{n}\right)^n \text{ dollars.}$$

dollars. A natural question is to wonder what happens to this sequence as n grows larger and larger, something we will call the limit of this sequence. It turns out that the limit of this sequence is $10\ 000 \cdot e^r$.

Some values of B_n, the balance after one year after n equally spaced interest payments, are given in the table below.

n	$r = 0.05$	$r = 0.1$	$r = 0.15$
1	10 500	11 000	11 500
2	10 506.25	11 025	11 556.25
3	10 508.3796	11 033.7037	11 576.25
6	10 510.5331	11 042.6042	11 596.9342
12	10 511.619	11 047.1307	11 607.5452
55	10 512.4722	11 050.7057	11 615.9705
360	10 512.6745	11 051.5557	11 617.9795
1 000	10 512.6978	11 051.6539	11 618.2117
10 000	10 512.7096	11 051.7037	11 618.3294
100 000	10 512.7108	11 051.7086	11 618.3411
$\vdots$			
10^9	$\approx 10\,000 \cdot e^{0.05}$	$\approx 10\,000 \cdot e^{0.1}$	$\approx 10\,000 \cdot e^{0.15}$

There are various ways to define a sequence. We might define the sequence **explicitly**, that is, a_n is explicitly defined in terms of n, as in the following example.

Example 1.2.3 The following are explicit sequences:

1. The sequence $\left\{\frac{1}{n+1}\right\}_{n=1}^{\infty}$ is $\{1/2, 1/3, 1/4, \dots\}$.
2. The sequence $\{(-1)^n\}_{n=1}^{\infty}$ is alternating between -1 and +1, i.e., $\{-1, 1, -1, 1, \dots\}$.

Alternatively, we might define the sequence **recursively**, that is, a_n is defined in terms of one or more of the previous terms $a_{n-1}, a_{n-2}, \dots$. In this case, we also need to define one or more starting values, depending on the depth of the recursion, which is the number of previous terms a_n depends on.

Example 1.2.4 The Fibonacci sequence, named after a famous Italian mathematician, is defined recursively as follows:

$$F_1 = 1, \quad F_2 = 1, \quad F_n = F_{n-1} + F_{n-2} \text{ for } n \geq 3.$$

That is, the next term is the sum of the two previous terms, which makes this a recursive sequence with depth 2. From the expression, we can see it takes the following values:

$$1, 1, 2, 3, 5, 8, 13, 21, \dots$$

The Fibonacci sequence models, for instance, the (idealized) growth of a rabbit population, where never-dying rabbits reproduce once they reach age one month, and starting their second month, give birth to a new pair of rabbits every month.

Denote by F_n the number of pairs of rabbits at the beginning of the nth month.

- At the beginning of the first month, we have one newly born pair of rabbits, so $F_1 = 1$.
- At the beginning of the second month, that one pair mates, so $F_2 = 1$.
- At the beginning of the third month, the pregnant pair gives birth to a new couple, so $F_3 = 2$, and immediately mate again.
- At the beginning of the fourth month, the (again pregnant) couple gives birth to a new pair, and the pair born at the beginning of the third month mates. Hence, $F_4 = 3$.
- At the beginning of the fifth month, 3 couples already existed, and two pregnant couples give birth, hence $F_5 = 5$.

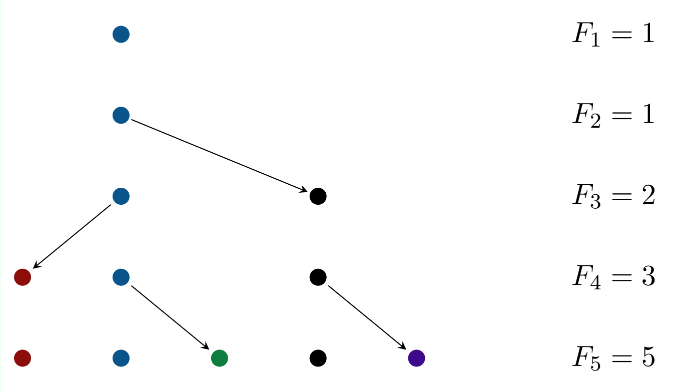
The scheme is visualized in the graphic below, where each dot represents one couple.

In this fashion, we find that at the beginning of the nth month, there will be

- the number of couples alive in the previous month (or, equivalently, at least one month old), given by F_{n-1},
- plus the number of births, which is the same as the number of couples alive two months prior (or equivalently, at least two months old), given by F_{n-2}.

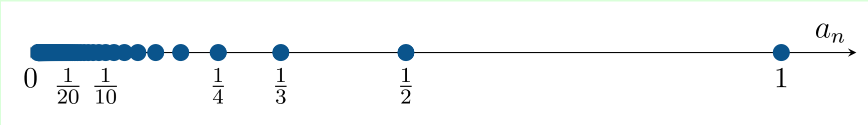
so overall

$$F_n = F_{n-1} + F_{n-2}, \quad n \geq 3.$$



It can be helpful to visualize sequences on a real line of numbers, which we illustrate in the following two examples.

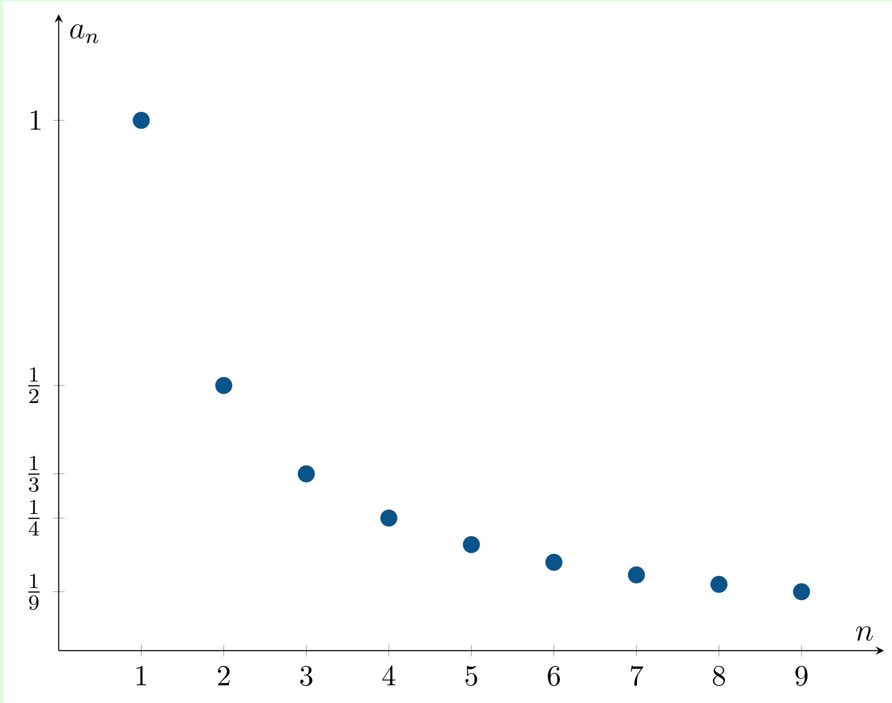
Example 1.2.5 Consider the sequence $\{a_n\}$ with $a_n = \frac{1}{n}$ for $n \in \mathbb{N}$. This sequence takes the values $\left\{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots, \frac{1}{20}, \ldots\right\}$, which is depicted in the figure below. We can see a concentration of points close to 0.



An alternative way to think of a sequence is as a function mapping from the natural numbers to the real numbers, i.e., as a function f with

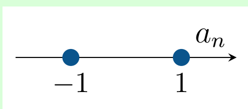
$$f : \mathbb{N} \to \mathbb{R} \quad \text{with} \quad f(n) = a_n.$$

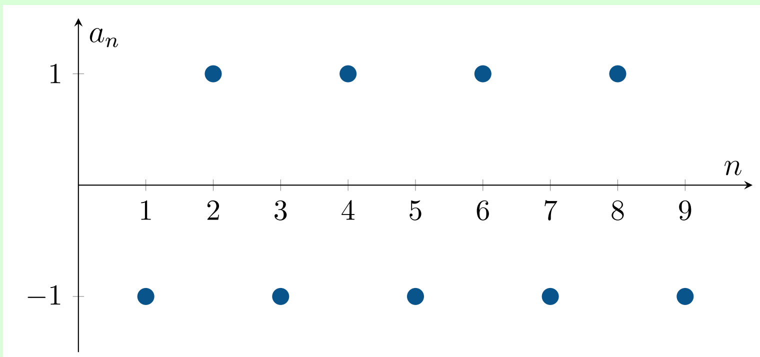
Example 1.2.6 For the sequence $\{1/n\}_{n=1}^{\infty}$, we have $a_n = 1/n$, or $f(n) = 1/n$. We can plot the function f in a 2-dimensional plot, displaying the pairs $(n, f(n))$, as depicted below.



The 2-dimensional plot is typically more informative, as we can see how the values of the sequence evolve for different n, rather than just seeing what values the sequence takes. This is particularly well-illustrated by the sequence in the next example.

Example 1.2.7 Consider the sequence $\{(-1)^n\}$, which alternates between -1 and $+1$. The sequence is displayed below: The 1-dimensional plot on the left just shows two values, while the 2-dimensional plot on the right shows how the sequence is oscillating.





Why are we studying sequences in calculus? As mentioned earlier, lots of continuous processes can be modelled with discrete data. Furthermore, recursive sequences can be used to approximate solutions to equations that cannot be solved explicitly. One such method is Newton's Method, which iteratively approximates the solution of an equation $f(x) = 0$.

A LOOK AHEAD

Consider the sequence given by

$$a_{n+1} = \frac{1}{2}a_n + \frac{1}{a_n}$$

with starting value $a_1 = 1$. The first terms of the sequence are, up to 4 digits, are

$$a_1 = 1, \quad a_2 = 1.5, \quad a_3 = 1.4167, \quad a_4 = 1.4142, \ldots$$

Note that $\sqrt{2} = 1.414213\ldots$, so this recursively defined sequence seems to approximate the number $\sqrt{2}$ with higher and higher accuracy as n increases. In fact, this sequence is an example of Newton's algorithm, and we will learn in Section 3.6 how to construct such sequences.

A natural question to ask is how the sequence $\{a_n\}$ behaves as the index n grows to infinity. For instance, the sequence with $a_n = 1/n$ will approach 0 as n grows larger, and we would think of 0 as the limit of the sequence. In contrast, the alternating sequence with $a_n = (-1)^n$ would not have a limit, as it does not approach one unique value as n grows.

Before making precise what we mean by the limit of a sequence, if it exists, we first study the definition of a subsequence, which will later help us study limiting behaviors of sequences.

Definition 1.2.8
Subsequence

Let $\{a_n\}$ be a sequence of real numbers and $\{n_1, n_2, n_3, \ldots\}$ be a sequence of natural numbers, where $n_1 < n_2 < n_3 < \cdots$. Then the sequence $\{a_{n_1}, a_{n_2}, a_{n_3}, \cdots\} = \{a_{n_k}\}$ is a **subsequence** of $\{a_n\}$.

Example 1.2.9

Consider $\{a_n\}$ with $a_n = (-1)^n$, and recall that this sequence is alternating between -1 and $+1$: $\{-1, 1, -1, \ldots\}$. This sequence naturally splits into two subsequences: the odd-indexed terms, that are all -1, and the even-indexed terms, that are all $+1$. Use Definition 1.2.8 to show this.

Solution:

First, let $n_k = 2k$ for $k \in \mathbb{N}$. Then $\{n_k\}_{k=1}^{\infty}$ is exactly $\{2, 4, 6, \ldots\}$ the sequence of even natural numbers, and the subsequence is

$$a_{n_k} = (-1)^{n_k} = (-1)^{2k} = \left((-1)^2\right)^k = 1, \quad k \in \mathbb{N}.$$

That is, the subsequence that consists only of the a_n with even indices is always 1. Similarly, let $n_k = 2k - 1$ for $k \in \mathbb{N}$, which is the sequence of odd numbers $\{1, 3, 5\ldots,\}$. Then

$$a_{n_k} = (-1)^{n_k} = (-1)^{2k-1} = (-1)^{-1}\left((-1)^2\right)^k = -1, \quad k \in \mathbb{N}.$$

Definition 1.2.10
Tail of a Sequence

Let $\{a_n\}$ be a given sequence and $k \in \mathbb{N}$. The subsequence $\{a_k, a_{k+1}, a_{k+2}, \ldots\}$ is called the **tail** of $\{a_n\}$ **with cutoff** k, and is the sequence $\{a_n\}$ with the first $k-1$ points deleted.

1.2.1 Limits of Sequences

We are now ready to define what the limit of a sequence is. In fact, we will define a number of different limits in this course, including limits of functions, but we will start with exploring the limiting behaviour of a sequence $\{a_n\}$ as n grows to infinity.

For example, we saw earlier that $\left\{\frac{1}{n}\right\}$ seems to converge to 0, or that 0 is the limit of this sequence. This could be seen from the 2-dimensional plot of the function $f(n) = 1/n$, or by substituting larger and larger numbers for n. We will come back to this example later once we are equipped with a formal definition.

Given a sequence $\{a_n\}$, what does it mean to say that $\{a_n\}$ converges to a limit, say $L \in \mathbb{R}$, as n goes to infinity? At first glance, we might say: as n gets larger, the value a_n gets closer to L. But this isn't precise enough: For the sequence $\{1/n\}$, as n grows larger, the sequence gets closer and closer to -1, or to -2, as well! The distinguishing property of the 0 in this example is that the sequence $1/n$ can get arbitrarily close to 0. To account for this, we introduce the following definition for the limit of a sequence:

Definition 1.2.11
Limit of a Sequence

Let $\{a_n\}$ be a sequence and $L \in \mathbb{R}$. We say that L is the **limit** of $\{a_n\}$ if for all $\varepsilon > 0$ there exists a real number N such that if $n > N$, then

$$|a_n - L| < \varepsilon.$$

If such an L exists, we say that $\{a_n\}$ **converges to** L and write

$$\lim_{n\to\infty} a_n = L \quad \text{or} \quad a_n \to L.$$

If no such L exists then we say that $\{a_n\}$ **diverges**.

Figure 1.2.4 below illustrates a convergent sequence with limit L: Here, if $n > N$ then $|a_n - L| < \varepsilon$, or equivalently $a_n \in (L-\varepsilon, L+\varepsilon)$.

Note that our definition 1.2.11 does not require the number N to be a natural number. Some sources give a slightly different definition, requiring N to be a natural number. It is an easy exercise to show that the two definitions are equivalent.

Let's look at some more examples.

Example 1.2.12

Consider $\{a_n\}$ with $a_n = \frac{1}{n^2}$. Intuitively, we guess the limit is $L = 0$, but how do we show this more formally?

Say $\varepsilon = \frac{1}{100}$. Then we need to find a large enough $N \in \mathbb{N}$ so that $\left|\frac{1}{n^2} - 0\right| < \frac{1}{100}$ for all $n > N$. We solve

$$\left|\frac{1}{n^2} - 0\right| < \frac{1}{100} \Leftrightarrow \frac{1}{n^2} < \frac{1}{100} \Leftrightarrow n^2 > 100 \Leftrightarrow n > 10.$$

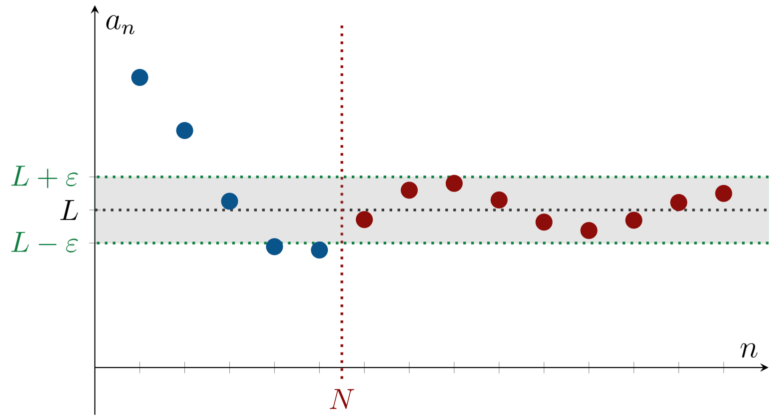


Figure 1.2.4: All points in the sequence with index $n > N$ lie within $(L-\varepsilon, L+\varepsilon)$.

So if we set $N = 10$, then indeed, for all $n > N$, we have the desired inequality $\left|\frac{1}{n^2} - 0\right| < \frac{1}{100}$.

To prove that $L = 0$ is truly the limit, by Definition 1.2.11 we need to show the inequality holds for every $\varepsilon > 0$, not just $\varepsilon = 1/100$. We can easily adapt our work and formally prove that $\lim\limits_{n\to\infty} \frac{1}{n^2} = 0$ as follows:

Proof: Let $\varepsilon > 0$ be given. We determine N so that $|a_n - L| < \varepsilon$ by solving

$$|a_n - L| < \varepsilon \Leftrightarrow \left|\frac{1}{n^2}\right| < \varepsilon \Leftrightarrow \frac{1}{n^2} < \varepsilon \Leftrightarrow n > \frac{1}{\varepsilon}.$$

Thus, for every $\varepsilon > 0$, let $N = \frac{1}{\sqrt{\varepsilon}}$. Then, for all $n > N$,

$$|a_n - L| = \frac{1}{n^2} < \frac{1}{\left(\frac{1}{\sqrt{\varepsilon}}\right)^2} = \varepsilon,$$

as desired. □

REMARK

Note that our definition states that $N \in \mathbb{R}$ though many sources will specify $N \in \mathbb{N}$. However, choosing N to be a natural number may require the use of an awkward ceiling or floor function since, in general, the bound on N is real.

Specifically, in the example above, we found $N = \frac{1}{\sqrt{\epsilon}}$ which is a real number for (most) values of ϵ. If our limit definition had that $N \in \mathbb{N}$, to be precise, we would stipulate $N = \lceil \frac{1}{\sqrt{\epsilon}} \rceil$ and this would ensure that N was a natural number.

Regardless of the details related what number set N belongs to, the take-home message is that no matter the tolerance ϵ, we can always find a cutoff N such that $|a_n - L| < \epsilon$ whenever $n > N$.

The example illustrates the general procedure of proving that L is the limit of a given sequence $\{a_n\}$:

1. Fix some given $\varepsilon > 0$.
2. In a side computation, find the bound N so that $|a_n - L| < \varepsilon$ holds for all $n > N$.
3. Then, show that for all $n > N$, the inequality $|a_n - L| < \varepsilon$ indeed holds.

We will look at another example:

Example 1.2.13 Prove that

$$\lim_{n\to\infty} \frac{n}{2n+3} = \frac{1}{2}.$$

Solution:

Let $\varepsilon > 0$ be given. Let $N = \frac{1}{4}\left(\frac{3}{\varepsilon} - 6\right)$. Then, for all $n > N$

$$\begin{aligned}
|a_n - L| &= \left|\frac{n}{2n+3} - \frac{1}{2}\right| = \left|\frac{-3}{4n+6}\right| = \frac{3}{4n+6} \\
&< \frac{3}{4N+6} \quad \text{since } n > N \\
&= \frac{3}{4\left(\frac{1}{4}\left(\frac{3}{\varepsilon} - 6\right)\right) + 6} \\
&= \varepsilon.
\end{aligned}$$

How did we find N? In a side computation, we find

$$|a_n - L| = \frac{3}{4n+6} < \varepsilon \Leftrightarrow \frac{3}{\varepsilon} < 4n + 6 \Leftrightarrow n > \frac{1}{4}\left(\frac{3}{\varepsilon} - 6\right).$$

Example 1.2.14 Show that

$$\lim_{n\to\infty} \frac{n^2}{3n^2+7n} = \frac{1}{3}.$$

Solution:

Let $\varepsilon > 0$ be given. Let $N = \frac{7}{9\varepsilon}$. Then, for $n > N$ we get:

$$\begin{aligned}
|a_n - L| &= \left|\frac{n^2}{3n^2+7n} - \frac{1}{3}\right| = \left|\frac{3n^2 - 3n^2 - 7n}{9n^2 + 21n}\right| \\
&= \left|\frac{-7n}{9n^2+21n}\right| = \frac{7n}{9n^2+21n} \\
&\leq \frac{7n}{9n^2} \\
&= \frac{7}{9n}
\end{aligned}$$

$$
\begin{aligned}
&< \frac{7}{9N} \quad \text{since } n > N \\
&= \frac{7}{9\frac{7}{9\varepsilon}} = \varepsilon.
\end{aligned}
$$

The previous examples illustrated the crucial steps in proving that $\{a_n\}$ converges to L using the definition of the limit. We need to choose ε arbitrarily, as we need that the sequence can get arbitrarily close to the limit L. Then, we need to show that $|a_n - L| < \varepsilon$ indeed holds for all $n > N$, for our proposed N (which typically depends on ε).

REMARK

In practice, unless you are asked to, we do not recommend you use this formal definition. Instead, we will develop simpler criteria and techniques to show that a sequence $\{a_n\}$ converges to L.

1.2.2 Equivalent Definitions of the Limit

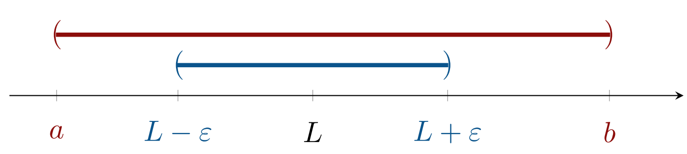
When proving $\lim_{n\to\infty} a_n = L$, we are given $\varepsilon > 0$ and we try to find N so that if $n > N$, the distance between a_n and the limit L satisfies $|a_n - L| < \varepsilon$.

Notice that this inequality is equivalent to $a_n \in (L-\varepsilon, L+\varepsilon)$ for $n > N$. Also, the collection of $\{a_n\}$ for which $n > N$ is the tail of the sequence with cutoff N! Motivated by this, we can give an equivalent definition:

Definition 1.2.15
Limit of a Sequence II

Let $\{a_n\}$ be a sequence and $L \in \mathbb{R}$. Then, $\lim_{n\to\infty} a_n = L$ if for all $\varepsilon > 0$, the interval $(L - \varepsilon, L + \varepsilon)$ contains a tail of the sequence $\{a_n\}$.

We can generalize this even further! Since the above is true for every $\varepsilon > 0$, if we pick any open interval (a, b) containing L, then we can find a small enough ε so that $(L-\varepsilon, L+\varepsilon) \subseteq (a, b)$. This is illustrated in the figure below. Therefore, any open interval containing L also contains a tail of $\{a_n\}$.



Let's collect all of these alternate (but equivalent) definitions together in the following theorem:

Theorem 1.2.16 **(Equivalent Definitions of the Limit of a Sequence)**

The following are equivalent:

1. $\lim_{n\to\infty} a_n = L$.
2. For every $\varepsilon > 0$, the interval $(L-\varepsilon, L+\varepsilon)$ contains a tail of $\{a_n\}$.
3. For every $\varepsilon > 0$, the number of elements of $\{a_n\}$ that do not lie in the interval $(L-\varepsilon, L+\varepsilon)$ is finite.
4. Every interval (a, b) containing L contains a tail of $\{a_n\}$.
5. Given any interval (a, b) containing L, the number of elements of $\{a_n\}$ that do not lie in (a, b) is finite.

In light of this theorem it is also clear that changing finitely many terms of a sequence $\{a_n\}$ does not affect convergence or the limit.

1.2.3 Uniqueness of Limits

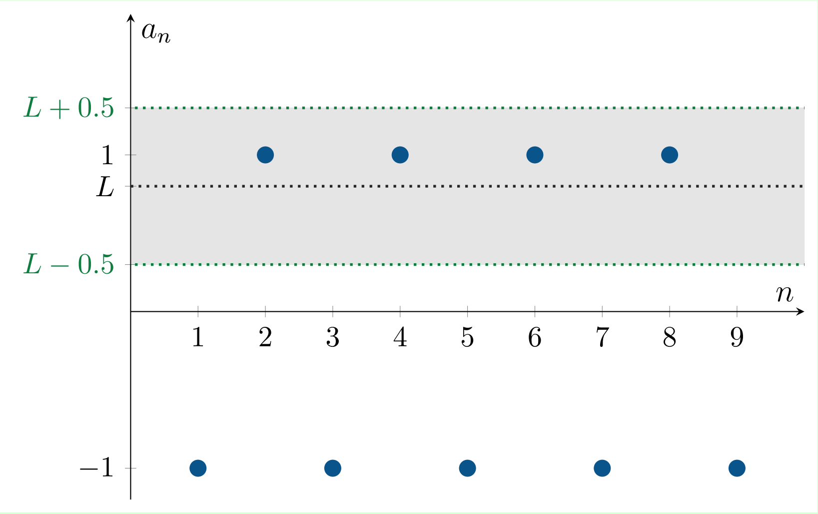
A natural question in mathematics concerns the uniqueness of objects we define. In this context, we might ask: can a sequence have more than one limit? We first examine an example.

Example 1.2.17 Consider $\{(-1)^n\} = \{-1, 1, -1, 1, \cdots\}$. It takes the values $+1$ and -1 infinitely often. Is it possible that both $+1$ and -1 are limits of this sequence? No! First, let's prove -1 is not a limit.

Consider, for instance, the interval $(-2, 0)$ with $-1 \in (-2, 0)$. If -1 were the limit, then the number of elements of $\{(-1)^n\}$ that do not lie in this interval is finite. But $(-2, 0)$ doesn't contain any of the infinitely many terms with $a_n = 1$. Hence, -1 cannot be the limit of this sequence by (5) in Theorem 1.2.16. A similar argument can be used with the interval $(0, 2)$ to show 1 is also not a limit.

This raises the question: Does the sequence $\{(-1)^n\}$ have a limit at all? We can prove it does not!

Proof: Let $\varepsilon = 1/2$, and suppose, for a contradiction, that the sequence converges to some limit $L \in \mathbb{R}$. This implies that the number of elements of $\{(-1)^n\}$ that do not lie in the interval $(L - 1/2, L + 1/2)$ is finite. That is, both $+1$ and -1 must lie in the interval $(L - 1/2, L + 1/2)$. But the interval is only 1 unit long, hence, cannot contain $+1$ and -1 simultaneously for any $L \in \mathbb{R}$. This is illustrated in the figure below: Infinitely many elements of the sequence lie outside the interval $(L - 1/2, L + 1/2)$. As such, the sequence $\{(-1)^n\}$ diverges. $\square$



We can use a similar argument that we used in this example to prove formally that the limit of a sequence, if it exists, is unique.

Theorem 1.2.18 **(Uniqueness of Sequence Limit)**

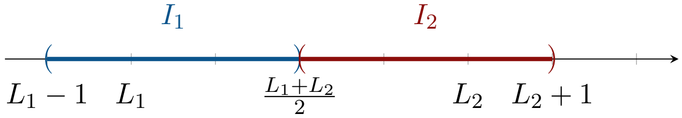
Let $\{a_n\}$ be a sequence. If $\{a_n\}$ has a limit L, then the limit L is unique.

Proof: Suppose for a contradiction that the sequence $\{a_n\}$ has two limits, say L_1, L_2, with $L_1 \neq L_2$. We can, without loss of generality, assume that $L_1 < L_2$.

Let's consider two intervals,

$$I_1 = \left(L_1 - 1, \frac{L_1 + L_2}{2}\right) \quad \text{and} \quad I_2 = \left(\frac{L_1 + L_2}{2}, L_2 + 1\right).$$

Note that the two intervals are disjoint, that is, $I_1 \cap I_2 = \emptyset$.



Since L_1 is, by assumption, a limit of the sequence, only finitely many terms of $\{a_n\}$ are not in I_1.

Similarly, since L_2 is limit of the sequence, only finitely many terms of $\{a_n\}$ are not in I_2.

But the sequence has infinitely many terms! Hence, at least one point must be in both I_1 and I_2. This is a contradiction, since by construction, $I_1 \cap I_2 = \emptyset$. □

REMARKS

If we know that a sequence $\{a_n\}$ can only take certain values, we can infer something about the limit of the sequence, so it exists.

- If $a_n \geqslant 0$ for all n then $\{a_n\}$ cannot converge to a negative number! If it did, say to $L < 0$ then the interval $(L-1, 0)$ would contain L, but not even a single term of the sequence.
- More generally: If $\alpha \leq a_n \leq \beta$ for all n and $\lim_{n\to\infty} a_n = L$ then $\alpha \leq L \leq \beta$.
- But can we also make the same statement using strict inequalities? That is, can we infer from $a_n > 0$ for all $n \in \mathbb{N}$ that then also $\lim_{n\to\infty} a_n = L$ satisfies $L > 0$? No! We have already seen the sequence with $a_n = 1/n$ which satisfies $a_n > 0$ for all $n \in \mathbb{N}$, and yet $\lim_{n\to\infty} a_n = 0$. In general, if $\{a_n\}$ has a limit L and $\alpha < a_n < \beta$ for all n, then $\alpha \leq L \leq \beta$.

1.2.4 Divergence to Infinity

Consider the sequence with $a_n = n$. It is clear that the terms of the sequence are getting larger and larger without bound, so we would conclude that $\lim_{n\to\infty} a_n$ does not exist. That is, $\{a_n\}$ diverges.

But we can say more! The behavior of the sequence $\{n\}$ (which diverges) is different from the behavior of the sequence $\{(-1)^n\}$. We know that the latter is alternating, while we know that the former gets larger without bound. We can make a definition to capture this:

Definition 1.2.19
Divergence to ∞

We say that a sequence $\{a_n\}$ **diverges** to ∞, or $\lim_{n\to\infty} a_n = \infty$, if, for all $m > 0$, we can find $N \in \mathbb{R}$ so that $a_n > m$ for all $n > N$. Equivalently, any interval of the form (m, ∞) contains a tail of $\{a_n\}$.

Similarly, $\lim_{n\to\infty} a_n = -\infty$ if, for all $m < 0$ there is $N \in \mathbb{R}$ so that $a_n < m$ for all $n > N$.

It does look strange to write " $= \infty$ ", but with the above definition we know it means that the limit does not exist, but the terms of the sequence get infinitely large.

Example 1.2.20

Show that
$$\lim_{n\to\infty} 1 - n = -\infty.$$

Solution:

Let $m < 0$. As an aside: $1 - n < m$ if and only if $1 - m < n$. Now, if we let $N = 1 - m$, then for all $n > N$,
$$a_n = 1 - n < 1 - N = 1 - (1 - m) = m,$$
as desired.

Section 1.2 Problems

1.2.1. Use the formal definition of the limit to show that

$$\lim_{n\to\infty} \frac{n}{n+9} = 1.$$

1.2.2. Use the formal definition of the limit to show that

$$\lim_{n\to\infty} \frac{n+(-1)^n}{n} = 1.$$

1.2.3. Use the formal definition of the limit to show that

$$\lim_{n\to\infty} \frac{n^2+1}{2n^2+3} = \frac{1}{2}.$$

1.2.4. Use the formal definition of the limit to show that

$$\lim_{n\to\infty} \frac{\sin(n)}{n^2+6} = 0.$$

1.2.5. Let $a, b, c, d \in \mathbb{R}$ with $c > 0$. Use the formal definition of the limit to prove that

$$\lim_{n\to\infty} \frac{an+b}{cn+d} = \frac{a}{c}.$$

1.2.6. Consider the function

$$f(n) = \begin{cases} 1 & \text{if } n \text{ is a perfect square,} \\ \frac{1}{n} & \text{otherwise.} \end{cases}$$

Prove that $\lim\limits_{n\to\infty} f(n)$ does not exist.

Hint: Assume that $\lim\limits_{n\to\infty} f(n)$ exists and derive a contradiction.

1.2.7. Use the formal definition of the limit to show that

$$\lim_{n\to\infty} \frac{n^2}{3n+1} = \infty.$$

1.3 Arithmetic Rules for Limits

As you are likely aware, using the formal definition of a limit is a tedious and non-trivial task. Thus, if we can avoid using the definition to find the limit L of a sequence $\{a_n\}$ we should do so. There are certain rules that we can derive using the formal definition that will allow us to compute sequence limits more quickly.

For instance, suppose that $\lim_{n\to\infty} a_n = a$ and $\lim_{n\to\infty} b_n = b$ (for $a, b \in \mathbb{R}$). How does the sequence $\{a_n + b_n\}$ behave as $n \to \infty$? Intuitively, for large n, a_n will be close to a and b_n will be close to b, so $a_n + b_n$ should be close to $a + b$. And this is indeed the case. We summarize this, and some other rules, in the following theorem.

Theorem 1.3.1 **(Arithmetic Rules for Limits)**

Let $\{a_n\}$ and $\{b_n\}$ be sequences. Assume $\lim_{n\to\infty} a_n = a$ and $\lim_{n\to\infty} b_n = b$, where $a, b \in \mathbb{R}$.

Then the following rules hold:

1. For any $c \in \mathbb{R}$, if $a_n = c$ for all n then $c = a$.
2. For any $c \in \mathbb{R}$, $\lim_{n\to\infty} ca_n = ca$.
3. $\lim_{n\to\infty} (a_n + b_n) = a + b$.
4. $\lim_{n\to\infty} a_n b_n = ab$.
5. $\lim_{n\to\infty} \frac{a_n}{b_n} = \frac{a}{b}$ if $b \neq 0$.
6. If $a_n \geqslant 0$ for all n and $\alpha > 0$, then $\lim_{n\to\infty} a_n^\alpha = a^\alpha$.
7. For any $k \in \mathbb{N}$, $\lim_{n\to\infty} a_{n+k} = a$.
8. If $\alpha > 0$, then $\lim_{n\to\infty} n^\alpha = \infty$.
9. If $\alpha < 0$ then $\lim_{n\to\infty} n^\alpha = 0$.

Proof: We only show the third property and leave the remainder as exercises.

Let $\varepsilon > 0$ be given. Since $\lim_{n\to\infty} a_n = a$, we can find N_1 So that if $n > N_1$, we have $|a_n - a| < \varepsilon/2$. Similarly, since $\lim_{n\to\infty} b_n = b$, we can find N_2 so that if $n > N_2$ we have $|b_n - b| < \varepsilon/2$. Now, let $N = \max\{N_1, N_2\}$. Then, if $n > N$, we can be sure that both inequalities hold; hence, from the Triangle Inequality II,

$$\begin{aligned} |(a_n + b_n) - (a + b)| &= |(a_n - a) + (b_n - b)| \\ &\leq |(a_n - a)| + |(b_n - b)| \\ &< \varepsilon/2 + \varepsilon/2 = \varepsilon \end{aligned}$$

$\square$

In order to apply the results in this theorem, the limits must exist. For instance, let $a_n = \frac{1}{n}$ and $b_n = (-1)^n$. Then we cannot apply Rule 4 to argue that $\lim\limits_{n\to\infty} \frac{(-1)^n}{n} = 0$, since the limit $\lim\limits_{n\to\infty} (-1)^n$ does not exist. We will develop another method in the next section that will help us deal with this limit!

Let's see how handy these rules are.

Example 1.3.2 Determine the limit of the following sequences:

1. $a_n = \dfrac{3n+7}{n+2}$.
2. $b_n = \dfrac{n^3+n^2+1}{2n^3+7n^2-1}$.
3. $c_n = \dfrac{n+1}{n^2+1}$.

Solution:

1. Using the arithmetic rules for limits, we find
$$\begin{aligned}\lim_{n\to\infty} \frac{3n+7}{n+2} &= \lim_{n\to\infty} \frac{3+\frac{7}{n}}{1+\frac{2}{n}} \\ &= \frac{\lim\limits_{n\to\infty} 3 + \lim\limits_{n\to\infty} \frac{7}{n}}{\lim\limits_{n\to\infty} 1 + \lim\limits_{n\to\infty} \frac{2}{n}} \\ &= \frac{\lim\limits_{n\to\infty} 3 + 7 \lim\limits_{n\to\infty} \frac{1}{n}}{\lim\limits_{n\to\infty} 1 + 2 \lim\limits_{n\to\infty} \frac{1}{n}} \\ &= \frac{3+0}{1+0} = 3,\end{aligned}$$
where in the first step we divided by n in both the numerator and the denominator.
2. As in the previous example, we divide by the largest appearing power of n (here n^3) and find
$$\lim_{n\to\infty} \frac{n^3+n^2+1}{2n^3+7n^2-1} = \lim_{n\to\infty} \frac{1+\frac{1}{n}+\frac{1}{n^3}}{2+\frac{7}{n}-\frac{1}{n^3}} = \frac{1+0+0}{2+0+0}.$$
3. The power in the denominator is larger than the power in the numerator, so the denominator grows faster and we would expect the sequence to converge to zero. Indeed, dividing by n^2 in the numerator and denominator gives
$$\lim_{n\to\infty} \frac{n+1}{n^2+1} = \lim_{n\to\infty} \frac{\frac{1}{n}+\frac{1}{n^2}}{1+\frac{1}{n^2}} = \frac{0+0}{1+0} = 0.$$

REMARKS

Recall Property 5 of Theorem 1.3.1 stated that $\lim_{n\to\infty} \frac{a_n}{b_n} = \frac{a}{b}$ so long as $b \neq 0$. But what if $b = 0$? In this case, anything can happen!

- For instance, the limit can go to a finite (non-zero) number:

$$\lim_{n\to\infty} \frac{\left(\frac{1}{n}\right)}{\left(\frac{1}{n}\right)} = 1$$

- Or can diverge to $\pm\infty$:

$$\lim_{n\to\infty} \frac{\left(\frac{1}{n}\right)}{\left(\frac{1}{n^2}\right)} = \lim_{n\to\infty} \frac{n^2}{n} = \lim_{n\to\infty} n = \infty.$$

- Or can converge to 0:

$$\lim_{n\to\infty} \frac{\left(\frac{1}{n^2}\right)}{\left(\frac{1}{n}\right)} = \lim_{n\to\infty} \frac{1}{n} = 0.$$

These three examples illustrate that the limit $\lim_{n\to\infty} \frac{a_n}{b_n}$ when $\lim_{n\to\infty} b_n = 0$ needs special consideration on an individual basis. There is one thing we can say, though, if we additionally assume that the $\lim_{n\to\infty} \frac{a_n}{b_n}$ exists, as the next theorem shows.

Theorem 1.3.3 If $\lim_{n\to\infty} b_n = 0$ and $\lim_{n\to\infty} \frac{a_n}{b_n}$ exists, then

$$\lim_{n\to\infty} a_n = 0.$$

Proof: Suppose $\lim_{n\to\infty} b_n = 0$, and assume $\lim_{n\to\infty} \frac{a_n}{b_n} = k$, where $k \in \mathbb{R}$.

Then,

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} \frac{a_n}{b_n} \cdot b_n = k \cdot 0 = 0,$$

as desired. □

An easy application of this result is the following: If $\lim_{n\to\infty} b_n = 0$ and $\lim_{n\to\infty} a_n \neq 0$ then $\lim_{n\to\infty} \frac{a_n}{b_n}$ does not exist. For example, $\lim_{n\to\infty} \frac{n^3+3n}{n^2+1}$ does not exist:

$$\lim_{n\to\infty} \frac{n^3+3n}{n^2+1} = \lim_{n\to\infty} \frac{1+\frac{3}{n^2}}{\frac{1}{n}+\frac{1}{n^3}}$$

The numerator converges to 1, while the denominator converges to 0. Hence $\lim_{n\to\infty} \frac{n^3+3n}{n^2+1} = \infty$ and the sequence diverges.

Theorem 1.3.4 If $\lim\limits_{n\to\infty} b_n = \pm\infty$ and $\lim_{n\to\infty} a_n = k$ for some $k \in \mathbb{R}$, then

$$\lim_{n\to\infty} \frac{a_n}{b_n} = 0.$$

We will leave the proof of this theorem as an exercise, but the result should be intuitive. In particular, the denominator grows without bound (either to ∞ or $-\infty$) while the numerator approaches some finite number, k. As a result, the overall limit approaches 0.

Next, let's compute the limit of any ratio of powers of n by using the limit rules!

Example 1.3.5 We compute

$$\lim_{n\to\infty} \frac{b_0 + b_1 n + b_2 n^2 + \cdots + b_j n^j}{c_0 + c_1 n + c_2 n^2 + \cdots + c_k n^k}$$
$$= \lim_{n\to\infty} \frac{n^j}{n^k} \left[\frac{\frac{b_0}{n^j} + \frac{b_1}{n^{j-1}} + \cdots + b_j}{\frac{c_0}{n^k} + \frac{c_1}{n^{k-1}} + \cdots + c_k} \right] = \begin{cases} \frac{b_j}{c_k} & \text{if } j = k, \\ 0 & \text{if } j < k, \\ \infty & \text{if } j > k \text{ and } \quad \frac{b_j}{c_k} > 0, \\ -\infty & \text{if } j > k \text{ and } \quad \frac{b_j}{c_k} < 0. \end{cases}$$

From this rule, we find, for instance,

$$\lim_{n\to\infty} \frac{3n+2}{2n-1} = \frac{3}{2}, \quad \lim_{n\to\infty} \frac{4n^2+5n}{n^3-1} = 0, \quad \lim_{n\to\infty} \frac{7-n^4}{1+n^3} = -\infty.$$

We close this section with two more examples, each showing a trick that can help us determine sequence limits.

Example 1.3.6 Consider the sequence $\{\sqrt{n^2+4} - n\}$. As n approaches ∞, this looks like an $\infty - \infty$ case, and we cannot directly apply our limit rules. But we can re-write the sequence by multiplying the numerator and the denominator by its conjugate:

$$\begin{aligned} \lim_{n\to\infty} \sqrt{n^2+4} - n &= \lim_{n\to\infty} \left(\sqrt{n^2+4} - n\right) \left(\frac{\sqrt{n^2+4} + n}{\sqrt{n^2+4} + n}\right) \\ &= \lim_{n\to\infty} \frac{n^2+4-n^2}{\sqrt{n^2+4}+n} = \lim_{n\to\infty} \frac{4}{\sqrt{n^2+4}+n} \\ &= 0, \end{aligned}$$

where in the last step we used Theorem 1.3.4.

While we will examine recursive sequences in more detail later, our final example shows that if we know that the sequence converges, we can use Property 8 of Theorem 1.3.1 to determine its limit.

Example 1.3.7 Let $\{a_n\}$ with $a_1 = 2$ and $a_{n+1} = \dfrac{5 + a_n}{2}$. Suppose we know it has a limit, say $\lim_{n\to\infty} a_n = L$. Then, using Property 7 of Theorem 1.3.1, we get:

$$L = \lim_{n\to\infty} a_n = \lim_{n\to\infty} a_{n+1} = \lim_{n\to\infty} \frac{5 + a_n}{2} = \frac{5 + L}{2}$$

This means $L = (5 + L)/2$, or $L = 5$.

We can also use the arithmetic rules to show that a limit does not exist as explored in the following example.

Example 1.3.8 Show that $\lim_{n\to\infty} \sin(n)$ does not exist.

Solution: Intuitively, we know that $\sin(x)$ oscillates infinitely between $[-1, 1]$ and this behaviour is echoed in $\sin(n)$ (recall that $\sin(n)$ is defined only at the natural numbers). Thus, we would expect that $\lim_{n\to\infty} \sin(n)$ does not exist. Here is a proof that relies on trig identities.

Proof: Assume (for a contradiction) that $\lim_{n\to\infty} \sin(n) = L$ for some $L \in \mathbb{R}$.

Since $\sin(n) \to L$, then $\sin(n + 1) \to L$ and $\sin(n - 1) \to L$ as well.

Now, using trig identities, we compute

$$\begin{aligned}\sin(n + 1) &= \sin(n)\cos(1) + \cos(n)\sin(1),\\ \sin(n - 1) &= \sin(n)\cos(1) - \cos(n)\sin(1).\end{aligned}$$

Adding the equations together, we get

$$\sin(n + 1) + \sin(n - 1) = 2\sin(n)\cos(1),$$

and taking the limit we find that

$$2L = 2L\cos(1) \Rightarrow \cos(1) = 1,$$

(assuming that $L \neq 0$) which is a contradiction.

If $L = 0$, taking the limit of either trig identity above implies that $\cos(n) \to 0$. That is, we have both $\sin(n) \to 0$ and $\cos(n) \to 0$. Applying the limit to $\sin^2(n) + \cos^2(n) = 1$ results in $0 = 1$ which is a contradiction.

Thus $\lim_{n\to\infty} \sin(n)$ does not exist. □

Section 1.3 Problems

1.3.1. Use the formal definition of the limit to prove the following.

$$\text{If } c \in \mathbb{R} \text{ and } \lim_{n\to\infty} a_n = a \text{ where } a \in \mathbb{R}, \text{ then } \lim_{n\to\infty} ca_n = ca.$$

Hint: Consider the $c = 0$ and $c \neq 0$ cases separately.

1.3.2. Use the formal definition of the limit to prove the following.

$$\text{If } \lim_{n\to\infty} a_n = a \text{ and } \lim_{n\to\infty} b_n = b \text{ where } a, b \in \mathbb{R}, \text{ then } \lim_{n\to\infty} (a_n - b_n) = a - b.$$

1.3.3. Use the formal definition of the limit to prove the following.

$$\text{If } \lim_{n\to\infty} b_n = \infty \text{ and } \lim_{n\to\infty} a_n = a \text{ for some } a \in \mathbb{R}, \text{ then } \lim_{n\to\infty} \frac{a_n}{b_n} = 0.$$

1.3.4. Let $\{a_n\}$ be a sequence.

(a) Use the formal definition of a limit to prove the following.

$$\text{If } \lim_{n\to\infty} a_n = L, \text{ then } \lim_{n\to\infty} |a_n| = |L|.$$

Hint: It may help to recall that $||a| - |b|| \leq |a - b|$ for all $a, b \in \mathbb{R}$, a fact proved in Exercise 1.5.1(b).

(b) Consider the statement

$$\text{If } \lim_{n\to\infty} |a_n| = |L|, \text{ then } \lim_{n\to\infty} a_n = L \text{ or } \lim_{n\to\infty} a_n = -L.$$

Is this statement true or false? If it is true, prove it. If it is false, provide a counterexample.

1.3.5. Let $a_n \geq 0$ for all n. Use the formal definition of the limit to prove the following:

$$\text{If } \lim_{n\to\infty} a_n = L, \text{ then } \lim_{n\to\infty} \sqrt{a_n} = \sqrt{L}.$$

Hint: Consider the $L = 0$ and $L \neq 0$ cases separately.

1.3.6. (**Challenge Problem!**) There are many ways to find the average of two positive numbers a and b. The familiar arithmetic average is $A = (a + b)/2$. The perhaps-less-familiar geometric average is $G = \sqrt{ab}$. Both take a single step to evaluate.

We can, however, use the arithmetic and geometric averages to create an iterative combination called the arithmetic-geometric mean, written as $AGM(a, b)$, and defined as the limit of the coupled, recursively-defined sequences $\{a_k\}$ and $\{b_k\}$ constructed via the following rules:

Initialize:

$$a_0 = a$$
$$b_0 = b$$

Iterate: $(k = 0, 1, 2, \ldots)$

$$a_{k+1} = \frac{a_k + b_k}{2}$$
$$b_{k+1} = \sqrt{a_k b_k}$$

Then a_k and b_k converge to the same limit,

$$\lim_{k\to\infty} a_k = \lim_{k\to\infty} b_k = AGM(a, b).$$

Although recursively-defined sequences are notoriously difficult to analyze (see Section 1.5), we can prove several useful properties of the arithmetic-geometric mean using the tools from this section.

(a) Use the AGM-algorithm definition and the limit rules to show that

$$AGM(a,b) = AGM\left(\frac{a+b}{2}, \sqrt{ab}\right) \tag{1.2}$$

and that for any positive real number λ,

$$AGM(\lambda a, \lambda b) = \lambda AGM(a,b) \tag{1.3}$$

(b) Use part 6a to show that

$$AGM(1,b) = \frac{1+b}{2}\, AGM\left(1, \frac{2\sqrt{b}}{1+b}\right)$$

(c) The $AGM(a,b)$ algorithm converges incredibly quickly – the number of correct decimal places approximately <u>doubles</u> with each iteration. Find the first three terms of the sequences $\{a_1, a_2, a_3\}$ and $\{b_1, b_2, b_3\}$ converging to $AGM(1, 1/\sqrt{2})$, and display your results to 16 decimal places.

(d) The arithmetic-geometric mean $AGM(a,b)$ forms the basis of one of the most powerful algorithms to calculate π, and up until 2009, it was the method-of-choice (last used by Daisuke Takahashi and Yasumasa Kanada to compute 2.58 trillion digits). Denoting the half-distance between the two sequences by $c_{k+1} = \frac{1}{2}(a_k - b_k) = -(a_{k+1} - a_k)$, the iterative scheme is based on the formula by Gauss,

$$\pi = \frac{2AGM^2\left(1, \frac{1}{\sqrt{2}}\right)}{\frac{1}{2} - \sum_{n=1}^{\infty} 2^n c_n^2} \tag{1.4}$$

If we write the partial sum of the series in the denominator of Eq. 1.4 as $s_k = \frac{1}{2} - \sum_{n=1}^{k} 2^n c_n^2$, then the algorithm runs as follows. Initialize:

$$\begin{aligned} a_0 &= 1 \\ b_0 &= \tfrac{1}{\sqrt{2}} \\ s_0 &= \tfrac{1}{2} \end{aligned}$$

Iterate: $(k = 0, 1, 2, \ldots, K-1)$

$$\begin{aligned} a_{k+1} &= (a_k + b_k)/2 \\ b_{k+1} &= \sqrt{a_k b_k} \\ c_{k+1}^2 &= (a_{k+1} - a_k)^2 \\ s_{k+1} &= s_k - 2^{k+1} c_{k+1}^2 \end{aligned}$$

Finally, using $AGM(1, 1/\sqrt{2}) \approx a_{K+1}$,

$$\pi \approx \pi_K = \frac{(a_K + b_K)^2}{2 s_K}$$

Run the algorithm to $K = 3$, and display your result to 16 decimal places. For reference,

$$\pi = 3.14159\ 26535\ 89793\ 23846\ 26433\ 83279\ 50288\ 41971\ 69399\ldots$$

1.4 Squeeze Theorem

As demonstrated earlier, when one of the limits of $\{a_n\}$ and $\{b_n\}$ does not exist, we cannot apply our limit rules such as $\lim\limits_{n\to\infty} a_n b_n = \left(\lim\limits_{n\to\infty} a_n\right)\left(\lim\limits_{n\to\infty} b_n\right)$.

Example 1.4.1

Consider the sequences

$$a_n = \sin(n), \quad b_n = \frac{1}{n}, \quad c_n = a_n \cdot b_n = \frac{\sin(n)}{n} \quad n \in \mathbb{N}.$$

We know that $\lim\limits_{n\to\infty} b_n = 0$ and we have demonstrated in Example 1.3.8 that $\lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} \sin(n)$ does not exist because of the oscillating nature of the sine function.

As such, we are not able to compute $\lim\limits_{n\to\infty} a_n \cdot b_n$ as directly as $(\lim\limits_{n\to\infty} a_n)(\lim\limits_{n\to\infty} b_n)$, because $\lim\limits_{n\to\infty} a_n$ does not exist.

Nevertheless, by plugging in large values of n, it seems that the sequence c_n does converge to zero. We will equip ourselves with a tool to show this, and revisit this example below.

An alternative to our limit rules that can help us show convergence is the Squeeze Theorem:

Theorem 1.4.2

(Squeeze Theorem for Sequences)

If $a_n \le b_n \le c_n$ for $n \in \mathbb{N}$ such that $n > M$ (for some $M \in \mathbb{R}$) and $\lim\limits_{n\to\infty} a_n = L = \lim\limits_{n\to\infty} c_n$, then $\lim\limits_{n\to\infty} b_n = L$ as well.

Proof: Let $\varepsilon > 0$ be given. Since $a_n \to L$ and $c_n \to L$, we can find a number N with $N > M$ such that if $n > N$ then $a_n \in (L-\varepsilon, L+\varepsilon)$ and $c_n \in (L-\varepsilon, L+\varepsilon)$. Then, for $n > N$, we have $L - \varepsilon < a_n \le b_n \le c_n < L + \varepsilon$, so $b_n \in (L-\varepsilon, L+\varepsilon)$ which means $\lim\limits_{n\to\infty} b_n = L$ as desired; see Figure 1.4.5 below for an illustration. □

This theorem is particularly appealing when dealing with a product of two sequences one of which is bounded. For instance, if we wish to compute $\lim\limits_{n\to\infty} a_n b_n$ where $\{a_n\}$ is bounded and $\{b_n\}$ converges to 0, then we are well set-up to use the Squeeze Theorem to conclude that $\lim\limits_{n\to\infty} a_n b_n = 0$.

We have motivated this subsection with the sequence $\{\sin(n)/n\}$, which we will now revisit.

Example 1.4.3

Consider again the sequence $b_n = \dfrac{\sin(n)}{n}$. We know that $-1 \le \sin(x) \le 1$ holds for all $x \in \mathbb{R}$, hence $-1 \le \sin(n) \le 1$ for all n. Dividing all parts of the inequality by n, we get

$$\underbrace{-\frac{1}{n}}_{=a_n} \le \underbrace{\frac{\sin(n)}{n}}_{=b_n} \le \underbrace{\frac{1}{n}}_{=c_n}.$$

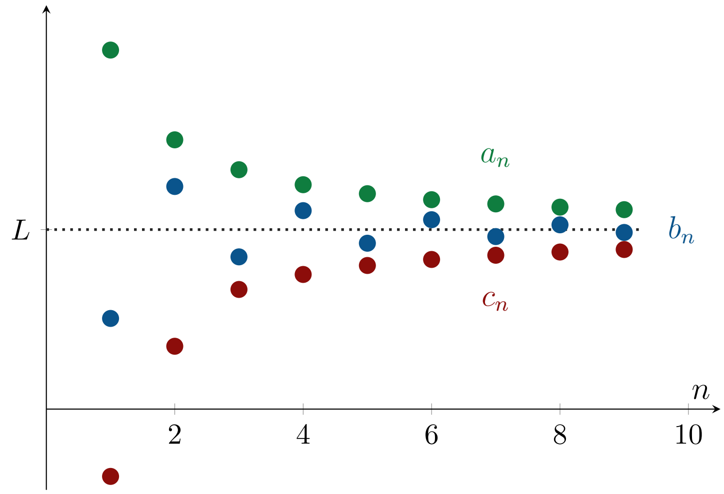
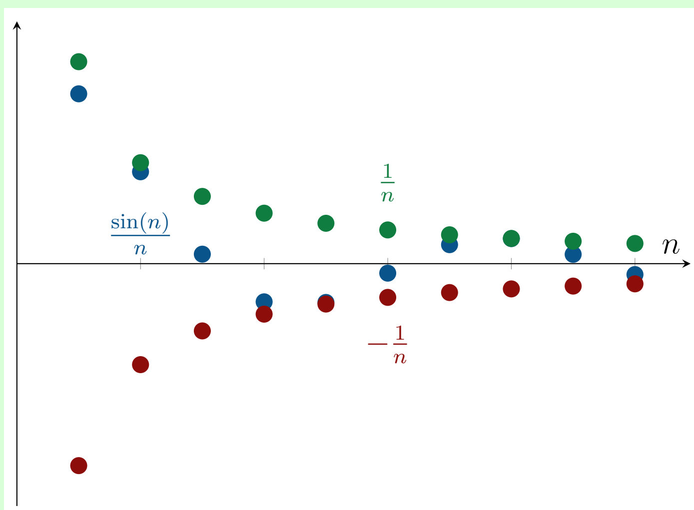


Figure 1.4.5: Illustration of the Squeeze Theorem: Note how the sequence b_n lies between a_n and c_n.

for all $n \in \mathbb{N}$. Since $\lim_{n\to\infty} a_n = \lim_{n\to\infty} c_n = 0$, the Squeeze Theorem implies that

$$\lim_{n\to\infty} b_n = \lim_{n\to\infty} \frac{\sin(n)}{n} = 0.$$

The figure below shows the sequences $a_n = -\frac{1}{n}$, $b_n = \frac{\sin(n)}{n}$ and $c_n = \frac{1}{n}$. Note how the sequence b_n is "squeezed" between a_n and c_n.



Example 1.4.4 Use the Squeeze Theorem to prove convergence of the following sequences:

1. $a_n = \dfrac{(-1)^n}{n^2+1}$,
2. $b_n = \dfrac{\cos\left(n^2+7\right)+7}{n} + 1$.

Solution:

1. We notice that
$$\frac{-1}{n^2+1} \leq \frac{(-1)^n}{n^2+1} \leq \frac{1}{n^2+1}$$
and
$$\lim_{n\to\infty} \frac{-1}{n^2+1} = 0 = \lim_{n\to\infty} \frac{1}{n^2+1},$$
hence, $\lim\limits_{n\to\infty} \frac{(-1)^n}{n^2+1} = 0$ by the Squeeze Theorem.

2. Recall that $-1 \leq \cos(x) \leq 1$ for all $x \in \mathbb{R}$. Hence, $-1 \leq \cos\left(n^2+7\right) \leq 1$, which implies $6 \leq \cos\left(n^2+7\right) + 7 \leq 8$ for all $n \in \mathbb{N}$. We therefore find the bounds
$$\frac{6}{n} + 1 \leq \frac{\cos\left(n^2+7\right)+7}{n} + 1 \leqslant \frac{8}{n} + 1,$$
and both the left-hand side and right-hand side converge to 1 as
$$\lim_{n\to\infty} \frac{6}{n} + 1 = \lim_{n\to\infty} \frac{8}{n} + 1 = 0 + 1 = 1.$$
By the Squeeze Theorem, we conclude $\lim\limits_{n\to\infty} \dfrac{\cos\left(n^2+7\right)+7}{n} = 1.$

REMARK

The squeeze theorem also applies when the bounding sequences $\{a_n\}$ and $\{b_n\}$ diverge. That is, if $a_n \leq b_n \leq c_n$ for all $n > M$ and $\lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} c_n = \infty$ (or $-\infty$), then $\lim\limits_{n\to\infty} b_n = \infty$ (or $-\infty$, respectively) as well.

Example 1.4.5

Consider the sequence $\{b_n\}$ with $b_n = n + \sin(n)$ for $n \in \mathbb{N}$. Does the sequence $\{b_n\}$ converge?

Solution:

Since $|\sin(n)| \leq 1$ for all $n \in \mathbb{N}$, we have
$$n - 1 \leq n + \sin(n) \leq n + 1$$
for all $n \in \mathbb{N}$. Since $\lim\limits_{n\to\infty}(n-1) = \lim\limits_{n\to\infty}(n+1) = \infty$, the sequence $\{b_n\}$ diverges to ∞ as well.

Section 1.4 Problems

1.4.1. Evaluate the following limits.

(a) $\displaystyle\lim_{n\to\infty} \frac{(-1)^n \cos(n)}{n}$

(b) $\displaystyle\lim_{n\to\infty} \frac{n + \sin(n)\cos(n)}{n}$

(c) $\displaystyle\lim_{n\to\infty} \frac{|n\sin(n) - 1|}{n^2+1}$

1.4.2. If $\{a_n\}$ and $\{b_n\}$ are sequences with $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = 0$, evaluate

$$\lim_{n\to\infty} \left(a_n \sin(n) + b_n \cos(n)\right).$$

1.4.3. Evaluate $\displaystyle\lim_{n\to\infty} \frac{n!}{n^n}$.

Hint: Write $n! = 1 \cdot 2 \cdot 3 \cdots n$ and $n^n = n \cdot n \cdot n \cdots n$ and use the fact that $0 < \frac{n!}{n^n}$ for all $n \in \mathbb{N}$.

1.4.4. Use the Squeeze Theorem to show that $\lim_{n\to\infty} (2^n + 3^n)^{1/n} = 3$.

1.4.5. Use the Squeeze Theorem to prove that if $\lim_{n\to\infty} |a_n| = 0$, then $\lim_{n\to\infty} a_n = 0$.

1.4.6. Let $\{a_n\}$ be a sequence and M a real number such that $|a_n - n| \leq M$ for all $n \in \mathbb{N}$. Evaluate

$$\lim_{n\to\infty} \frac{a_n}{n}.$$

1.4.7. Let $A_1 \geq A_2 \geq \cdots \geq A_k \geq 0$ where $k \in \mathbb{N}$ and define $S_n = (A_1^n + A_2^n + \ldots A_k^n)^{1/n}$. In this problem, we will use the squeeze theorem to evaluate the limit

$$\lim_{n\to\infty} S_n = \lim_{n\to\infty} (A_1^n + A_2^n + \ldots + A_k^n)^{1/n}$$

(a) Consider $(A_1, A_2, A_3) = (7, 5, 2)$ and use a calculator to compute $S_n = (A_1^n + A_2^n + A_3^n)^{1/n}$ for $n = 5$ and $n = 50$ to 3 decimal places. What does this suggests for a lower bound on S_n and what does this suggest for the limit of S_n?

(b) Consider $(A_1, A_2, A_3) = (2.1, 2, 1.9)$ and use a calculator to compute $S_n = (A_1^n + A_2^n + A_3^n)^{1/50}$ for $n = 5$ and $n = 50$ to 3 decimal places. What does this suggests for a lower bound on S_n and what does this suggest for the limit of S_n?

(c) Based on the computations above, bound S_n above and below by appropriate bounds and compute the limit using the squeeze theorem.

1.5 Recursive Sequences

We have encountered recursive sequences, such as the Fibonacci Sequence in Example 1.2.4 which clearly diverges to ∞. We've also seen in Example 1.3.7 how we can use limit rules to determine the limit of a recursive sequence, if that limit exists. This raises the question: How do we determine if a recursive sequence converges?

Before we describe how to determine if a recursive sequence converges, we first introduce some terminology.

Definition 1.5.1
Upper and Lower Bound

Let $S \subseteq \mathbb{R}$. We say that α is an **upper bound** of S if $x \leq \alpha$ for all $x \in S$. We call such a set **bounded above**.

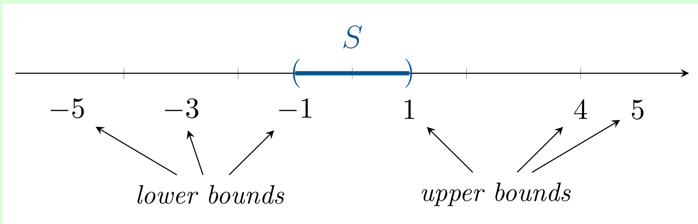
Similarly, β is a **lower bound** of S if $x \geq \beta$ for all $x \in S$. In this case S is **bounded below**.

We call S **bounded** if it is both bounded above and bounded below.

For a bounded set S, we can find $m \in \mathbb{R}$ so that $S \subseteq [-m, m]$. The choice of m, and the choice of a lower/upper bound in general, is not unique.

Example 1.5.2

The set $S = (-1, 1)$ has, for instance, 4 as an upper bound, and -3 as a lower bound, as depicted on the number line below. We also have that $S \subseteq [-5, 5]$ or $S \subseteq [-7, 7]$ or $S \subseteq [-1, 1]$ so S is bounded.



Definition 1.5.3
Least Upper Bound (Supremum)

Let $S \subseteq \mathbb{R}$. Then α is called the **least upper bound** of S if:

1. α is an upper bound, and
2. α is the smallest upper bound, that is, if α' is another upper bound of S, then $\alpha' \geq \alpha$

We often refer to the least upper bound of S as $\text{lub}(S)$. This is also called the **supremum** of S, or $\sup(S)$.

Very similarly, we can define the greatest lower bound:

Definition 1.5.4
Greatest Lower Bound (Infimum)

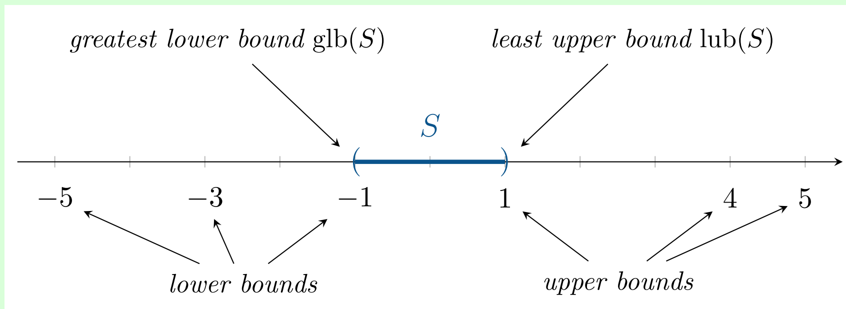
Let $S \subseteq \mathbb{R}$. Then β is called the **greatest lower bound** of S if:

1. β is a lower bound, and
2. β is the largest lower bound, that is, if β' is another lower bound of S, then $\beta' \leq \beta$

We often refer to the greatest lower bound of S as $\mathrm{glb}(S)$. This is also called the **infimum** of S, or $\inf(S)$.

Example 1.5.5

If $S = (-1, 1)$, then $\mathrm{glb}(S) = -1$ and $\mathrm{lub}(S) = 1$, as illustrated in the figure below. Note that in this example, neither $\mathrm{glb}(S)$ nor $\mathrm{lub}(S)$ are elements of S, i.e., $\mathrm{glb}(S) \notin S$ and $\mathrm{lub}(S) \notin S$. However, if we consider the set $S' = (-1, 1]$, then we still have $\mathrm{glb}(S') = -1 \notin S$, but $\mathrm{lub}(S') = 1 \in S$.



The previous example illustrates that the greatest lower bound and least upper bound of a set, if they exist, need not be in the set.

We make note of one axiom of the real numbers that we will need when discussing the Monotone Convergence Theorem.

Axiom 1

If $S \subseteq \mathbb{R}$ is nonempty and bounded above, then S has a least upper bound. Similarly, if S is bounded below, then S has a greatest lower bound.

An important class of sequences are monotonic sequences, which we define formally as follows:

Definition 1.5.6

Increasing and Decreasing Sequences

We say that a sequence $\{a_n\}$ is:

- **increasing** if $a_n \leq a_{n+1}$ for all $n \in \mathbb{N}$.
- **strictly increasing** if $a_n < a_{n+1}$ for all $n \in \mathbb{N}$.
- **decreasing** if $a_n \geq a_{n+1}$ for all $n \in \mathbb{N}$.
- **strictly decreasing** if $a_n > a_{n+1}$ for all $n \in \mathbb{N}$.

Finally, a sequence $\{a_n\}$ is **monotonic** (or **monotone**) if it is either decreasing or increasing.

REMARK

A sequence $\{a_n\}$ can be increasing and decreasing at the same time: In this case,

$$a_n \leq a_{n+1} \leq a_n \quad \text{for all } n \in \mathbb{N},$$

which means

$$a_n = a_{n+1} \quad \text{for all } n \in \mathbb{N}.$$

But this implies that

$$a_n = c \quad \text{for all } n \in \mathbb{N},$$

where $c \in \mathbb{R}$ is a real constant. Hence, a sequence that is increasing and decreasing at the same time must be a constant sequence.

The following theorem allows us to conclude whether monotonic sequences converge and is an important tool we will use for recursive sequences.

Theorem 1.5.7 **(Monotone Convergence Theorem (MCT))**

Let $\{a_n\}$ be an increasing sequence.

1. If $\{a_n\}$ is bounded above, then $\{a_n\}$ converges to $L = \operatorname{lub}(\{a_n\})$.
2. If $\{a_n\}$ is not bounded above, then $\{a_n\}$ diverges to ∞.

Similarly, let $\{b_n\}$ be a decreasing sequence.

1. If $\{b_n\}$ is bounded below, then $\{b_n\}$ converges to $L = \operatorname{glb}(\{b_n\})$.
2. If $\{b_n\}$ is not bounded below, then $\{b_n\}$ diverges to $-\infty$.

Proof: We will prove the case where $\{a_n\}$ is increasing and bounded above. The other case is similar.

Suppose $\{a_n\}$ is increasing.

1. Suppose $\{a_n\}$ is bounded above and let $L = \operatorname{lub}(\{a_n\})$. Let $\varepsilon > 0$ be given. Then $L - \varepsilon < L$, hence, $L - \varepsilon$ is *not* an upper bound of $\{a_n\}$ (since L is the *least* upper bound). Therefore, there exists $M \in \mathbb{N}$ such that if $L - \varepsilon < a_M$ (i.e., since $L - \varepsilon$ is not an upper bound, there must be at least one term, a_M, that is *larger* than $L - \varepsilon$). Further, since $\{a_n\}$ is increasing, this means that

$$L - \varepsilon < a_M \leq a_n \leq L \quad \text{for all } n > M.$$

But this also implies that if $n > M$, then $L - \varepsilon < a_n \leqslant L < L + \varepsilon$ so the tail of $\{a_n\}$ is in $(L - \varepsilon, L + \epsilon)$. This means $\lim\limits_{n\to\infty} a_n = L$.

2. Suppose now $\{a_n\}$ is not bounded above. Let $m \in \mathbb{R}$ be given. We can find $M \in \mathbb{N}$ so that $m < a_M$. Then, if $n > M$, we have

$$m < a_M \leq a_n$$

because $\{a_n\}$ is increasing. This shows $\lim_{n\to\infty} a_n = \infty$.

□

Our strategy for recursive sequences will be to show that a sequence is monotonic and bounded (either above or below) and this, combined with Rule 8 of the Arithmetic Rules for Sequence Limits in Theorem 1.3.1, will allow us to determine the limit. This raises the question of how can we show that a recursive sequence is monotonic and bounded? For that, we introduce *mathematical induction.*

1.5.1 Mathematical Induction

Before we can use the MCT, we need to develop a proof technique called **Mathematical Induction**, a technique often used to show properties that need to hold for all natural numbers. Induction allows us to prove an infinite number of related statements! In fact, it helps us prove a sequence of statements P_n for $n \in \mathbb{N}$. For instance,

$$\text{Statement } P_n : 1 + 2 + \cdots + n = \frac{n(n+1)}{2}, \quad n \in \mathbb{N}$$

is a statement giving the formula for the sum of the first n natural numbers, for each $n \in \mathbb{N}$.

Method
Mathematical Induction

Suppose we have a sequence of statements $P_1, P_2, P_3, \ldots, P_n, \ldots$, where $n \in \mathbb{N}$. If we can:

1. Prove P_1 is true, **(base case)**
2. Prove: If P_k is true for some $k \in \mathbb{N}$ **(inductive hypothesis)**, then P_{k+1} is true, **(inductive step)**,

then we can conclude that P_n is true for all $n \in \mathbb{N}$.

You can think of mathematical induction as a sequence of dominoes falling: If the first one falls, and each domino can cause the next one to fall, then all dominoes will fall. Similarly, if the first statement P_1 is true, and any statement P_k being true causes P_{k+1} to be true, then all statements will be true. You will explore Mathematical Induction further in MATH 135.

Example 1.5.8

As an example, we show the formula for the sum of the first n integers:

$$P_n : 1 + 2 + \cdots + n = \frac{n(n+1)}{2}, \quad n \in \mathbb{N}.$$

- Base Case: For $n = 1$, we indeed find

$$1 = \frac{1 \cdot 2}{2},$$

so the statement P_1 holds.

- Inductive Hypothesis: Suppose that the statement P_k holds, i.e., assume that

$$1 + 2 + \cdots + k = \frac{k(k+1)}{2}.$$

- Inductive Step: We show, using the inductive hypothesis, that the statement P_{k+1} holds for some $k \in \mathbb{N}$.

$$\begin{aligned} 1 + 2 + \cdots + k + k + 1 &= (1 + 2 + \cdots + k) + (k+1) \\ &= \frac{k(k+1)}{2} + \frac{2(k+1)}{2} \\ &= \frac{(k+1)(k+2)}{2} \end{aligned}$$

where in the second line, we have used the inductive hypothesis. This means that the statement P_{k+1} holds as well.

By the principle of mathematical induction, we find that the statement P_n holds for all $n \in \mathbb{N}$.

In this course, we will use the MCT and induction to find limits of recursive sequences. To do this, we follow these steps:

Method

Recursive Sequence Limits via MCT

Here is our step-by-step algorithm for finding the limit of a recursive sequence.

1. Prove the sequence is monotonic.
2. Prove the sequence is bounded (above or below depending on if the sequence is increasing or decreasing).
3. Conclude the sequence converges by MCT.
4. Find the limit using Property 7 from Theorem 1.3.1.

Example 1.5.9

Consider the sequence $\{a_n\}$ recursively defined via

$$a_1 = 1, \quad a_{n+1} = \frac{3 + a_n}{2} \quad (n \geqslant 1).$$

We prove that this sequence converges and find its limit.

1. **Monotonicity**. We inspect the first few terms:

$$a_1 = 1, \quad a_2 = 2, \quad a_3 = 2.5, \quad a_4 = 2.75, \quad \ldots$$

so it seems that the sequence $\{a_n\}$ is increasing. We will prove this by Mathematical Induction as follows:

Claim: $a_n \le a_{n+1}$ for all $n \in \mathbb{N}$.

Proof:

- Base Case: Since $a_1 = 1 \le 2 = a_2$ we indeed have that $a_1 \le a_2$.
- Inductive Hypothesis: Suppose $a_k \leqslant a_{k+1}$ for some $k \in \mathbb{N}$.
- Inductive Step: Since $a_k \le a_{k+1}$, we also have $3 + a_k \le 3 + a_{k+1}$, which in turn implies $\frac{3+a_k}{2} \le \frac{3+a_{k+1}}{2}$. From the definition of the sequence, this means $a_{k+1} \le a_{k+2}$.

By Mathematical Induction, the sequence is increasing. $\square$

2. **Bounded above.** We need to show that the sequence is bounded above; but this does not mean that we need to find the least upper bound at this point. To show the sequence is bounded, any upper bound will do. Computing a couple of terms as in the proof of monotonicity above, we might guess that $a_n \le 5$ for all $n \in \mathbb{N}$. Let's prove it by Mathematical Induction.

   Claim: $a_n \le 5$ for all $n \in \mathbb{N}$.

   Proof:

   - Base Case: Since $a_1 = 1 \le 5$ we indeed have that $a_1 \le 5$.
   - Inductive Hypothesis: Suppose $a_k \le 5$ for some $k \in \mathbb{N}$.
   - Inductive Step: Since $a_k \le 5$, we have $3 + a_k \le 8$, and thus $a_{k+1} = \frac{3+a_k}{2} \le 4$. Thus, $a_{k+1} \le 4 \le 5$.

   By Mathematical Induction, we conclude that $\{a_n\}$ is indeed bounded above by 5. $\square$

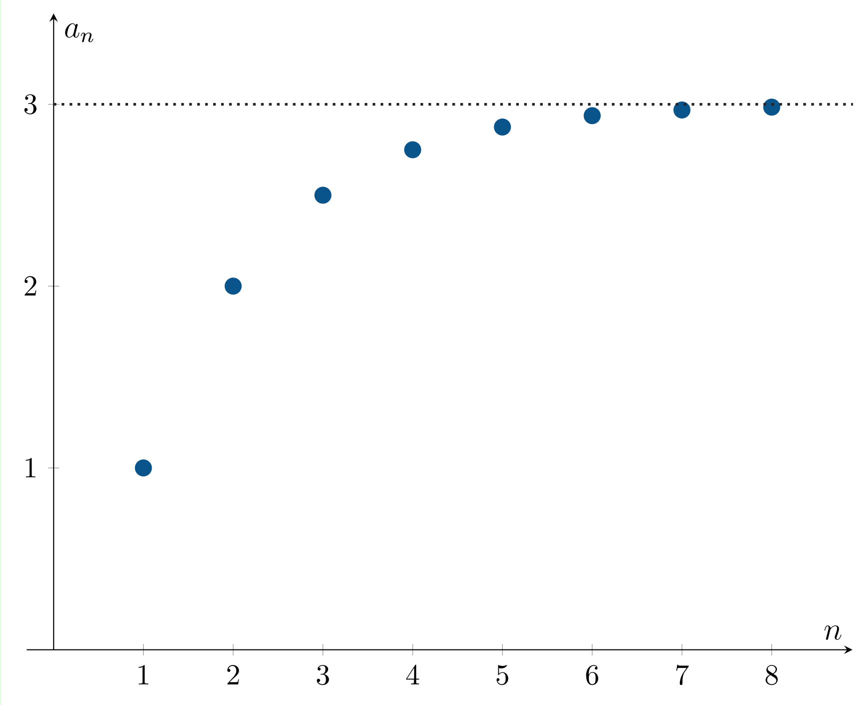
3. **Monotone Convergence Theorem.** Since $\{a_n\}$ is bounded above and increasing, the MCT implies that $\{a_n\}$ converges and has a limit L.

4. **Finding the limit.** From Step 3, we know the limit exists. Let $L = \lim_{n\to\infty} a_n$. But then also $L = \lim_{n\to\infty} a_{n+1}$, so

$$L = \lim_{n\to\infty} a_{n+1} = \lim_{n\to\infty} \frac{3+a_n}{2} = \frac{3+\lim_{n\to\infty} a_n}{2} = \frac{3+L}{2}$$

   Solving $L = \frac{3+L}{2}$ for L gives $2L = 3 + L$, or $L = 3$.

We conclude that $\lim_{n\to\infty} a_n = 3$.



Notice that both claims, $a_n \leq a_{n+1}$ and $a_n \leq 5$, were proven using the same steps. In fact, we can prove both increasing and bounded above in a single use of Mathematical Induction as seen in the following example.

Example 1.5.10 Consider the sequence $\{a_n\}$ recursively defined via

$$a_1 = 2, \quad a_{n+1} = \sqrt{7 + a_n} \quad (n \geqslant 1).$$

We prove that this sequence converges and find its limit. We first inspect a couple of terms:

$$a_1 = 2, \quad a_2 = \sqrt{9} = 3, \quad a_3 = \sqrt{10}, \quad a_4 = \sqrt{7 + \sqrt{10}}, \quad \ldots$$

so it seems that the sequence $\{a_n\}$ is increasing. We will prove that the sequence is increasing and bounded from above with one application of Mathematical Induction:

Claim: $a_n \leq a_{n+1} \leq 9$ for all $n \in \mathbb{N}$.

Proof:

- Base Case: Since $a_1 = 2 \leq 3 = a_2$ and $a_2 = 3 \leq 9$, we indeed have that $a_1 \leq a_2 \leq 9$.
- Inductive Hypothesis: Suppose $a_k \leqslant a_{k+1} \leq 9$ for some $k \in \mathbb{N}$.

- Inductive Step: Starting from the Inductive Hypothesis, we find

$$\begin{aligned}
& a_k \leqslant a_{k+1} \leq 9 \\
\Rightarrow\ & 7 + a_k \leqslant 7 + a_{k+1} \leq 7 + 9 = 16 \\
\Rightarrow\ & \sqrt{7 + a_k} \leqslant \sqrt{7 + a_{k+1}} \leq \sqrt{16} \\
\Rightarrow\ & a_{k+1} \leq a_{k+2} \leq 4 \leq 9
\end{aligned}$$

By Mathematical Induction, the sequence is increasing and bounded above.

□

Since the sequence in increasing and bounded above, the MCT implies that $\{a_n\}$ converges, say $L = \lim_{n\to\infty} a_n$. But then also $L = \lim_{n\to\infty} a_{n+1}$ holds, which yields the equation

$$L = \lim_{n\to\infty} a_{n+1} = \lim_{n\to\infty} \sqrt{7 + a_n} = \sqrt{7 + L}.$$

We can solve this equation as follows:

$$\begin{aligned}
& L = \sqrt{7 + L} \\
\Rightarrow\ & L^2 = 7 + L \\
\Rightarrow\ & L^2 - L - 7 = 0 \\
\Rightarrow\ & L = \frac{1 \pm \sqrt{29}}{2}
\end{aligned}$$

We obtain two solutions: $L_1 = \frac{1+\sqrt{29}}{2}$ and $L_2 = \frac{1-\sqrt{29}}{2}$. We know that the limit L satisfies $L = \text{lub}(\{a_n\})$. But $L_2 = \frac{1-\sqrt{29}}{2} < 2 = a_1$, hence, L_2 is not a upper bound (let alone the lowest upper bound). Hence, we must have that $L = L_1 = \frac{1+\sqrt{29}}{2}$.

Overall, we conclude that $\lim_{n\to\infty} a_n = \frac{1+\sqrt{29}}{2}$.

REMARK

The order of the steps does matter! We cannot perform the last step, unless we know that $\lim_{n\to\infty} a_n$ exists.

For instance, consider the sequence $\{a_n\}$ with

$$a_1 = 1, \quad a_{n+1} = 2a_n + 1.$$

The first few terms are

$$a_1 = 1, \quad a_2 = 3, \quad a_3 = 7, \quad a_4 = 15, \quad \dots$$

Had we carelessly started with the last step, and argued $\lim_{n\to\infty} a_n = \lim_{n\to\infty} a_{n+1} = L$, we would have obtain the ridiculous result $L = 2 \cdot L + 1$ or $L = -1$. But the given sequence $\{a_n\}$ does not converge (and in fact, is never negative). While the given sequence is monotonic, it is not bounded, and we cannot apply MCT.

Section 1.5 Problems

1.5.1. Define a sequence $\{a_n\}$ by $a_1 = 1$ and $a_{n+1} = \dfrac{7 + a_n}{6}$ for $n \geq 1$.

(a) Prove by induction that $\{a_n\}$ is increasing.

(b) Prove by induction that $\{a_n\}$ is bounded above by 2.

(c) Prove that $\{a_n\}$ is convergent and find $\lim\limits_{n\to\infty} a_n$.

1.5.2. Define a sequence $\{a_n\}$ by $a_1 = 1$ and $a_{n+1} = \dfrac{1}{3 - a_n}$ for $n \geq 1$.

(a) Prove that $\{a_n\}$ is monotonic.

(b) Prove that $\{a_n\}$ is bounded.

(c) Prove that $\{a_n\}$ is convergent and find $\lim\limits_{n\to\infty} a_n$.

1.5.3. Consider the sequence $\{a_n\}$ defined by $a_1 = 19$ and $a_{n+1} = \dfrac{\sqrt{9a_n - 2}}{3}$ for $n \in \mathbb{N}$. Prove that $\{a_n\}$ converges and find its limit.

Hint: As usual, your main tool will be the Monotone Convergence Theorem. However, in this example, it will be difficult to identify the limit of the sequence without knowing sharp bounds for a_n. It may therefore help to figure out what the possible limits are – something you would normally do at the *end* of such a problem – and then use this information to guess (and prove!) bounds for the sequence accordingly.

1.5.4. The Monotone Convergence Theorem states that if a sequence is both bounded and monotone, then the sequence must converge. In this exercise, we investigate the converse of this statement – which turns out to be false.

(a) Demonstrate by example that there exist convergent sequences that are not monotonic. Can you find an example of a convergent sequence such that no tail of the sequence is monotonic?

(b) However, prove that if $\{a_n\}$ converges to a limit L, then $\{a_n\}$ must be bounded. That is, prove that there exist numbers m and M such that $m \leq a_n \leq M$ for all n.

1.5.5. The Fibonacci sequence is given by $1, 1, 2, 3, 5, 8, 13, 21, \ldots$ where each subsequent term is given by the sum of the preceding two terms, $a_{n+2} = a_{n+1} + a_n$ $(n \geq 1)$, starting with $a_1 = a_2 = 1$. It is clear that $\lim\limits_{n\to\infty} a_n = \infty$. However, the limit of the ratio of consecutvie Fibonacci numbers, a_{n+1}/a_n, *does* exist!

In the following two problems, we will explore some methods to determine this limit.

(a) Show that the terms in the Fibonacci sequence are related as

$$\frac{a_{n+2}}{a_{n+1}} = 1 + \frac{a_n}{a_{n+1}}, \; n \geq 1.$$

(b) Assume that the sequence $b_n = a_{n+1}/a_n$ converges and find the value of the limit $\lim_{n\to\infty} b_n = \varphi$. Start by constructing a recursive relationship for b_{n+1}.

(c) (**Challenge Problem!**) Prove that the sequence $\{b_n = a_{n+1}/a_n\}$ indeed converges by proving that the sequence is a contraction.

(i) First, show that

$$b_{n+1} = 1 + \frac{b_{n-1}}{1+b_{n-1}} \quad (n \geq 2),$$

and explain why this implies that $1 < b_n < 2$ for $n \geq 3$.

(ii) Second, using the recursion formula for b_n in part (a), and the limit φ from part (b), show that

$$|b_{n+1} - \varphi| = \left|\frac{1}{b_n} - \frac{1}{\varphi}\right| < \frac{1}{\varphi}|b_n - \varphi|,$$

for $n \geq 3$.

(iii) Finally, show that

$$|b_{n+1} - \varphi| < \frac{1}{\varphi^{n-2}}|b_3 - \varphi|,$$

and explain why we can conclude that $b_n \to \varphi$ as $n \to \infty$.

1.5.6. (**Challenge Problem!**) As in the previous problem, the Fibonacci sequence is given by $1, 1, 2, 3, 5, 8, 13, 21, \ldots$ where each subsequent term is given by the sum of the preceding two terms,

$$a_{n+2} = a_{n+1} + a_n \ (n \geq 0), \tag{1.5}$$

starting with $a_0 = 0$ and $a_1 = 1$.

Recursively-defined sequences are notoriously difficult to analyze because we cannot simply use limit rules to determine convergence. The goal of this question is to convert the recursive sequence defined by Eq. 1.5 into an explicit sequence.

(a) Assume that the terms in the Fibonacci sequence are given explicitly by $a_n = c\lambda^n$ where c and λ are constants. Substitute this expression into Eq. (1.5), and solve for λ (you should find two values, λ_1 and λ_2).

(b) The full expression for a_n is written as the linear combination, $a_n = c_1\lambda_1^n + c_2\lambda_2^n$. Use the initial terms in the sequence, $a_0 = 0$ and $a_1 = 1$, to determine c_1 and c_2. Then, use a calculator to compute a few terms of a_n using this expression and compare with the Fibonacci sequence.

(c) Using the explicit expression for a_n from part (b), determine the limit of the ratio of subsequent Fibonnaci terms, $\lim_{n\to\infty} \frac{a_{n+1}}{a_n}$.

(d) Using induction, prove that for $n \geq 1$,

$$a_1^2 + a_2^2 + \cdots + a_n^2 = a_n a_{n+1}.$$

1.5.7. (**Challenge Problem!**) The implicitly-defined sequence, with growth parameter $1 < r < 4$, and initial state $0 < X_0 < 1$,

$$X_{n+1} = rX_n(1 - X_n), \quad n = 0, 1, 2, 3, \ldots$$

was popularized by the mathematical biologist Robert May. Although it is a simple-looking sequence, the limiting-behaviour has an unusual dependence on the growth parameter r.

For values of $1 < r < 4$, initialize the sequence at $X_0 = \frac{1}{2}$, generate 500 terms and plot the two-hundred-term 'tail' of the sequence $\{X_n\}_{n=300}^{500}$ as a function of the parameter r. Pay careful attention to the parameter ranges $1 < r < 3$, $3 \leq r < 3.57$ and $3.57 \leq r < 4$.

1.5.8. (**Challenge Problem!**) The sequence

$$a_{n+1} = \begin{cases} a_n/2, & a_n \text{ is even} \\ 3a_n + 1, & a_n \text{ is odd} \end{cases}$$

is called the hail stone sequence because, for many initial (positive integer) values a_0, the terms of the sequence make large excursions up and down, like hail stones in a storm cloud. Once the sequence reaches $a_k = 1$, it will oscillate forever as $\{4, 2, 1\}$. The index k at which the sequence first reaches $a_k = 1$ is called the total stopping time of the sequence.

(a) For initial points $a_0 = [1, 50]$, plot the total stopping time k as a function of a_0.

(b) You should find that the largest stopping time in part (a) is $k = 111$ when $a_0 = 27$. Plot the sequence $\{a_n\}$ as a function of the index n starting at $a_0 = 27$. How large does the sequence climb?

Collatz (1937) conjectured that every initial point a_0 has a finite total stopping time. As of 2020, Collatz's conjecture has been verified up to $a_0 = 2^{68}$, but its proof remains an infamous unsolved problem in mathematics. The great 20^{th} century mathematician, Paul Erdös (1913-1996), said that the search for a proof of the Collatz conjecture is 'hopeless, absolutely hopeless' and that 'mathematics may not be ripe for such problems.'

Chapter 2

Function Limits and Continuity

Our journey into calculus began with thinking about sequences and their behaviour as $n \to \infty$. In this chapter, we will consider real valued functions defined over intervals (or even the entire real line), and formalize what we mean by the limit of a function $f(x)$, as the argument x approaches a fixed values $a \in \mathbb{R}$ and as x grows towards $\pm\infty$. We will demonstrate many connections between sequence limits and function limits. Towards the end of the chapter, we will use the notion of a function limit to define what it means for a function to be continuous. Roughly speaking, continuous functions are essentially functions whose graph can be drawn without lifting the pen, and we will formalize this idea through the use of the limit. These functions play a key role in mathematics, and many physical processes can be modelled through continuous functions. Finally, we will develop a method that can help us not only guarantee the existence of a solution to an equation, but even iteratively approximate this solution.

By the end of this chapter, you will be able to

- define and compute the limit of a function;
- determine if a given function is continuous;
- approximate roots of continuous functions via the bisection method.

2.1 Introduction to Function Limits

We now turn our attention to functions and their limits. We will use our knowledge of sequence limits to develop a formal definition of the limit of a function and, from there, we will see how our knowledge of sequence limit properties extends to limits of functions.

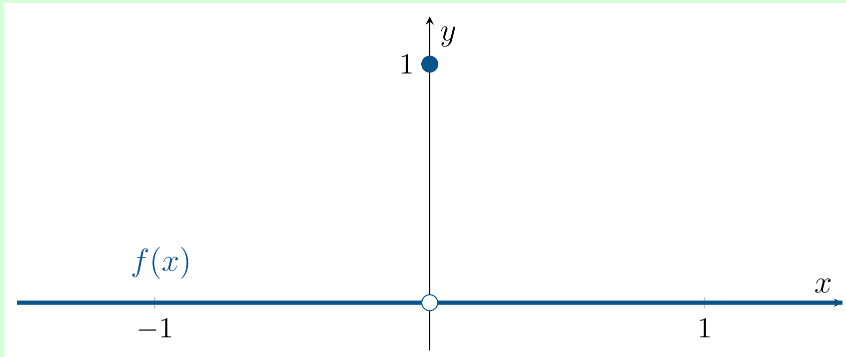
More precisely, for $a \in \mathbb{R}$, the limit of $f(x)$ as x approaches a is meant to describe the behavior of $f(x)$ (i.e., of the y values), as x gets arbitrarily close, but not necessarily equal to $x = a$. Before giving a formal definition, let us illustrate what we mean by the function limit in the following example.

Example 2.1.1

Consider the function

$$f(x) = \begin{cases} 1, & \text{if } x = 0, \\ 0, & \text{otherwise.} \end{cases}$$

The graph of the function f is shown below.



Let us investigate the behavior of the function near $x = 0$. Note how $f(0) = 1$ but $f(x) = 0$ for all $x \neq 0$. Inspecting function values near $x = 0$ we see

$$f(1) = 0, \quad f(0.1) = 0, \quad f(0.01) = 0, \quad \ldots, \quad f(0.0000001) = 0$$

and

$$f(-1) = 0, \quad f(-0.1) = 0, \quad f(-0.01) = 0, \quad \ldots, \quad f(-0.0000001) = 0.$$

Thus, as x gets arbitrarily close to 0 from the left or the right (while still satisfying $x \neq 0$), the function values $f(x)$ approach 0, which we will soon define to be the limit of $f(x)$ at x approaches 0.

Motivated by this example, we will now formalize what we mean by the limit of a function.

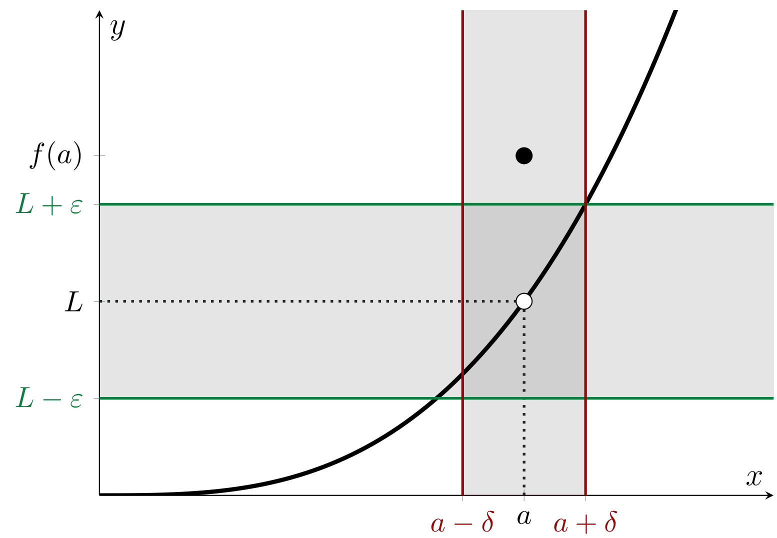
Suppose we are given a function $f : \mathbb{R} \to \mathbb{R}$ and wish to examine the statement $\lim\limits_{x\to a} f(x) = L$ where $a, L \in \mathbb{R}$. Intuitively, the statement means that the function value $f(x)$ gets infinitely close to the limit L as x gets infinitely close to a (while still $x \neq a$). Let us translate this to a more formal definition

Definition 2.1.2
Function Limit

Let $f : \mathbb{R} \to \mathbb{R}$ be a function and $a \in \mathbb{R}$. We say $\lim\limits_{x\to a} f(x) = L$ if for all $\varepsilon > 0$ there exists $\delta > 0$ such that if $0 < |x - a| < \delta$, then $|f(x) - L| < \varepsilon$.

Note that this definition requires that $0 < |x - a|$, which means in particular that $x \neq a$. The function limit is thus unaffected by the function value at $x = a$. Furthermore, for the limit to exist, the function needs to approach L regardless of whether x approaches a from the left or the right. (Note that we will also refer to this behaviour as "approaches from below or from above.")

The definition of a function limit is illustrated in the figure below. Note how $|f(x) - L| < \varepsilon$ for all x satisfying $0 < |x - a| < \delta$ (but not necessarily for $x = a$).



Next, let us consider how to actually use this formal definition to prove that the limit of a function exists.

Example 2.1.3 Show that

$$\lim_{x\to 2} 5x+1 = 11.$$

Solution:

Proof: Let $\varepsilon > 0$ be given. Let $\delta = \frac{\varepsilon}{5}$. Then, if $0 < |x-2| < \delta$, we get

$$\begin{aligned} |(5x+1)-11| &= |5x-10| = 5|x-2| \\ &< 5\cdot\delta = 5\cdot\frac{\varepsilon}{5} \\ &= \varepsilon, \end{aligned}$$

as desired. □

As an aside: How did we find that the choice $\delta = \frac{\varepsilon}{5}$ works? We wanted that

$$5|x-2| < \varepsilon.$$

But this holds if and only if

$$|x-2| < \varepsilon/5 = \delta.$$

That is, similar to sequences, we control the value of δ and choose it in a manner that results in $|f(x)-L| < \varepsilon$.

Example 2.1.4 Show that

$$\lim_{x \to 5} x^2 = 25.$$

Solution:

Proof: Let $\varepsilon > 0$ be given. Let $\delta = \min\{\frac{\varepsilon}{11}, 1\}$ (we will comment on the use of the minimum function below). We first note that

$$\left|x^2 - 25\right| = |x-5||x+5|.$$

Now, suppose that $0 < |x-5| < \delta$, then

$$\left|x^2 - 25\right| < \delta \cdot |x+5|.$$

Note how we control the size of $|x-5|$ via δ, but not the size of $|x+5|$. The choice $\delta = \frac{\varepsilon}{|x+5|}$, which would lead to $\left|x^2 - 25\right| < \varepsilon$, is invalid, as it depends on x.

Hence, we additionally assume that $\delta \leq 1$, as this will help us find an upper bound on $|x+5|$ as follows: If $|x-5| < \delta$, and $\delta \leq 1$, then we must also have $|x-5| < 1$, which implies $4 < x < 6$. If $4 < x < 6$, then

$$|x+5| \leq |x| + |5| < 6 + 5 = 11.$$

Overall, if $|x-5| < \min\{\frac{\varepsilon}{11}, 1\}$, which will ensure that $\delta \leq 1$ and $\delta \leq \frac{\varepsilon}{11}$, then

$$\begin{aligned}
\left|x^2 - 25\right| &= |x-5||x+5| \\
&< \delta \cdot 11 \\
&\leq \frac{\varepsilon}{11} \cdot 11 \\
&= \varepsilon
\end{aligned}$$

as desired. $\square$

Example 2.1.5 As in Example 2.1.1, consider the function

$$f(x) = \begin{cases} 1, & \text{if } x = 0, \\ 0, & \text{otherwise.} \end{cases}$$

Show that $\lim_{x \to 0} f(x) = 0$, which is not equal to $f(0) = 1$.

Solution:

Let $\varepsilon > 0$. Let $\delta = \varepsilon$, though we could in fact choose any $\delta > 0$. Then, if $0 < |x-0| < \delta$, we have in particular $x \neq 0$, so

$$|f(x) - 0| = |0 - 0| = 0 < \varepsilon,$$

as desired.

Note that the function $f(x)$ given above and the constant function $g(x) = 0$ are not equal everywhere: While $\lim_{x\to 0} f(x) = 0 = \lim_{x\to 0} g(x)$ holds true, we have $g(0) \neq f(0)$. The graph of f is displayed in Example 2.1.1: The filled circle at $(0, 1)$ indicates that the function value at $x = 0$ is $f(0) = 1$.

Similar to limits of sequences, it is tricky to work with the formal Definition 2.1.2. In the next section, We will develop some better techniques for determining the limit of a function! We close this section with a few remarks.

REMARK

1. For $\lim_{x\to a} f(x)$ to exist, the function f must be defined in an open interval, (α, β), containing a (except possibly at $x = a$).
2. The function value $f(a)$ at $x = a$ does not affect the limit $\lim_{x\to a} f(x)$.
3. If $f(x) = g(x)$ for all x, except possibly at $x = a$, then $\lim_{x\to a} f(x) = \lim_{x\to a} g(x)$.

Section 2.1 Problems

2.1.1. Use the $\varepsilon - \delta$ definition of limits to show that $\lim_{x\to 3}(4x + 1) = 13$.

2.1.2. Use the $\varepsilon - \delta$ definition of limits to show that $\lim_{x\to -1}(1 - 9x) = 10$.

2.1.3. Use the $\varepsilon - \delta$ definition of limits to show that $\lim_{x\to 2}\left(x^2 - 4x + 4\right) = 0$.

2.1.4. Use the $\varepsilon - \delta$ definition of limits to show that $\lim_{x\to 4}\sqrt{x} = 2$.

2.1.5. Use the $\varepsilon - \delta$ definition of limits to show that $\lim_{x\to 1} x^3 = 1$.

Hint: $a^3 - b^3 = (a - b)(a^2 + ab + b^2)$.

2.1.6. Use the $\varepsilon - \delta$ definition of limits to show that $\lim_{x\to 3}\dfrac{1}{x^2} = \dfrac{1}{9}$.

2.1.7. Let $f(x)$ and $g(x)$ be functions such that $\lim_{x\to a} f(x) = L$ and $\lim_{x\to a} g(x) = M$. Use the $\varepsilon - \delta$ definition of limits to show that $\lim_{x\to a}(f(x) + g(x)) = L + M$.

Note: We saw the same phenomenon when working with limits of sequences, and the proof there followed a similar structure. In the next section, you'll learn about the Sequential Characterization of Function Limits, which lets us transfer many familiar results from sequences to functions—without having to reprove everything from scratch using the ε–δ definition!

2.2 Sequential Characterization of Limits

In this section we will establish, and then use, a connection between limits of sequences and limits of functions. This connection will allow us to immediately deduce that function limits are unique and prove commonly used properties of function limits.

Theorem 2.2.1 **(Sequential Characterization of Function Limits)**

Let $a \in \mathbb{R}$. Let the function $f(x)$ be defined on an open interval containing $x = a$, except possibly $x = a$. Then, the following are equivalent

1. $\lim_{x \to a} f(x) = L$
2. For all sequences $\{x_n\}$ that satisfy $\lim_{n \to \infty} x_n = a$ and $x_n \neq a$ for all $n \in \mathbb{N}$, we have that $\lim_{n \to \infty} f(x_n) = L$.

Proof: We show (1) $\Rightarrow$ (2). Suppose $\lim_{x \to a} f(x) = L$.

Let $\varepsilon > 0$ be given. By definition of the function limit, there exists $\delta > 0$ so that if $0 < |x - a| < \delta$, we have $|f(x) - L| < \varepsilon$.

Next, let $\{x_n\}$ be any sequence of real numbers such that $\lim_{n \to \infty} x_n = a$ and $x_n \neq a$. By definition of sequence limits, there exists $N \in \mathbb{R}$ such that if $n > N$ we have $|x_n - a| < \delta$. Therefore, if $n > N$, we have

$$|f(x_n) - L| < \varepsilon,$$

which implies $\lim_{n \to \infty} f(x_n) = L$, as desired.

We leave the other direction (2) $\Rightarrow$ (1) as a tricky exercise to think about. □

This theorem is very powerful as it means that a function limit can be interpreted as the limit of a sequence and many of the results of sequence limits *immediately* apply to function limits!

For example, recall that the limit of a sequence (if it exists) is unique, see Theorem 1.2.18.

Theorem 2.2.1 therefore immediately implies that the limit of functions is unique as well, which we summarize in the following theorem.

Theorem 2.2.2 **(Uniqueness of Function Limits)**

The limit of a function is unique. That is, if $\lim_{x \to a} f(x) = L$ and $\lim_{x \to a} f(x) = M$, then we must have $L = M$.

The sequential characterization of function limits can help us prove that a function limit does not exist. Two common strategies to prove that a function limit $\lim_{x \to a} f(x)$ does not exist are as follows:

1. Find a sequence $\{x_n\}$ with $x_n \to a$ and $x_n \neq a$, for which $\lim_{n \to \infty} f(x_n)$ does not exist.

2. Find two sequences $\{x_n\}$ and $\{y_n\}$ with $x_n \to a, x_n \neq a, y_n \to a, y_n \neq a$, for which $\lim_{n\to\infty} f(x_n) \neq \lim_{n\to\infty} f(y_n)$. That is, find two sequences whose limits $f(x_n)$ and $f(y_n)$ differ.

Example 2.2.3 Let

$$f(x) = \frac{|x|}{x}.$$

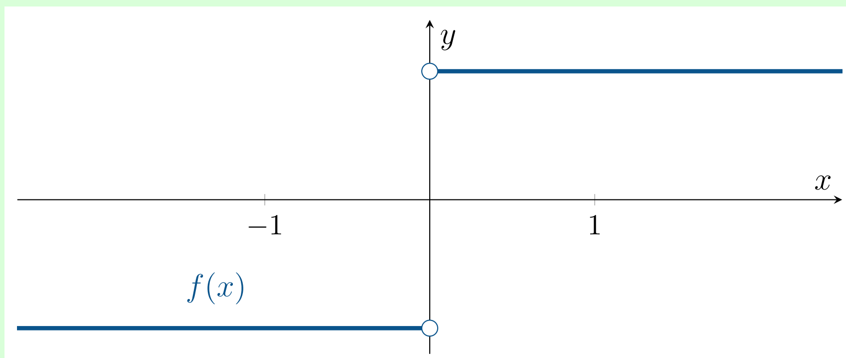
Show that $\lim_{x\to 0} f(x)$ does not exist.

Solution:

Note that

$$f(x) = \frac{|x|}{x} = \begin{cases} 1, & \text{if } x > 0, \\ -1, & \text{if } x < 0. \end{cases}.$$

The graph of the function f is shown below.



The fact that $f(x) = 1$ for all $x > 0$ and $f(x) = -1$ for all $x < 0$ motivates us to choose a positive sequence $\{x_n\}$ and a negative sequence $\{y_n\}$, both of which converge to 0 as $n \to \infty$.

Let

$$x_n = \frac{1}{n}, \quad y_n = -\frac{1}{n}, \quad n \in \mathbb{N}.$$

We show that applying the function f will lead to different limits of $f(x_n)$ and $f(y_n)$.

Clearly, $\{x_n\}$ and $\{y_n\}$ satisfy the requirements that $x_n \neq 0$, $y_n \neq 0$ and $\lim_{n\to\infty} x_n = \lim_{n\to\infty} y_n = 0$. But

$$\lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} f\left(\frac{1}{n}\right) = \lim_{n\to\infty} 1 = 1$$

while

$$\lim_{n\to\infty} f(y_n) = \lim_{n\to\infty} f\left(-\frac{1}{n}\right) = \lim_{n\to\infty} -1 = -1.$$

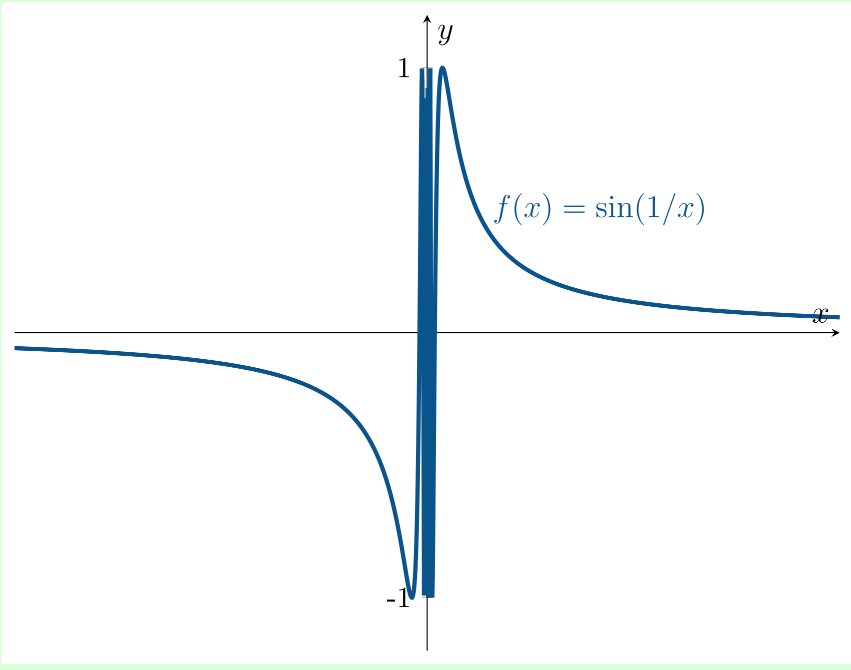
Since $\lim_{n\to\infty} f(x_n) \neq \lim_{n\to\infty} f(y_n)$, this means that $\lim_{x\to 0} f(x)$ cannot exist.

Example 2.2.4 Let
$$f(x) = \sin\left(\frac{1}{x}\right).$$
Show that $\lim_{x\to 0} f(x)$ does not exist.

Solution:

As we can see from the plot below, the function is oscillating increasingly rapidly as we approach zero. This suggests that the limit does not exist. To show this, let
$$x_n = \frac{1}{\pi/2 + 2n\pi}, \quad y_n = \frac{1}{3\pi/2 + 2n\pi}, \quad n \in \mathbb{N}.$$
Then $x_n \neq 0$, $y_n \neq 0$ and $\lim_{n\to\infty} x_n = \lim_{n\to\infty} y_n = 0$. Applying the function f gives
$$\lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} \sin(\pi/2 + 2n\pi) = \lim_{n\to\infty} 1 = 1$$
while
$$\lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} \sin(3\pi/2 + 2n\pi) = \lim_{n\to\infty} -1 = -1,$$
so again, the limit $\lim_{x\to 0} f(x)$ cannot exist.

Alternatively, we can define a single sequence
$$x_n = \frac{1}{\frac{\pi}{2}n}, \quad n \in \mathbb{N}.$$
Clearly, $x_n \to 0$ and $x_n \neq 0$. We calculate $\sin(\frac{1}{x_n}) = \sin(\frac{\pi}{2}n)$ and this corresponds to the sequence $\{1, 0, -1, 0, 1, 0, \ldots\}$. Thus, $\lim_{n\to\infty} \sin\left(\frac{\pi}{2}n\right)$ does not exist and, therefore, $\lim_{x\to 0} \sin\left(\frac{1}{x}\right)$ does not exist.



2.2.1 Arithmetic Rules for Function Limits

Similar to the arithmetic rules for sequence limits, we can use Theorem 2.2.1 to develop a set of arithmetic rules for function limits!

Theorem 2.2.5 **(Arithmetic Rules for Function Limits)**

Let f and g be functions and $a \in \mathbb{R}$. Assume $\lim_{x \to a} f(x) = L$, $\lim_{x \to a} g(x) = M$, where $L, M \in \mathbb{R}$.

Let $c \in \mathbb{R}$. Then,

1. If $f(x) = c$ for all $x \in \mathbb{R}$, then $L = c$.
2. $\lim_{x \to a} cf(x) = cL$.
3. $\lim_{x \to a} f(x) \pm g(x) = L \pm M$.
4. $\lim_{x \to a} f(x)g(x) = LM$.
5. $\lim_{x \to a} \dfrac{f(x)}{g(x)} = \dfrac{L}{M}$, provided $M \neq 0$.
6. $\lim_{x \to a} (f(x))^{\alpha} = L^{\alpha}$ for any $\alpha > 0, L > 0$.
7. If $M = 0$ and $\lim_{x \to a} \dfrac{f(x)}{g(x)}$ exists, then $L = 0$.

We can use the above limit rules to quickly find limits of certain kinds of functions.

Theorem 2.2.6 **(Limits of Polynomials)**

If $p(x) = a_0 + a_1 x + \cdots + a_n x^n$ is a polynomial, then

$$\lim_{x \to a} p(x) = p(a).$$

Proof: We only sketch the proof and leave the details as an exercise for the reader. First, define $f_j(x) = a_j x^j$ for $j = 0, 1, \ldots, n$ (noting that $x^0 = 1$ for $x \in \mathbb{R}$). You can show that $\lim_{x \to a} f_j(x) = f_j(a)$ using Properties 2. and 4. from Theorem 2.2.5. Next, since $p(x) = f_0(x) + f_1(x) + \cdots + f_n(x)$, you can use Property 3 from Theorem 2.2.5 to show that $\lim_{x \to a} p(x) = p(a)$. $\square$

We close this section by considering rational functions, which we recall are fractions of polynomials.

Theorem 2.2.7 **(Limit of a Rational Function)**

Let $f(x) = \dfrac{p(x)}{q(x)}$, where p, q are polynomials. Let $a \in \mathbb{R}$. Then,

1. If $q(a) \neq 0$, then $\lim\limits_{x \to a} \dfrac{p(x)}{q(x)} = \dfrac{p(a)}{q(a)}$.

2. If $\lim\limits_{x \to a} q(x) = 0$ and $\lim\limits_{x \to a} p(x) \neq 0$, then $\lim\limits_{x \to a} \dfrac{p(x)}{q(x)}$ does not exist.

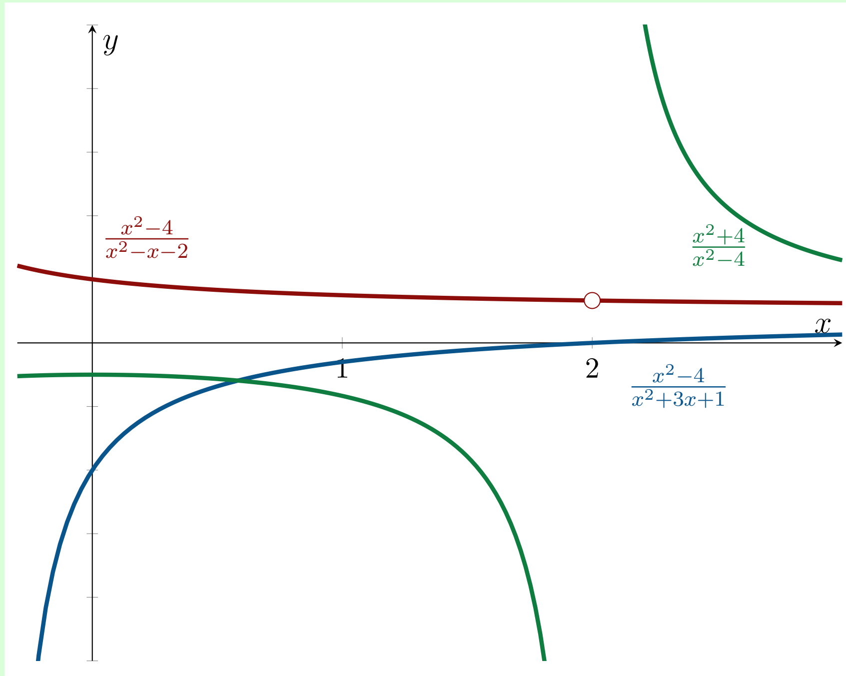
3. If $\lim\limits_{x \to a} q(x) = q(a) = 0 = p(a) = \lim\limits_{x \to a} p(x)$, then $(x - a)$ is a common factor of the two polynomials p and q. In other words, we can write $p(x) = (x - a)p^*(x)$ and $q(x) = (x - a)q^*(x)$, and we get
$$\lim_{x \to a} \frac{p(x)}{q(x)} = \lim_{x \to a} \frac{(x - a)p^*(x)}{(x - a)q^*(x)} = \lim_{x \to a} \frac{p^*(x)}{q^*(x)}$$
and we can go back to 1), 2), or 3) with the new rational function $\dfrac{p^*(x)}{q^*(x)}$.

Example 2.2.8 We illustrate each case from the previous theorem in an example:

1. Let $f(x) = \dfrac{x^2 - 4}{x^2 + 3x + 1}$ and $a = 2$. Then
$$\lim_{x \to 2} \frac{x^2 - 4}{x^2 + 3x + 1} = \frac{0}{11} = 0.$$

2. Let $f(x) = \dfrac{x^2 + 4}{x^2 - 4}$ and $a = 2$. Then the denominator converges to 0 as $x \to 2$, while the numerator converges to 8. Hence, $\lim\limits_{x \to \infty} \dfrac{x^2 + 4}{x^2 - 4}$ does not exist.

3. Let $f(x) = \dfrac{x^2 - 4}{x^2 - x - 2}$ and let $a = 2$. Note that $x = 2$ is a common factor of both $x^2 - 4$ and $x^2 - x - 2$. Hence,
$$\lim_{x \to 2} \frac{x^2 - 4}{x^2 - x - 2} = \lim_{x \to 2} \frac{(x - 2)(x + 2)}{(x - 2)(x + 1)} = \lim_{x \to 2} \frac{x + 2}{x + 1} = \frac{4}{3}.$$



Example 2.2.9 Let $f(x) = x^2$. Compute the following limit:

$$\lim_{h\to 0} \frac{f(2+h)-f(2)}{h}.$$

Solution:

We find

$$\begin{aligned}\lim_{h\to 0} \frac{f(2+h)-f(2)}{h} &= \lim_{h\to 0} \frac{(2+h)^2-2^2}{h}\\ &= \lim_{h\to 0} \frac{4h-h^2}{h}\\ &= \lim_{h\to 0} 4-h\\ &= 4.\end{aligned}$$

We will return to this important limit in Definition 3.2.2.

2.2.2 The Squeeze Theorem for Functions

We end this section by remarking that the Squeeze Theorem for sequences can be extended to a Squeeze Theorem for functions, by using Theorem 2.2.1.

Theorem 2.2.10 **(Squeeze Theorem for Functions)**

Let $f(x), g(x), h(x)$ be three functions defined in an open interval I around a, except possibly at a itself. If for all $x \in I$ (except possibly $x = a$) we have

$$g(x) \leq f(x) \leq h(x)$$

and

$$\lim_{x \to a} g(x) = L = \lim_{x \to a} h(x)$$

then $\lim_{x \to a} f(x) = L$ as well.

Example 2.2.11 Show that

$$\lim_{x \to 0} x \sin\left(\frac{1}{x}\right) = 0.$$

Solution:

We first observe that, if $x > 0$,

$$-x \leq x \sin\left(\frac{1}{x}\right) \leq x,$$

because $-1 \leq \sin\left(\frac{1}{x}\right) \leq 1$.

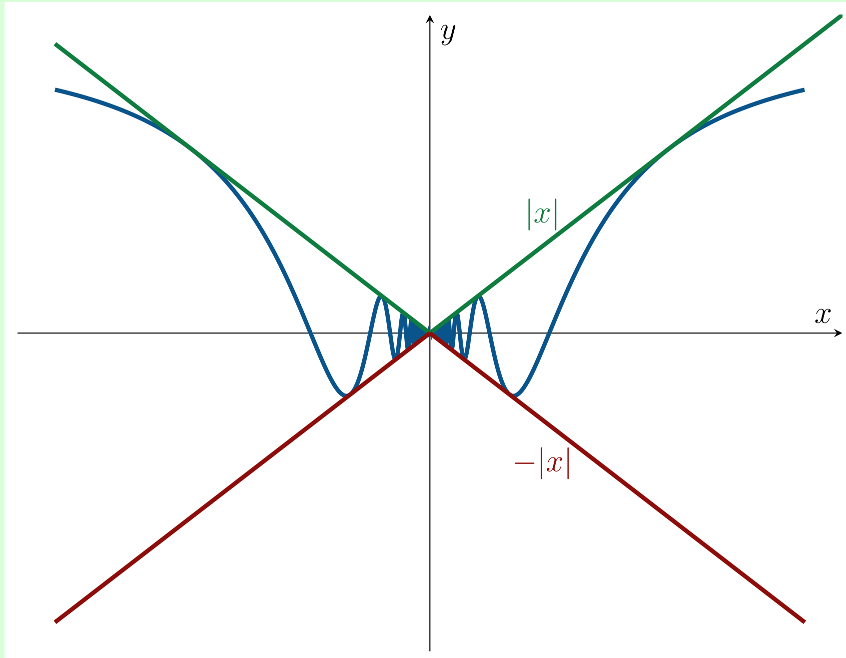
This inequality does not hold if $x < 0$ (to see this, plug in $x = -1$). Indeed, if $x < 0$, then

$$x \leq x \sin\left(\frac{1}{x}\right) \leq -x.$$

Since $|x| = x$ for $x > 0$ and $|x| = -x$ for $x < 0$, we can combine the two cases and obtain

$$-|x| \leq x \sin\left(\frac{1}{x}\right) \leq |x|$$

for all $x \in \mathbb{R}$ with $x \neq 0$. But $\lim_{x \to 0} -|x| = \lim_{x \to 0} |x| = 0$. Hence also $\lim_{x \to 0} x \sin\left(\frac{1}{x}\right) = 0$ by the squeeze theorem. The graph of the function f, along with its lower and upper bounds $-|x|$ and $|x|$, is displayed below.



Section 2.2 Problems

2.2.1. Use the Sequential Characterization of Function Limits to evaluate

$$\lim_{x\to 4} \frac{\sqrt{x}}{1+x}.$$

2.2.2. Consider the function

$$f(x) = \cos^2\left(\frac{1}{x}\right).$$

Find sequences $\{x_n\}$ and $\{y_n\}$ that both converge to 0, $x_n \neq 0$ and $y_n \neq 0$ for all $n \in \mathbb{N}$, and $\lim\limits_{n\to\infty} f(x_n) \neq \lim\limits_{n\to\infty} f(y_n)$. Clearly explain why this implies that $\lim\limits_{x\to 0} f(x)$ does not exist.

2.2.3. Consider the function

$$f(x) = \frac{x^2-9}{|x-3|}.$$

Find sequences $\{x_n\}$ and $\{y_n\}$ that both converge to 3, $x_n \neq 3$ and $y_n \neq 3$ for all $n \in \mathbb{N}$, and $\lim\limits_{n\to\infty} f(x_n) \neq \lim\limits_{n\to\infty} f(y_n)$. Clearly explain why this implies that $\lim\limits_{x\to 3} f(x)$ does not exist.

2.2.4. For $x \in \mathbb{R}$, define the *floor* of x, denoted $\lfloor x \rfloor$, to be the largest integer less than or equal to x. For instance,

$$\lfloor 2 \rfloor = 2, \quad \lfloor 1.003 \rfloor = 1, \quad \lfloor \pi \rfloor = 3, \quad \lfloor -7.4 \rfloor = -8.$$

Consider the function

$$f(x) = \begin{cases} x & \text{if } \lfloor x \rfloor \text{ is even,} \\ -x & \text{if } \lfloor x \rfloor \text{ is odd.} \end{cases}$$

(a) Sketch the graph of $f(x)$.

(b) Prove that $\lim_{x \to 2} f(x)$ does not exist.

2.2.5. Prove the Squeeze Theorem for Functions using the Sequential Characterization of Function Limits and the Squeeze Theorem for Sequences.

2.3 One-Sided Limits

Recall the function $f(x) = \frac{|x|}{x}$ from Example 2.2.3. We saw that $f(x) = 1$ for $x > 0$ and $f(x) = -1$ for $x < 0$, with an undefined limit as x approaches 0. For such a function, it is reasonable to examine the limit as $x \to 0$ from the right, and as $x \to 0$ from the left.

More generally, we may want to examine the behaviour of a function at a point but only from one side, instead of both sides at the same time. Let's see how to do that, and what the behaviour means for the overall limit.

Definition 2.3.1

Right-side and Left-side Limits

Let f be a function and let $a, L \in \mathbb{R}$.

1. We say that L is the **right-side limit** of f at a, $\lim_{x\to a^+} f(x)$, if for any $\varepsilon > 0$ there exists $\delta > 0$ such that if $|x - a| < \delta$ and $x > a$ then $|f(x) - L| < \varepsilon$.
2. We say that L is the **left-side limit** of f at a, $\lim_{x\to a^-} f(x)$, if for any $\varepsilon > 0$ there exists $\delta > 0$ such that if $|x - a| < \delta$ and $x < a$ then $|f(x) - L| < \varepsilon$.

Example 2.3.2

Let

$$f(x) = \frac{|x|}{x} = \begin{cases} 1, & \text{if } x > 0, \\ -1, & \text{if } x < 0. \end{cases}$$

In Example 2.2.3 we have shown that $\lim_{x\to 0} f(x)$ does not exist. Show that $\lim_{x\to 0^+} f(x) = 1$ and $\lim_{x\to 0^-} f(x) = -1$.

Solution:

We only show $\lim_{x\to 0^+} f(x) = 1$; the other case is shown analogously. Let $\varepsilon > 0$ and $\delta = \varepsilon$. Then, if $|x| < \delta$ and $x > 0$, this means $0 < x < \delta$, and we have

$$|f(x) - 1| = |1 - 1| = 0 < \varepsilon.$$

As before, using the formal definition is quite cumbersome and we typically can use our limit properties to quickly deduce one-sided limits. Indeed, all of the arithmetic rules, the sequential characterization, and the Squeeze Theorem hold for one-sided limits as well. For example, if $\lim_{x\to a^+} f(x) = L$ and $\lim_{x\to a^+} g(x) = M$ then $\lim_{x\to a^+} f(x) + g(x) = L + M$.

Example 2.3.3

Let

$$f(x) = \begin{cases} -1, & \text{if } x \le 0 \\ 0, & \text{if } 0 < x \le 1, \\ x^2 - 1, & \text{if } x > 1. \end{cases}$$

The graph of the function is shown below. We find

$$\lim_{x\to 1^+} f(x) = \lim_{x\to 1^+} (x^2 - 1) = 1 - 1 = 0$$

and

$$\lim_{x\to 1^-} f(x) = \lim_{x\to 1^-} 0 = 0,$$

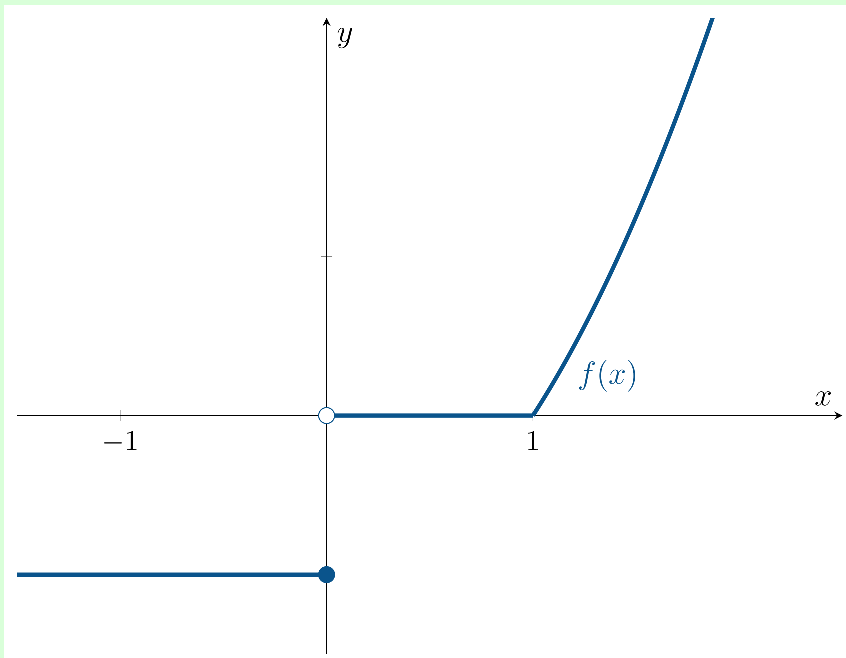
and both of these one-sided limits are equal to $f(1) = 0$.

Furthermore,

$$\lim_{x\to 0^+} f(x) = \lim_{x\to 0^+} 0 = 0$$

while

$$\lim_{x\to 0^-} f(x) = \lim_{x\to 0^-} -1 = -1 = f(0).$$



From the knowledge of one-sided limits, what can we infer about the overall limit? The proof of the following is an easy exercise:

Theorem 2.3.4 **(Function Limit Characterization through One-Sided Limits)**

Let $f(x)$ be defined in an open interval around a, except possibly a itself. The following are equivalent:

1. $\lim_{x\to a} f(x) = L$
2. Both $\lim_{x\to a^+} f(x) = L$ and $\lim_{x\to a^-} f(x) = L$.

As mentioned above, the Squeeze Theorem also applies to one-sided limits. We use this in the following example, which is similar to Example 2.2.11.

Example 2.3.5 Compute

$$\lim_{x \to 0^+} \sqrt{x} \sin\left(\frac{1}{x}\right).$$

Solution:

Let $x > 0$. Since $-1 \leq \sin\left(\frac{1}{x}\right) \leq 1$ and $\sqrt{x} > 0$ for all $x > 0$ we find

$$-\sqrt{x} \leq \sqrt{x} \sin\left(\frac{1}{x}\right) \leq \sqrt{x}.$$

Next, we show that $\lim_{x \to 0^+} \sqrt{x} = 0$. Let $\varepsilon > 0$ and $\delta = \varepsilon^2$. Then, for any $0 < x < \delta$,

$$|\sqrt{x} - 0| = \sqrt{x} < \sqrt{\delta} = \sqrt{\varepsilon^2} = \varepsilon.$$

Because of the limit rules, it also holds that $\lim_{x \to 0^+} -\sqrt{x} = -0 = 0$, so that overall the Squeeze Theorem implies that

$$\lim_{x \to 0^+} \sqrt{x} \sin\left(\frac{1}{x}\right) = 0.$$

Example 2.3.6 Show that

(1) $\lim_{x \to 0} \sin(x) = 0$

(2) $\lim_{x \to 0} \cos(x) = 1$

(3) $\lim_{x \to 0} \tan(x) = 0$

Solution: You may have an intuitive sense that the limits are "obvious." Here, we present a geometric argument to show the limit in (1) and use some of the limit concepts developed thus far to show the limits in (2) and (3).

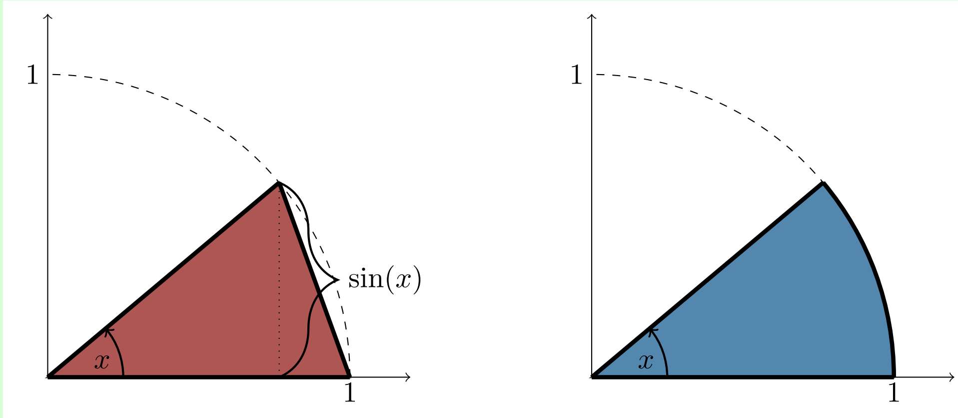
First, we show that $\lim_{x \to 0} \sin(x) = 0$ using a geometric argument with one-sided limits. The proof below shows that $\lim_{x \to 0^+} \sin(x) = 0$.

For $x \in (0, \frac{\pi}{2})$, we construct a unit circle and consider two areas as shown below. First, the red triangle has a height of $\sin(x)$ and a base of 1, resulting in an area of $\frac{1}{2}\sin(x)$. Second, the blue area of the sector of circle is given by $\frac{1}{2}x$.

Based on the figure below, we have that, for $x \in (0, \frac{\pi}{2})$,

$$0 \leq \frac{1}{2}\sin(x) \leq \frac{1}{2}x \Rightarrow 0 \leq \sin(x) \leq x.$$

Finally, since $\lim_{x \to 0^+} x = 0$, then by the Squeeze Theorem, $\lim_{x \to 0^+} \sin(x) = 0$. A similar argument can be used to show that $\lim_{x \to 0^-} \sin(x) = 0$. Thus, we have shown that $\lim_{x \to 0} \sin(x) = 0$.



Next, we show that $\lim_{x\to 0}\cos(x) = 1$. Luckily (or unluckily), we can do this algebraically! Since $\sin^2(x)+\cos^2(x) = 1$, then $\cos(x) = \sqrt{1-\sin^2(x)}$. Taking the limit and applying our arithmetic rules for limits of functions, we have that

$$\lim_{x\to 0}\cos(x) = \lim_{x\to 0}\sqrt{1-\sin^2(x)} = \sqrt{1-\lim_{x\to 0}\sin^2(x)} = \sqrt{1-0} = 1.$$

Lastly, we again apply our arithmetic rules to find that

$$\lim_{x\to 0}\tan(x) = \lim_{x\to 0}\frac{\sin(x)}{\cos(x)} = \frac{0}{1} = 0.$$

Section 2.3 Problems

2.3.1. Evaluate $\displaystyle\lim_{x\to 2^-}\frac{x-2}{\sqrt{x^2-4x+4}}$.

2.3.2. Evaluate $\displaystyle\lim_{x\to 0^+}\sqrt{x}\cos\left(\frac{\pi}{x}\right)$.

2.3.3. Evaluate $\displaystyle\lim_{x\to 4}\frac{x^2-8x+16}{|4-x|}$.

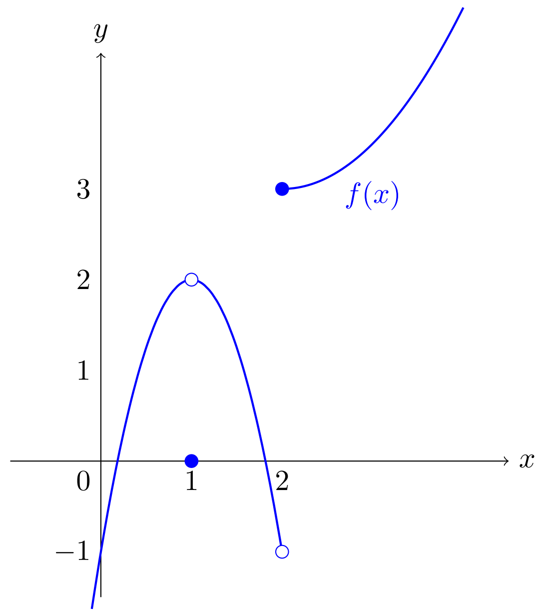
2.3.4. Evaluate $\displaystyle\lim_{x\to 1}\frac{|x-1|+|x+1|-2}{x-1}$.

2.3.5. Let a be a real number and consider the function defined by

$$f(x) = \begin{cases} \dfrac{x^2-1}{x^2+(a-1)x+a} & \text{if } 0 < x < 1, \\[2ex] \dfrac{\sqrt{x}-1}{x-1} & \text{if } x > 1. \end{cases}$$

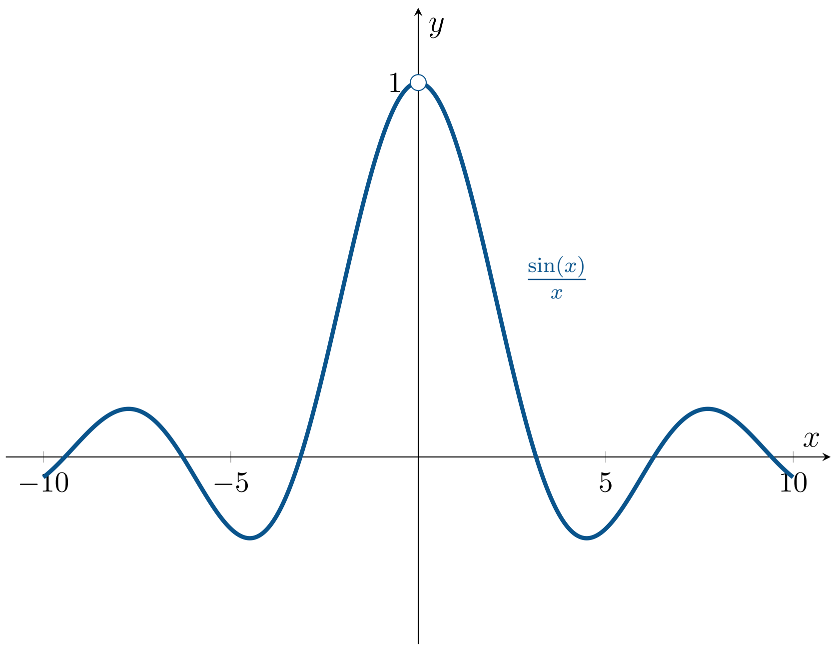
Determine the value of a such that $\lim_{x\to 1} f(x)$ exists.

2.3.6. Consider the graph of a function $y = f(x)$ shown below. Determine $\lim_{x\to 1} f(f(x))$. Explain your reasoning.



2.4 Fundamental Trigonometric Limit

We next consider the limit $\lim_{x \to 0} \dfrac{\sin(x)}{x}$. This is an important limit that will be useful when we talk about differentiation in chapter 3. A plot of the function $\frac{\sin(x)}{x}$ is shown below and it is clear from the graph that the limit is 1, but how do we go about showing this mathematically? Similar to example 2.3.6, we proceed geometrically!

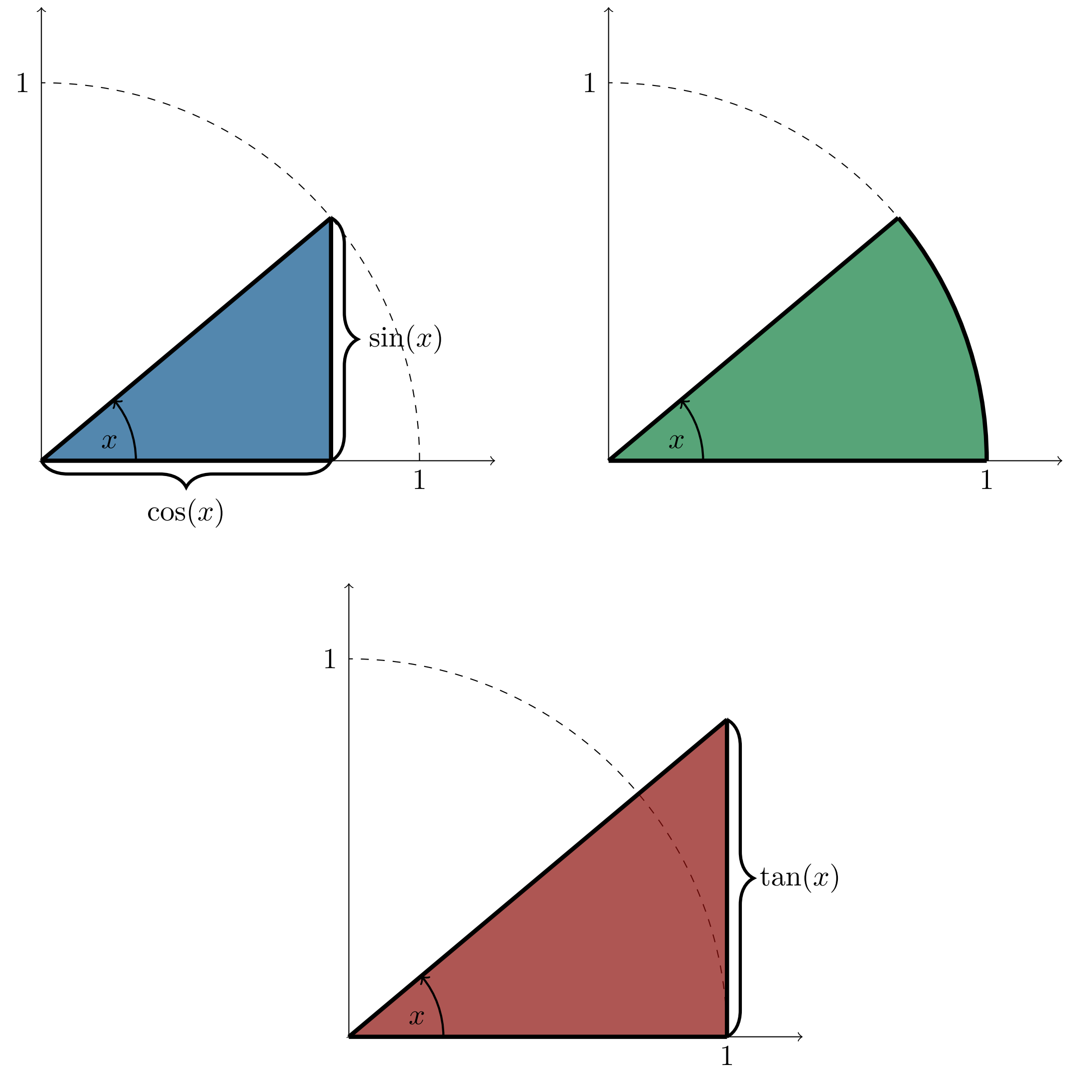


Theorem 2.4.1 **(The Fundamental Trigonometric Limit)**

It holds that

$$\lim_{x \to 0} \frac{\sin(x)}{x} = 1.$$

Proof: Consider the unit circles depicted below, which depicts three regions: a small blue triangle, a sector of the circle in green, and a large red triangle.



The small blue triangle has a height of $\sin(x)$ and a base length of $\cos(x)$, and thus total area $\frac{1}{2}\sin(x)\cos(x)$.

The total unit circle has area $\pi \cdot 1^2 = \pi$. Thus, the sector of the circle in green, resembling a piece of pie with angle $x \in [0, 2\pi)$, has area $\frac{x}{2\pi} \cdot \pi = \frac{x}{2}$.

The large red triangle is a right triangle with base length 1. Recall that in a right triangle, $\tan(x) = \dfrac{\text{length opposite}}{\text{length adjacent}}$. Since the length of the adjacent side is 1, we immediately get that the length of the opposite side, or height of the triangle, is $\tan(x)$. Hence, the large triangle has area $\frac{1}{2} \cdot 1 \cdot \tan(x)$.

Clearly, the area of the small blue triangle is smaller than or equal to the area of the sector of the circle in green, which in turn is smaller than or equal to the area of the large red triangle. That is,

$$\frac{1}{2}\sin(x)\cos(x) \leqslant \frac{x}{2} \leqslant \frac{\tan(x)}{2}$$

which implies

$$\cos(x) \leqslant \frac{x}{\sin(x)} \leqslant \frac{1}{\cos(x)}$$

for $x \in (0, \pi/2)$.

Next, $\lim_{x \to 0^+} \cos(x) = 1$, and thus also $\lim_{x \to 0^+} \frac{1}{\cos(x)} = \frac{1}{1} = 1$. The Squeeze Theorem implies that

$$\lim_{x \to 0^+} \frac{\sin(x)}{x} = 1.$$

A similar argument can be used to show that $\lim_{x \to 0^-} \frac{\sin(x)}{x} = 1$.

Therefore, we have $\lim_{x \to 0} \frac{\sin(x)}{x} = 1$. □

With this limit at hand, we can solve certain limits involving trig functions.

Example 2.4.2 Compute

$$\lim_{x \to 0} \frac{\sin(5x)}{\sin(2x)}.$$

Solution:

We find, by multiplying and dividing by terms of the form kx to get terms of the form $\frac{\sin(kx)}{kx}$, that:

$$\begin{aligned}
\lim_{x \to 0} \frac{\sin(5x)}{\sin(2x)} &= \lim_{x \to 0} \left(\frac{\sin(5x)}{5x}\right) \cdot \left(\frac{2x}{\sin(2x)}\right) \cdot \left(\frac{5x}{2x}\right) \\
&= 1 \cdot 1 \cdot \frac{5}{2} \\
&= \frac{5}{2}.
\end{aligned}$$

Example 2.4.3 Compute

$$\lim_{x \to 0} \frac{\tan(3x)}{\sin(x)}.$$

Solution:

Similarly as in the previous example,

$$\begin{aligned}
\lim_{x \to 0} \frac{\tan(3x)}{\sin(x)} &= \lim_{x \to 0} \left(\frac{x}{\sin(x)}\right) \cdot \left(\frac{\sin(3x)}{3x}\right) \cdot \left(\frac{1}{\cos(3x)}\right) \cdot \left(\frac{3x}{x}\right) \\
&= 1 \cdot 1 \cdot 1 \cdot 3 \\
&= 3.
\end{aligned}$$

Section 2.4 Problems

2.4.1. Evaluate the following limits.

(a) $\lim_{x \to 0} \dfrac{\sin(3x)}{x}$

(b) $\lim_{x \to 0} \dfrac{x^2 + 2x}{x\sin(x) + 2\sin(x)}$

(c) $\lim_{x \to 0} \dfrac{2x}{\sin(x) + \sin(5x)}$

2.4.2. Prove the following for all real numbers a and b with $a, b \neq 0$.

(a) $\lim_{x \to 0} \dfrac{\sin(ax)}{\sin(bx)} = \dfrac{a}{b}$

(b) $\lim_{x \to 0} \dfrac{\tan(ax)}{\tan(bx)} = \dfrac{a}{b}$

(c) $\lim_{x \to 0} \dfrac{\tan(ax)}{\sin(bx)} = \dfrac{a}{b}$

2.4.3. Evaluate $\lim_{x \to 0} \dfrac{x^3 \sin\left(\frac{1}{x}\right)}{\sin(x)}$.

2.4.4. Evaluate $\lim_{x \to 0} \dfrac{\sin(\sin(x))}{x}$.

2.4.5. Evaluate $\lim_{x \to 0} \dfrac{\cos(x) - 1}{x}$.

Hint: Rewrite the limit by multiplying and dividing by $\cos(x) + 1$. This limit will be useful when calculating the derivative of $\sin(x)$ in Chapter 3.

2.5 Limits at Infinity and Horizontal Asymptotes

So far, we've discussed limits of functions at a point $a \in \mathbb{R}$ and limits that either exist (that is, evaluate to a value $L \in \mathbb{R}$) or that the one-sided limits exist but do not equal.

In this section, we discuss limits at infinity. That is, we consider the limits

$$\lim_{x\to\infty} f(x) \quad \text{and} \quad \lim_{x\to-\infty} f(x).$$

Let's mimic Definition 1.2.11 of the limit of a sequence $\{a_n\}$, where we let $n \to \infty$, to define the limit of a function $f(x)$, as $x \to \pm\infty$:

Definition 2.5.1
Limit at $\pm\infty$

Let $L \in \mathbb{R}$. We say that $\lim_{x\to\infty} f(x) = L$ if, for all $\varepsilon > 0$ there exists $N \in \mathbb{R}$ such that if $x > N$, then $|f(x) - L| < \varepsilon$.

Similarly $\lim_{x\to-\infty} f(x) = L$ if, for all $\varepsilon > 0$ there exists $N \in \mathbb{R}$ such that if $x < N$, then $|f(x) - L| < \varepsilon$.

Example 2.5.2

Use the formal definition of a limit to show that

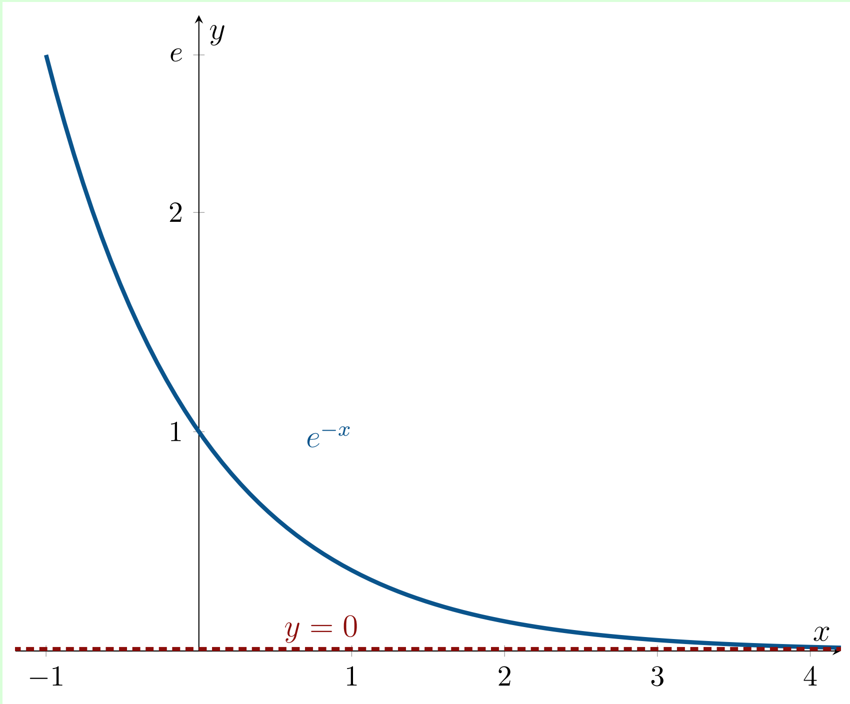
$$\lim_{x\to\infty} e^{-x} = 0.$$

Solution:

Recall that $e^{-x} = \frac{1}{e^x}$. Let $\varepsilon > 0$. Let $N = -\ln(\varepsilon)$ and $x > N$. Because the function e^x is increasing, we have $e^x > e^N$ if $x > N$. This implies $\frac{1}{e^x} < \frac{1}{e^N}$, or $e^{-x} < e^{-N}$. Hence,

$$|f(x) - 0| = |e^{-x}| = e^{-x} < e^{-N} = \varepsilon,$$

which shows the limit. The function is displayed below; note how the graph of the function approaches the horizontal line $y = 0$, as $x \to \infty$.



As seen in the example, if $\lim_{x\to\infty} f(x) = L$, then the graph of $f(x)$ approaches the line $y = L$ as x gets large. We have a name for such lines:

Definition 2.5.3
Horizontal Asymptote

If $\lim_{x\to\infty} f(x) = L$ or $\lim_{x\to-\infty} f(x) = L$ for some $L \in \mathbb{R}$, then we say $y = L$ is a **horizontal asymptote** of f.

The concept of horizontal asymptotes will be useful when we explore curve sketching later in the course.

We can also define what it means for $f(x)$ to diverge to $\pm\infty$ as $x \to \pm\infty$.

Definition 2.5.4
Divergence of $\lim_{x\to\pm\infty} f(x)$

We say that $\lim_{x\to\infty} f(x) = \infty$ if, for all $M > 0$, there exists $N \in \mathbb{R}$ so that if $x > N$ we have $f(x) > M$. The formal definitions of $\lim_{x\to\infty} f(x) = -\infty$ and $\lim_{x\to-\infty} f(x) = \pm\infty$ are defined similarly.

The good news is that even with limits at infinity or infinite limits, the Squeeze Theorem still applies!

Theorem 2.5.5

(Squeeze Theorem for functions at $\pm\infty$)

If $g(x) \leqslant f(x) \leqslant h(x)$ for all $x \geqslant N$ for some $N \in \mathbb{R}$, and if $\lim_{x\to\infty} g(x) = L = \lim_{x\to\infty} h(x)$, then $\lim_{x\to\infty} f(x) = L$ as well. The statement also holds if $L = \pm\infty$.

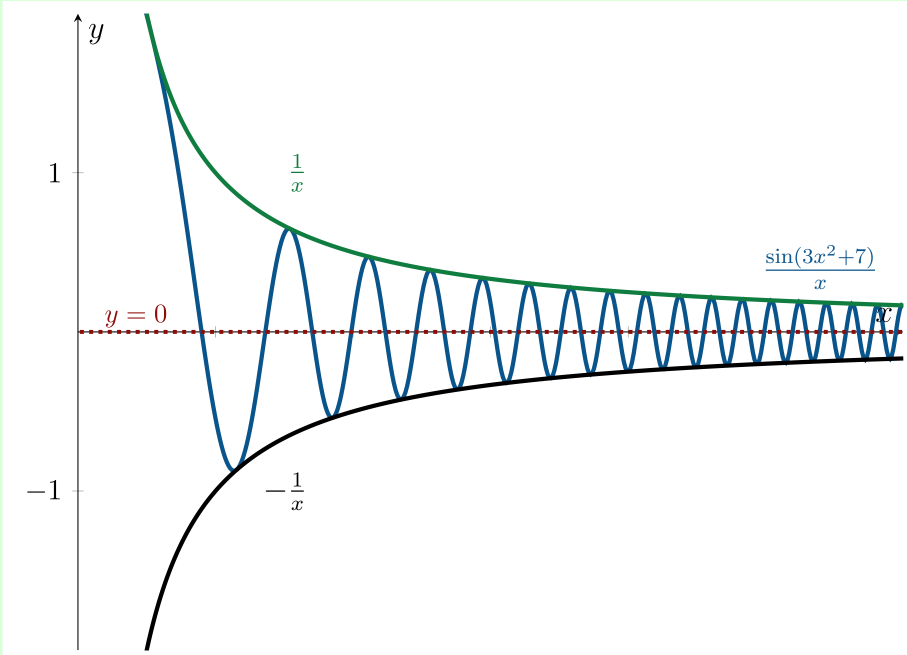
Example 2.5.6

Compute
$$\lim_{x\to\infty} \frac{\sin\left(3x^2+7\right)}{x}.$$

Solution:

Note that $-1 \le \sin\left(3x^2+7\right) \le 1$. Hence,
$$-\frac{1}{x} \le \frac{\sin\left(3x^2+7\right)}{x} \le \frac{1}{x}.$$

Also, $\lim_{x\to\infty} \frac{1}{x} = 0 = \lim_{x\to\infty} -\frac{1}{x}$. By the squeeze theorem, we get $\lim_{x\to\infty} \frac{\sin\left(3x^2+7\right)}{x} = 0$. The graph below displays the graph of function f, the horizontal asymptote $y = 0$, and the upper and lower bounds $\frac{1}{x}$ and $-\frac{1}{x}$, respectively.



REMARK

A common misconception is that the graph of a function cannot cross a horizontal asymptote. But this is not the case: The function $f(x) = \dfrac{\sin\left(3x^2+7\right)}{x}$ from Example 2.5.6 crosses its horizontal asympote $y = 0$ infinitely many times as $x \to \infty$.

2.5.1 Limits of Rational Functions

Suppose we are considering the limit of a rational function, say

$$f(x) = \frac{a_n x^n + \cdots + a_1 x + a_0}{b_m x^m + \cdots + b_1 x + b_0}.$$

The value and existence of $\lim\limits_{x\to\pm\infty} f(x)$ will depend on the relative magnitudes of the degree in the numerator, n, and the degree in the denominator, m.

For example, if $n > m$, then the numerator grows faster than the denominator, and the limit will not exist. If $n < m$, the opposite is the case, and the fraction will converge to 0. If $n = m$, then the fraction will converge to a finite limit $\dfrac{a_n}{b_n}$. We formalize this finding in the following theorem.

Theorem 2.5.7 **(Rational Function Limits at Infinity)**

Let
$$f(x) = \frac{a_n x^n + \cdots + a_1 x + a_0}{b_m x^m + \cdots + b_1 x + b_0},$$

where $a_n, b_m \neq 0$.

- If $n = m$, then $\lim\limits_{x \to \pm\infty} f(x) = \dfrac{a_n}{b_m}$.
- If $n < m$, then $\lim\limits_{x \to \pm\infty} f(x) = 0$.
- If $n > m$, then $\lim\limits_{x \to \infty} f(x) = \pm\infty$ depending on the sign of $\frac{a_n}{b_m}$.
- If $n > m$, then $\lim\limits_{x \to -\infty} f(x) = \pm\infty$ depending on the sign of $\frac{a_n}{b_m}$ as well as whether $n - m$ is an even or odd integer.

Proof: We consider the case $n = m$, the other cases are left as exercises. By factoring out and cancelling the largest appearing power of x in the numerator and denominator, we get

$$\begin{aligned}
\lim_{x \to \infty} f(x) &= \lim_{x \to \infty} \frac{a_n x^n + \cdots + a_1 x + a_0}{b_m x^m + \cdots + b_1 x + b_0} \\
&= \lim_{x \to \infty} \frac{x^n \left(a_n + a_{n-1} \cdot \frac{1}{x} + \cdots + a_1 \cdot \frac{1}{x^{n-1}} + a_0 \cdot \frac{1}{x^n}\right)}{x^n \left(b_n + b_{n-1} \cdot \frac{1}{x} + \cdots + b_1 \cdot \frac{1}{x^{n-1}} + b_0 \cdot \frac{1}{x^n}\right)} \\
&= \lim_{x \to \infty} \frac{a_n + a_{n-1} \cdot \frac{1}{x} + \cdots + a_1 \cdot \frac{1}{x^{n-1}} + a_0 \cdot \frac{1}{x^n}}{b_n + b_{n-1} \cdot \frac{1}{x} + \cdots + b_1 \cdot \frac{1}{x^{n-1}} + b_0 \cdot \frac{1}{x^n}} \\
&= \frac{a_n + a_{n-1} \cdot 0 + \cdots + a_1 \cdot 0 + a_0 \cdot 0}{b_n + b_{n-1} \cdot 0 + \cdots + b_1 \cdot 0 + b_0 \cdot 0} \\
&= \frac{a_n}{b_n}.
\end{aligned}$$

□

Example 2.5.8 Find the following limits:

1. $\lim\limits_{x \to \infty} \dfrac{2x^2 - 3x + 7}{x^2 - 4x + 5}$,
2. $\lim\limits_{x \to -\infty} \dfrac{x^2 + 2x + 1}{x - 7}$.

Solution:

1. Since the degree of the numerator equals the degree of the denominator, Theorem 2.5.7 implies the limit exists and is the ratio of the leading coefficients, given by $\frac{2}{1}$. Indeed, we find
$$\lim_{x \to \infty} \frac{2x^2 - 3x + 7}{x^2 - 4x + 5} = \lim_{x \to \infty} \frac{x^2 \left(2 - \frac{3}{x} + \frac{7}{x^2}\right)}{x^2 \left(1 - \frac{4}{x} + \frac{5}{x^2}\right)} = \frac{2}{1} = 2.$$
2. Similarly, since the degree of the numerator is larger than that of the denominator, we known from Theorem 2.5.7 that the limit will not exist. Indeed,
$$\lim_{x \to -\infty} \frac{x^2 + 2x + 1}{x - 7} = \lim_{x \to -\infty} \frac{x + 2 + \frac{1}{x}}{1 - \frac{7}{x}} = -\infty.$$

EXERCISE

Compute

$$\lim_{x\to-\infty} \frac{\cos(3x+2)+2}{x^3+1}.$$

Section 2.5 Problems

2.5.1. Evaluate the following limits.

(a) $\displaystyle\lim_{x\to\infty} \frac{2x^4-3x+4}{x^4+x+8}$

(b) $\displaystyle\lim_{x\to\infty} \frac{1}{\sqrt{x}-\sqrt{x+1}}$

(c) $\displaystyle\lim_{x\to-\infty} \frac{\cos(x+1)}{x+2}$

2.5.2. The following approach allows one to convert limits to $-\infty$ into limits to $+\infty$ (which some students may find easier to think about):

$$\lim_{x\to-\infty} f(x) = \lim_{x\to\infty} f(-x).$$

Use this approach to compute

$$\lim_{x\to-\infty} \frac{1+x+(-x)^{3/2}}{1+2(-x)^{3/2}}.$$

2.5.3. How many horizontal asymptotes can a function have? Explain your reasoning.

2.5.4. Find all horizontal asymptotes of the function

$$f(x) = \frac{\sqrt{x^2+1}}{x+3}.$$

2.5.5. Find all horizontal asymptotes for the function

$$f(x) = \frac{e^{2x}+1}{e^x+1}.$$

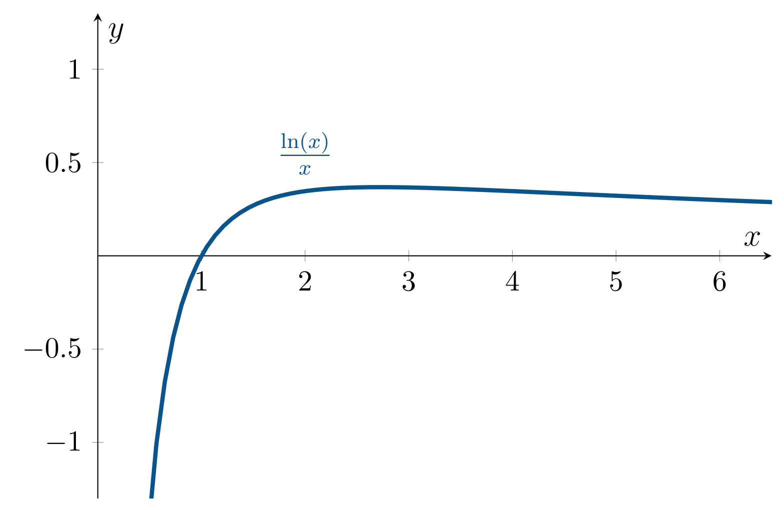
2.6 Fundamental Log Limit

Our goal in this section is to use the Squeeze Theorem to derive the limit $\lim_{x\to\infty} \frac{\ln(x)}{x}$. Due to its importance, we summarize it in the following theorem.

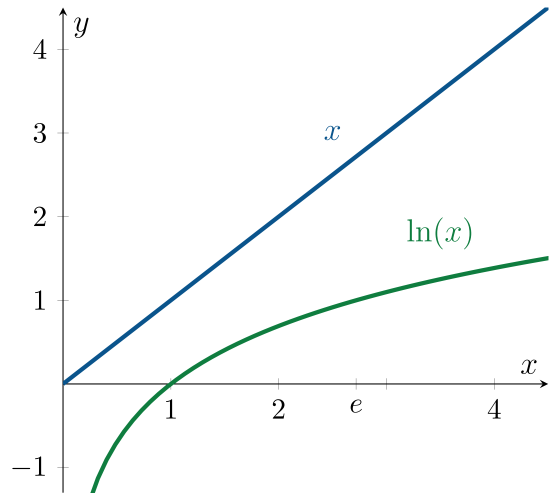
Theorem 2.6.1 (Fundamental Logarithmic Limit)

It holds that

$$\lim_{x\to\infty} \frac{\ln(x)}{x} = 0.$$



Proof: By inspecting the graphs of the functions $y = x$ and $y = \ln(x)$, we see that $\ln(x) < x$ for all $x > 0$, which implies $\frac{\ln(x)}{x} \leq 1$ for all $x > 0$.



As we are taking the limit $x \to \infty$, we can assume that $x \geq 1$. On this interval, $\ln(x) \geq 0$ and we get $\frac{\ln(x)}{x} \geq 0$. For the upper bound, note that

$$0 \le \frac{\ln(x)}{x} = \frac{\ln\left(\sqrt{x}^2\right)}{\sqrt{x}\cdot\sqrt{x}} = \frac{2}{\sqrt{x}} \cdot \frac{\ln(\sqrt{x})}{\sqrt{x}} \le \frac{2}{\sqrt{x}}$$

because $\frac{\ln(\sqrt{x})}{\sqrt{x}} \le 1$. Since $\lim_{x\to\infty} \frac{2}{\sqrt{x}} = 0$, and because $0 \le \frac{\ln(x)}{x} \le \frac{2}{\sqrt{x}}$, the Squeeze Theorem implies $\lim_{x\to\infty} \frac{\ln(x)}{x} = 0$. □

Theorem 2.6.1 indicates that the linear function x grows much faster than the logarithmic function. What about other powers of x?

Example 2.6.2 Determine the limit

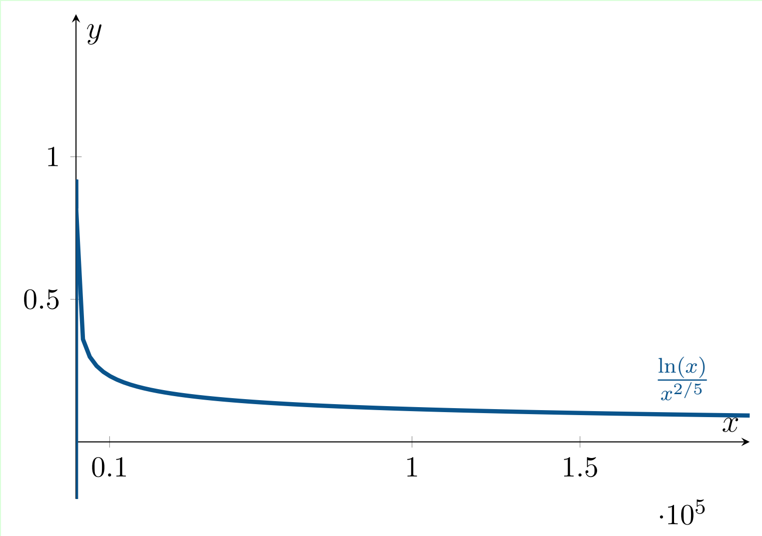
$$\lim_{x\to\infty} \frac{\ln(x)}{x^{2/5}}.$$

Solution:

Because of the logarithm rule $\ln(x^a) = a\ln(x)$, we find

$$\lim_{x\to\infty} \frac{\ln(x)}{x^{2/5}} = \lim_{x\to\infty} \frac{\frac{5}{2}\ln\left(x^{2/5}\right)}{x^{2/5}} = \frac{5}{2}\lim_{x\to\infty} \frac{\ln\left(x^{2/5}\right)}{x^{2/5}} = \frac{5}{2}\cdot 0 = 0.$$

The function $x^{2/5}$ increases very slowly, and as such, the function $\frac{\ln(x)}{x^{2/5}}$ converges very slowly to 0, as can be seen from the graph below.



EXERCISE

Show that

$$\lim_{x\to\infty} \frac{\ln(x)}{x^p} = 0$$

for any $p > 0$.

Example 2.6.3 Find the following limits:

1. $\lim_{x\to\infty} \frac{\ln\left(x^p\right)}{x}$.
2. $\lim_{x\to\infty} \frac{\ln\left(x^{100}\right)}{\sqrt{x}}$.

Solution:

- Using the logarithm rules and the limit from Theorem 2.6.1, we get
$$\lim_{x\to\infty} \frac{\ln\left(x^p\right)}{x} = \lim_{x\to\infty} p\frac{\ln(x)}{x} = p \cdot 0 = 0.$$
- Noting that $\sqrt{x} = x^{1/2}$, we get
$$\lim_{x\to\infty} \frac{\ln\left(x^{100}\right)}{\sqrt{x}} = \lim_{x\to\infty} \frac{\ln\left((x^{1/2})^{200}\right)}{\sqrt{x}} = \lim_{x\to\infty} 200\frac{\ln\left(\sqrt{x}\right)}{\sqrt{x}} = 200 \cdot 0 = 0.$$

Since the exponential function is the inverse of the logarithmic function, we can use Theorem 2.6.1 to find limits of exponential functions, as the following example illustrates.

Example 2.6.4 Let $p \in \mathbb{R}$ with $p > 0$. Find
$$\lim_{x\to\infty} \frac{x^p}{e^x}.$$

Solution:

We let $u = e^x$, so that $x = \ln(u)$. If $x \to \infty$, then $u = e^x \to \infty$ as well. This means
$$\begin{aligned}\lim_{x\to\infty} \frac{x^p}{e^x} &= \lim_{u\to\infty} \frac{\ln(u)^p}{u}\\ &= \lim_{u\to\infty} \left(\frac{\ln(u)}{u^{1/p}}\right)^p\\ &= 0^p = 0.\end{aligned}$$

We can also compute limits as $x \to 0^+$, by using a substituition:

Example 2.6.5 Let $p \in \mathbb{R}$ with $p > 0$. Find
$$\lim_{x\to 0^+} x^p \ln(x).$$

Solution:

Let $u = 1/x$ or $x = 1/u$. So as x approaches 0 from the right, i.e., as $x \to 0^+$, we have $u \to \infty$. Thus,

$$\begin{aligned}\lim_{x\to 0^+} x^p \ln(x) &= \lim_{u\to\infty} \frac{\ln\left(\frac{1}{u}\right)}{u^p} \\ &= \lim_{u\to\infty} -\frac{\ln(u)}{u^p} = 0,\end{aligned}$$

where we used that for $a, b > 0$, we have $\ln(a/b) = \ln(a) - \ln(b)$; in particular, $\ln(1/x) = \ln(1) - \ln(x) = -\ln(x)$. This example shows that x^p decreases faster to 0 than $\ln(x)$ goes to $-\infty$ as $x \to 0^+$.

Section 2.6 Problems

2.6.1. Use the fact that $\lim\limits_{x\to\infty} \dfrac{\ln(x)}{x} = 0$ and $\lim\limits_{x\to\infty} \dfrac{e^x}{x} = \infty$ to evaluate the following limits.

(a) $\lim\limits_{x\to\infty} \dfrac{\ln(x)}{x^3}$

(b) $\lim\limits_{x\to\infty} \dfrac{\ln(x)}{x + \sqrt{x}}$

(c) $\lim\limits_{x\to\infty} \dfrac{\ln(x)}{e^x}$

2.6.2. Use the fact that $\lim\limits_{x\to\infty} \dfrac{\ln(x)}{x^p} = 0$ and $\lim\limits_{x\to\infty} \dfrac{e^x}{x^p} = \infty$ for all $p > 0$ to evaluate the following limits.

(a) $\lim\limits_{x\to\infty} \dfrac{1 + e^x + e^{x^2}}{1 + x + x^2}$

(b) $\lim\limits_{x\to\infty} \dfrac{x + 1}{x + \sqrt{x}\,\ln(x) + 1}$

(c) $\lim\limits_{x\to\infty} e^{-x}\left(1 - x\sqrt{e^x}\right)$

2.6.3. Let $a > 0$, $a \neq 1$. Given $p > 0$, use the fact that $\lim\limits_{x\to\infty} \dfrac{\ln(x)}{x^p} = 0$ to prove that

$$\lim_{x\to\infty} \frac{\log_a(x)}{x^p} = 0.$$

2.6.4. Let $a > 1$. Given $p > 0$, use the fact that $\lim\limits_{x\to\infty} \dfrac{e^x}{x^p} = \infty$ to prove that

$$\lim_{x\to\infty} \frac{a^x}{x^p} = \infty$$

2.7 Infinite Limits and Vertical Asymptotes

If we examine a function near a point $a \in \mathbb{R}$, one- or both-sided limits could go to $\pm\infty$, a situation which we make more precise in the following:

Definition 2.7.1
Vertical Asymptote

We say that $\lim_{x \to a^+} f(x) = \infty$ if, for all $m > 0$ there exists $\delta > 0$ so that if $a < x < a + \delta$ we have $f(x) > m$.

Similarly, we say that $\lim_{x \to a^-} f(x) = \infty$ if, for all $m > 0$ there is $\delta > 0$ so that if $a - \delta < x < a$, then $f(x) > m$.

Finally, we say $\lim_{x \to a} f(x) = \infty$ if

$$\lim_{x \to a^+} f(x) = \infty = \lim_{x \to a^-} f(x).$$

If $\lim_{x \to a^-} f(x) = \pm\infty$ or $\lim_{x \to a^+} f(x) = \pm\infty$, then we say the line $x = a$ is a **vertical asymptote** of f.

This definition focuses on the case $\lim_{x \to a} f(x) = \infty$. As an exercise, write the definitions for $\lim_{x \to a^+} f(x) = -\infty$, $\lim_{x \to a^-} f(x) = -\infty$, and $\lim_{x \to a} f(x) = -\infty$. We also remind the reader that saying a limit is infinity means that the limit does not exist as a real number but instead grows without bound.

Example 2.7.2

Determine the following limits:

1. $\lim_{x \to 1^-} \dfrac{x^2 + 1}{x - 1}$,
2. $\lim_{x \to 3^+} \dfrac{(x + 1)(x - 7)}{(x - 3)(x - 1)}$.

Solution:

1. We already know the limit is $\pm\infty$, because the numerator goes to $2 \neq 0$ while the denominator approaches 0 as $x \to 1^-$. But is the limit $+\infty$ or $-\infty$? If $x \to 1^-$, then $x \to 1$ and $x < 1$, so $x^2 + 1 > 0, x - 1 < 0$, which means the whole function is negative, therefore $\lim_{x \to 1^-} \dfrac{x^2 + 1}{x - 1} = -\infty$.
2. Note as $x \to 3$ while $x > 3$, we have that the denominator $(x-3)(x-1) > 0$ approaches zero from above while the numerator $(x+1)(x-7)$ approaches $4 \cdot (-4) = -16$. Overall, we find
$$\lim_{x \to 3^+} \frac{(x + 1)(x - 7)}{(x - 3)(x - 1)} = -\infty.$$

REMARK

In the previous example, we argued that $\lim_{x\to 3^+} \frac{(x+1)(x-7)}{(x-3)(x-1)} = -\infty$ by noting that in the numerator, $\lim_{x\to 3^+}(x+1)(x-7) = -16 < 0$ and $\lim_{x\to 3^+}(x-3)(x-1) = 0$ from above. It can be helpful to make use of the abuse of notation $\lim_{x\to 3^+}(x-3)(x-1) = 0^+$, interpreted as a tiny positive number. Similarly, $\lim_{x\to 3^+}(3-x)(x-1) = 0^-$, interpreted as a tiny negative number. With this abuse of notation, we can write

$$\lim_{x\to 3^+} \frac{(x+1)(x-7)}{(x-3)(x-1)} = \frac{(4)(-3)}{(0^+)(2)};$$

from this expression it is easy to see that the numerator is negative, while the denominator is positive (with limit 0), which overall gives a negative limit $\lim_{x\to 3^+} \frac{(x+1)(x-7)}{(x-3)(x-1)} = -\infty$.

In the next example, we combine what we have learned in the previous sections and find all vertical and horizontal asymptotes.

Example 2.7.3 Find all vertical/horizontal asymptotes for

$$f(x) = \frac{x-3}{x-1}.$$

Solution:

We first inspect the behaviour of f as x approaches infinity. We find

$$\lim_{x\to\pm\infty} \frac{x-3}{x-1} = \lim_{x\to\pm\infty} \frac{1-\frac{3}{x}}{1-\frac{1}{x}} = 1,$$

so f has a horizontal asymptote at $y = 1$. Furthermore,

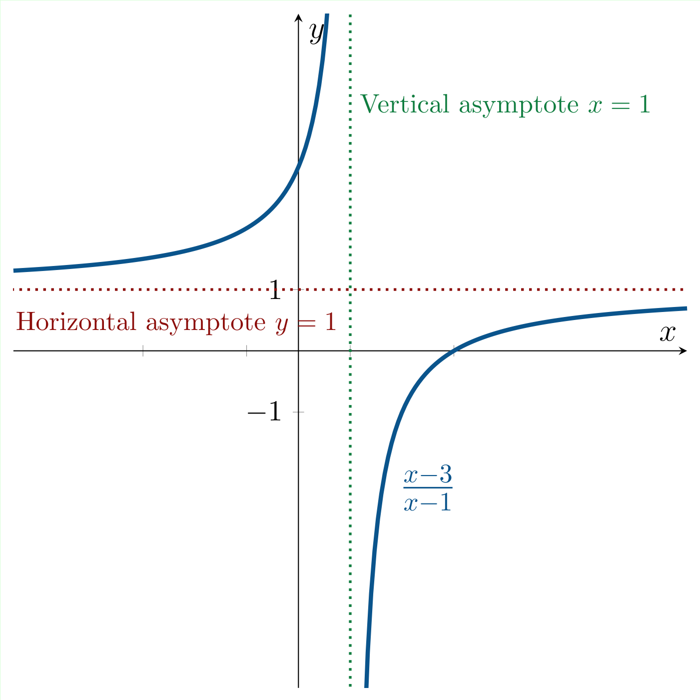
$$\lim_{x\to 1^+} \frac{x-3}{x-1} = \frac{-2}{0^+} = -\infty.$$

This guarantees that f has a vertical asymptote at $x = 1$.

Though it is not required to inspect the both one-sided limits, we do so as it allows us to better understand the shape of the graph. We find

$$\lim_{x\to 1^-} \frac{x-3}{x-1} = \frac{-2}{0^-} = +\infty.$$

The graph of f, along with the two asymptotes, are shown below.



Example 2.7.4 Find all vertical/horizontal asymptotes for

$$f(x) = \frac{x}{\sqrt{x^2+1}}.$$

Solution:

Again, we first inspect the behaviour of f as x approaches infinity. First note that

$$\frac{x}{\sqrt{x^2+1}} = \frac{x}{\sqrt{x^2}\sqrt{1+\frac{1}{x^2}}} = \frac{x}{|x|\sqrt{1+\frac{1}{x^2}}}.$$

The absolute value will, in fact, lead to there being two horizontal asymptotes for this function. Indeed, if $x \to \infty$, we have $x > 0$, so $|x| = x$. This implies

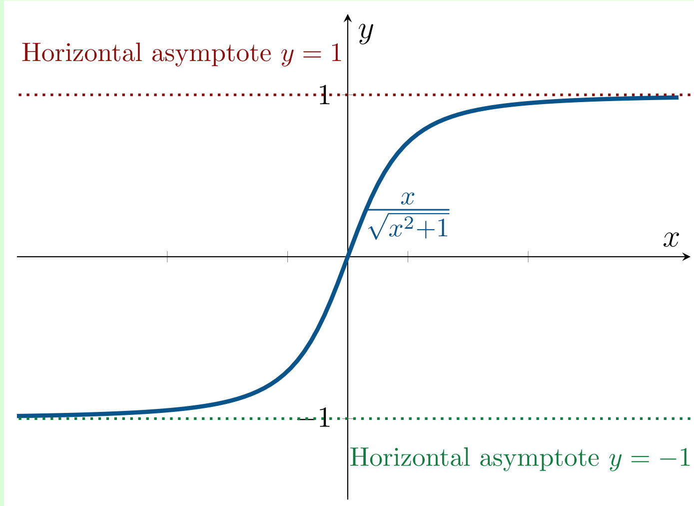
$$\lim_{x\to+\infty} f(x) = \lim_{x\to+\infty} \frac{x}{|x|\sqrt{1+\frac{1}{x^2}}} = \lim_{x\to\infty} \frac{1}{\sqrt{1+\frac{1}{x^2}}} = 1.$$

If $x \to -\infty$, then $x < 0$, so $|x| = -x$. In this case,

$$\lim_{x\to-\infty} f(x) = \lim_{x\to-\infty} \frac{x}{|x|\sqrt{1+\frac{1}{x^2}}} = \lim_{x\to-\infty} \frac{-1}{\sqrt{1+\frac{1}{x^2}}} = -1.$$

Thus, the function f has two horizontal asymptotes at $y = 1$ and $y = -1$.

Since $x^2 \geq 0$ for all $x \in \mathbb{R}$, we have $\sqrt{x^2+1} > 0$ for all $x \in \mathbb{R}$, so the denominator is never zero. There are no vertical asymptotes.



REMARK

If a function has a vertical asymptote, it is possible for its graph to touch the asymptote at a single point (but, by definition of a function, the graph cannot touch it more than once). An example of this would be the function

$$f(x) = \begin{cases} \frac{1}{x}, & \text{if } x \neq 0, \\ 0, & \text{otherwise,} \end{cases}$$

which has a vertical asymptote at $x = 0$ and $f(0) = 0$.

Section 2.7 Problems

2.7.1. Evaluate $\lim\limits_{x \to 1} \dfrac{3-x}{(x-1)^2}$.

2.7.2. Evaluate $\lim\limits_{x \to 3^+} \dfrac{x^2+10}{x^2-7x+12}$.

2.7.3. Determine all vertical asymptotes of the function $f(x) = \dfrac{\sqrt{x+2}}{x^2+4x+4}$.

2.7.4. Determine all vertical asymptotes of the function $f(x) = \dfrac{\sin(x)}{2x^2-\pi x}$.

2.7.5. A line $y = mx + b$ is said to be a **slant asymptote** (also known as an **oblique asymptote**) of a function f if

$$\lim_{x \to \infty} [f(x) - (mx+b)] = 0 \quad \text{or} \quad \lim_{x \to -\infty} [f(x) - (mx+b)] = 0.$$

Graphically, if $y = mx + b$ is a slant asymptote of f, then the graph of f becomes very close to the line $y = mx + b$ as $x \to \infty$ or $x \to -\infty$.

Find the slant asymptote for each function below. Try graphing these functions using software to see its behavior along the slant asymptote.

(a) $f(x) = x + \dfrac{1}{x}$

(b) $f(x) = \dfrac{2x^2 + 10\sin(x)}{x}$

(c) $f(x) = \dfrac{-x^3 + x^2 - x + 31}{x^2 + 1}$

Hint: Polynomial long division

2.8 Continuity

2.8.1 Introduction to Continuous Functions

We have encountered situations in this chapter where the limit $\lim_{x \to a} f(x)$ and the function value $f(a)$ at $x = a$ do not coincide and cases where they do coincide. In the latter case we call the function continuous.

Definition 2.8.1
Continuous Function

A function f is **continuous** at $x = a$ if:

1. The function $f(x)$ is defined at $x = a$,
2. $\lim_{x \to a} f(x)$ exists, and
3. $\lim_{x \to a} f(x) = f(a)$

Otherwise, we say f is **discontinuous** $x = a$, or that $x = a$ is a point of discontinuity for f.

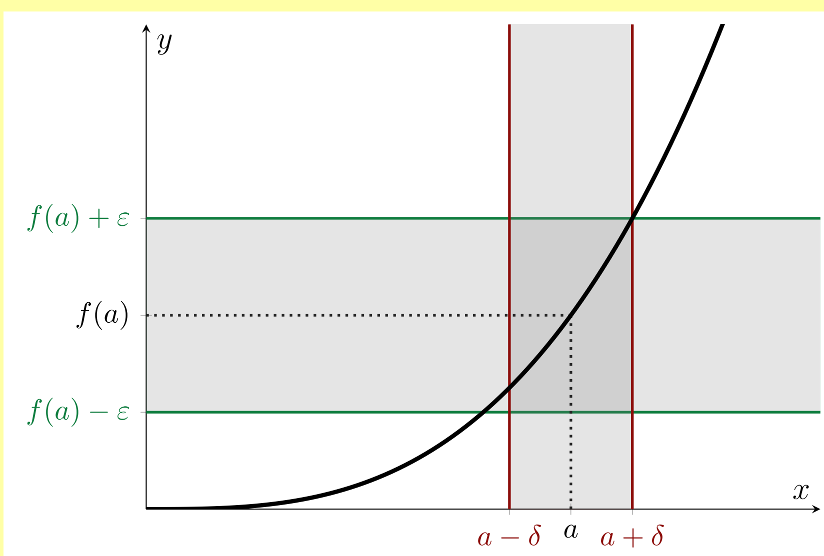
Intuitively, a function is continuous at $x = a$ if its behavior at $x = a$ is determined by its behaviour near $x = a$. Roughly speaking, we can draw the graph of a continuous function without lifting the pen.

We can also give an equivalent definition of a continuous function using $\varepsilon - \delta$ notation, which is illustrated below.

Definition 2.8.2
$\varepsilon - \delta$ Definition of a Continuous Function

A function f is continuous at $x = a$ if for all $\varepsilon > 0$ there exists $\delta > 0$ such that if $|x - a| < \delta$, then

$$|f(x) - f(a)| < \varepsilon.$$



Using the relationship between limits and sequences, we can give a third definition:

Definition 2.8.3

Sequential Definition of a Continuous Function

A function f is continuous at $x = a$ if, whenever $\{x_n\}$ is a sequence where $\lim_{n\to\infty} x_n = a$, we have $\lim_{n\to\infty} f(x_n) = f(a)$.

Example 2.8.4

Is the function

$$f(x) = \frac{x+1}{x-7}$$

continuous at $x = 1$?

Solution:

We find that $f(1) = \frac{2}{-6} = -\frac{1}{3}$. On the other hand,

$$\lim_{x\to 1} \frac{x+1}{x-7} = \frac{2}{-6} = \frac{-1}{3}.$$

Hence, f is continuous at $x = 1$.

Example 2.8.5

Is the function

$$f(x) = |x|$$

continuous at $x = 0$?

Solution:

Recall that

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

Since the function is piecewise defined, we compute one-sided limits at $x = 0$. Hence,

$$\lim_{x\to 0^+} |x| = \lim_{x\to 0^+} x = 0$$

and

$$\lim_{x\to 0^-} |x| = \lim_{x\to 0^-} -x = 0.$$

As such,

$$\lim_{x\to 0} |x| = 0 = |0| = f(0)$$

which means that $f(x) = |x|$ is continuous at $x = 0$.

Example 2.8.6

Is the function

$$f(x) = \frac{1}{x}$$

continuous at $x = 0$?

Solution:

There are two reasons why this function is not continuous at $x = 0$: First, the function f is not even defined at $x = 0$. Second, $\lim_{x \to 0} \frac{1}{x}$ does not exist. Therefore, $f(x)$ is discontinuous at $x = 0$.

2.8.2 Continuity of Basic Functions

Theorem 2.8.7 **(Continuity of Polynomials)**

Let $p(x) = a_n x^n + \cdots + a_1 x + a_0$ be a polynomial. Then p is continuous for all $x \in \mathbb{R}$.

Proof: In Theorem 2.2.6, we have shown that $\lim_{x \to a} p(x) = p(a)$ for all $a \in \mathbb{R}$ for a polynomial p. Hence, p is continuous by definition. □

A useful observation to take the limit $\lim_{x \to a} f(x)$ is the following: If we write $x = a + h$ for some $h \in \mathbb{R}$ with $h \neq 0$, then $x \to a$ if and only if $h \to 0$. Hence, we can say that f is continuous at $x = a$ if $\lim_{x \to a} f(x) = \lim_{h \to 0} f(a + h) = f(a)$.

Theorem 2.8.8 **(Continuity of $\sin(x)$ and $\cos(x)$)**

The functions $\sin(x)$ and $\cos(x)$ are continuous everywhere.

Proof: First, let's show $\lim_{x \to 0} \sin(x) = \sin(0) = 0$. Note that for $0 < x < \frac{\pi}{2}$, we have $0 < \sin(x) < x$. Since $\lim_{x \to 0^+} 0 = 0 = \lim_{x \to 0^+} x = 0$, it follows from the Squeeze Theorem that $\lim_{x \to 0^+} \sin(x) = 0$.

Next, we know $\sin(-x) = -\sin(x)$, and if $x \to 0^-$ then $-x \to 0^+$, so

$$\lim_{x \to 0^-} \sin(x) = \lim_{x \to 0^-} -\sin(-x) = \lim_{(-x) \to 0^+} -\sin(-x) = (-1)(0) = 0.$$

Overall, we get $\lim_{x \to 0} \sin(x) = 0 = \sin(0)$, so that $\sin(x)$ is continuous at $x = 0$.

Next, recall that $\sin^2(x) + \cos^2(x) = 1$. Thus, around $x = 0$, we have $\cos(x) = \sqrt{1 - \sin^2(x)}$, where here we are using the positive square root since $\cos(x)$ is positive near $x = 0$. Consequently,

$$\lim_{x \to 0} \cos(x) = \lim_{x \to 0} \sqrt{1 - \sin^2(x)} = \sqrt{1 - 0} = 1 = \cos(0),$$

which shows that $\cos(x)$ is continuous at $x = 0$.

Second, let $a \in \mathbb{R}$ be given. Let's prove that $\lim_{x \to a} \sin(x) = \sin(a)$. Indeed, using $\sin(a + b) = \sin(a)\cos(b) + \sin(b)\cos(a)$, we get

$$\begin{aligned}
\lim_{x \to a} \sin(x) &= \lim_{h \to 0} \sin(a + h) \\
&= \lim_{h \to 0} \sin(a)\cos(h) + \sin(h)\cos(a)
\end{aligned}$$

$$\begin{aligned} &= \sin(a)(1) + (0)\cos(a) \\ &= \sin(a). \end{aligned}$$

Thus, $\sin(x)$ is continuous at $x = a$. As an exercise, you can show that $\lim_{x \to a} \cos(x) = \cos(a)$. □

Theorem 2.8.9 **(Continuity of e^x)**

The exponential function e^x is continuous everywhere.

Proof: This result is surprisingly hard to show! For a rigorous proof, we would need more tools to work the function e^x, such as power series and Taylor series from MATH 138. To simplify things, we will take the following fact for granted:

$$e^x \text{ is continuous at } x = 0, \text{ i.e., } \lim_{x \to 0} e^x = e^0 = 1.$$

With this fact at hand, we can now prove that for all $a \in \mathbb{R}$, it holds that $\lim_{x \to a} e^x = e^a$. We can assume that $a \neq 0$, because we know the statement holds for $a = 0$. Then

$$\lim_{x \to a} e^x = \lim_{h \to 0} e^{a+h} = \lim_{h \to 0} e^a e^h = e^a \cdot 1 = e^a.$$

□

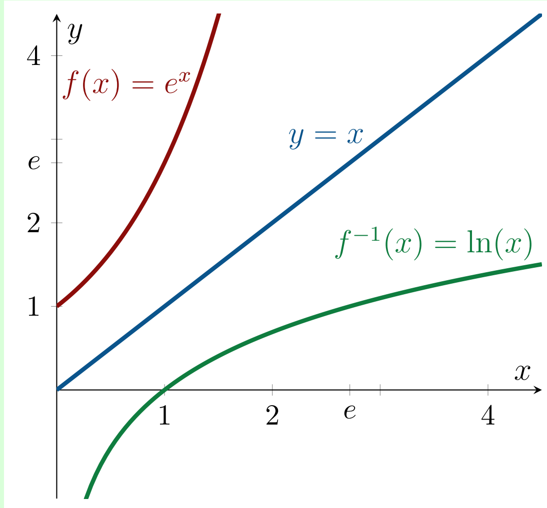
Next, we wish to show that the logarithmic function $\ln(x)$ is continuous. This will follow because $\ln(x)$ is the inverse function of e^x, and because e^x is continuous, so is $\ln(x)$. We show this relationship in the following theorem:

Theorem 2.8.10 **(Continuity of the Inverse Function)**

Let f be continuous and invertible at $x = a$ with $f(a) = b$. Then the inverse $f^{-1}(y)$ is continuous at $y = b$.

Rather than giving a formal proof, we give an intuitive argument: Recall that in order to get the graph of $f^{-1}(x)$, we would reflect the graph of $f(x)$ over the line $y = x$; this is illustrated in the example below. So, if $f(x)$ is continuous, reflecting it wont create any discontinuities, meaning that the inverse function $f^{-1}(y)$ is also continuous.

Example 2.8.11 It follows from Theorem 2.8.10 that $\ln(x)$ is continuous at $x = a$ for all $a > 0$, since it is the inverse of the function $f(x) = e^x$ which is continuous at $x = a$ for all $a \in \mathbb{R}$.



2.8.3 Arithmetic Rules for Continuity

Similarly to how we derived arithmetic rules for limits, we can derive rules for continuous functions.

Theorem 2.8.12 **(Arithmetic Rules for Continuous Functions)**

Let f and g be continuous functions at $x = a$. Then,

1. For any constant $c \in \mathbb{R}$, $cf(x)$ is continuous at $x = a$.
2. $f(x) + g(x)$ is continuous at $x = a$.
3. $f(x) \cdot g(x)$ is continuous $x = a$.
4. If $g(a) \neq 0$, then $\frac{f(x)}{g(x)}$ is continuous $x = a$.

Proof: The proof of each rule is an immediate consequence of the corresponding limit rules. $\square$

Example 2.8.13 Determine where the function

$$f(x) = \begin{cases} -\frac{3}{2}, & \text{if } x = 1, \\ 1, & \text{if } x = 3, \\ \frac{x^2+x-2}{x^2-4x+3} & \text{otherwise.} \end{cases}$$

is continuous.

Solution:

Note that for $a \notin \{1,3\}$, the function $\dfrac{x^2+x-2}{x^2-4x+3} = \dfrac{(x-1)(x+2)}{(x-1)(x-3)}$ is continuous at $x = a$ because all component functions are continuous and the denominator satisfies $(x-1)(x-3) \neq 0$ for $x \notin \{1,3\}$.

As such, the only possible discontinuities are at $x = 1$ and $x = 3$.

For $x = 1$, we find

$$\lim_{x\to 1} f(x) = \lim_{x\to 1} \frac{(x-1)(x+2)}{(x-1)(x-3)} = \lim_{x\to 1} \frac{x+2}{x-3} = -\frac{3}{2} = f(1),$$

so the function $f(x)$ is continuous at $x = 1$.

Next, for $x = 3$, we find

$$\lim_{x\to 3^+} f(x) = \lim_{x\to 3^+} \frac{(x-1)(x+2)}{(x-1)(x-3)} = \frac{(2)(5)}{(2)(0^+)} = +\infty$$

which shows that f is discontinuous at $x = 3$.

Overall, f is continuous at $x = a$ for every $a \in \mathbb{R}$ with $a \neq 3$.

Theorem 2.8.14 **(Continuity of the Composition of Functions)**

If f is continuous at $x = a$ and g is continuous at $x = f(a)$, then $h = g \circ f$ is continuous at $x = a$.

Proof: Let's prove this theorem using the sequential characterization of continuity. Suppose f is continuous at $x = a$, and that g is continuous at $x = f(a)$. Let $h = g \circ f$, that is, $h(x) = g(f(x))$. Suppose $\{x_n\}$ is any sequence with $\lim_{n\to\infty} x_n = a$. Then, since f is continuous at $x = a$, we know

$$\lim_{n\to\infty} f(x_n) = f(a).$$

As such, $\{f(x_n)\}$ is a sequence with $\lim_{n\to\infty} f(x_n) = f(a)$. Next, since the function g is continuous at $x = f(a)$, we get

$$\lim_{n\to\infty} g(f(x_n)) = g(f(a)).$$

Overall,

$$\lim_{n\to\infty} h(x_n) = \lim_{n\to\infty} g(f(x_n)) = g(f(a)) = h(a).$$

Since the sequence $\{x_n\}$ was arbitrary, we conclude that h is continuous at $x = a$. $\square$

Example 2.8.15 Determine whether $f(x) = \cos\left(e^{x^2}\right)$ is continuous at each $x \in \mathbb{R}$.

Solution:

Note that f is a composition of the functions $\cos(x)$, e^x and x^2, all of which are continuous functions. Hence, f is continuous by Theorem 2.8.14.

A useful technique to compute limits is the following:

Lemma 2.8.16

Suppose $\lim_{x\to a} f(x) = L$ and that the function $g(y)$ is continuous at $y = L$. Then

$$\lim_{x\to a} g\left(f(x)\right) = g(L) = g\left(\lim_{x\to a} f(x)\right)$$

Proof: This follows immediately from the continuity of the function g at $y = L$. $\square$

The significance of Lemma 2.8.16 is that we can bring limits inside continuous functions, as the following example illustrates.

Example 2.8.17

Compute

$$\lim_{x\to 1} \sin\left(\frac{\pi(x-1)}{x^2-1}\right).$$

Solution:

By Lemma 2.8.16, we find

$$\begin{aligned}
\lim_{x\to 1} \sin\left(\frac{\pi(x-1)}{x^2-1}\right) &= \sin\left(\lim_{x\to 1} \frac{\pi(x-1)}{(x-1)(x+1)}\right) && \text{(since } \sin(x) \text{ is continuous)} \\
&= \sin\left(\lim_{x\to 1} \frac{\pi}{x+1}\right) \\
&= \sin\left(\frac{\pi}{2}\right) \\
&= 1.
\end{aligned}$$

2.8.4 Continuity on an Interval

Continuity on open and closed intervals are treated differently. Firstly, we use the following definition for continuity on an open interval.

Definition 2.8.18
Continuous Function on (a, b)

We say that a function f is **continuous on** (a, b) if f is continuous for all $x \in (a, b)$. Similarly, a function f is said to be continuous on $\mathbb{R}$, if it is continuous for all $x \in \mathbb{R}$.

What about closed intervals? The problem is at the endpoints, as the function f may not be defined outside of the interval.

Example 2.8.19

Let $f(x) = \sqrt{x}$. Then the domain of f is $[0, \infty)$. In this case, we know $\lim_{x\to 0^+} \sqrt{x} = 0 = \sqrt{0}$, but the left limit $\lim_{x\to 0^-} \sqrt{x}$ is not defined (i.e., $\sqrt{x}$ is undefined for $x < 0$), which is why $\lim_{x\to 0} \sqrt{x}$ does not exist. But we would still like to say that $\sqrt{x}$ is continuous at $x = 0$, and just ignore $x < 0$.

Definition 2.8.20

Continuous Function on $[a, b]$

We say that a function f is **continuous on** $[a, b]$ if all of the following hold:

1. f is continuous on (a, b),
2. $\lim_{x \to a^+} f(x) = f(a)$, and
3. $\lim_{x \to b^-} f(x) = f(b)$.

In other words, we only consider continuity (and limits) as we approach from inside the interval in question. Based on this definition, we can say that $\sqrt{x}$ is indeed continuous on $[0, \infty)$.

Now that we know what it means for a function to be continuous, let's look at the various ways it can be discontinuous.

Section 2.8 Problems

2.8.1. Evaluate the following limits, making appropriate reference to the continuity of familiar functions.

(a) $\lim_{x \to 0} \sin\left(\frac{\sin(x)}{x}\right)$

(b) $\lim_{x \to 1} e^{\frac{x-1}{\sqrt{x}-1}}$

(c) $\lim_{x \to \infty} (\ln(3x + 1) - \ln(x + 1))$

2.8.2. Let $a > 0$, $a \neq 1$.

(a) Use the fact that $\ln(x)$ is continuous to prove that $\log_a(x)$ is continuous on $(0, \infty)$,

(b) Prove that a^x is continuous on $(-\infty, \infty)$.

2.8.3. Use the $\varepsilon - \delta$ definition of continuity to prove that $f(x) = 2x^2 + 9$ is continuous at $x = 2$.

2.8.4. Let f be a function defined as

$$f(x) = \begin{cases} \dfrac{x^2 - 4}{x^2 + x - 6} \cos(x^2) & \text{if } x \neq -3, 2, \\ 0 & \text{if } x = -3, 2. \end{cases}$$

Find the intervals where f is continuous. Justify your answer.

2.8.5. Find all values of a and b such that

$$f(x) = \begin{cases} x^2 - 1 & \text{if } x < -1, \\ ax - 2b & \text{if } -1 \leq x < 1, \\ x^2 - bx + a & \text{if } x \geq 1, \end{cases}$$

is continuous for all $x \in \mathbb{R}$. Justify your answer.

2.8.6. Show that if a function is continuous at $x = 0$ and satisfies:

(a) $f(x + y) = f(x) + f(y)$ then it is continuous everywhere;

(b) $f(x + y) = f(x)f(y)$ then it is continuous everywhere.

2.9 Types of Discontinuities

In this section, we discuss various types of discontinuities. Recall that in order for a function f to be continuous at $x = a$, we need that $f(a)$ is defined and

$$\lim_{x \to a} f(x) = f(a).$$

We define four different kinds of discontinuities.

Definition 2.9.1
Removable Discontinuity

We say that a function $f(x)$ has **removable discontinuity** at $x = a$, if

$$\lim_{x \to a} f(x) \text{ exists, but } \lim_{x \to a} f(x) \neq f(a).$$

Note that f can have a removable discontinuity at $x = a$ without being defined at $x = a$. This type of discontinuity is called removable because we could re-define $f(x)$ at $x = a$ to equal the existing limit $\lim\limits_{x \to a} f(x)$, and thereby removing the discontinuity, as the following example illustrates. These are the least serious kinds of discontinuities.

Example 2.9.2

Show that the following function has a removable discontinuity at $x = 1$.

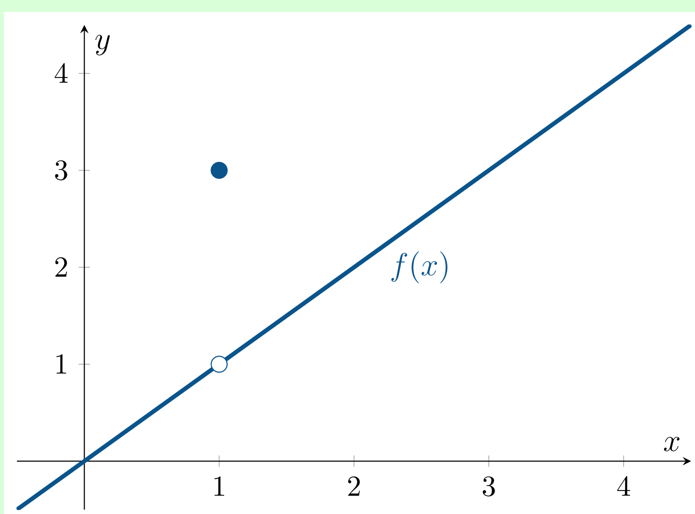
$$f(x) = \begin{cases} x, & \text{if } x \neq 1, \\ 3, & \text{if } x = 1. \end{cases}$$

Solution:

It is easy to see that

$$f(1) = 3 \neq \lim_{x \to 1} f(x) = \lim_{x \to 1} x = 1.$$

Hence, f has a removable discontinuity at $x = 1$, as can be seen from the graph below.



Definition 2.9.3
Finite Jump Discontinuity

We say that a function $f(x)$ has a **jump discontinuity** at $x = a$, if

$$\lim_{x\to a} f(x) \text{ does not exist, but both } \lim_{x\to a^+} f(x) \text{ and } \lim_{x\to a^-} f(x) \text{ exist.}$$

Note that if $f(x)$ has a jump discontinuity at $x = a$, then $\lim_{x\to a^\pm} f(x)$ existing means that these limits are finite.

Example 2.9.4

Show that the following function has a finite jump discontinuity at $x = 0$:

$$f(x) = \begin{cases} x, & \text{if } x \le 0 \\ 3, & \text{if } x > 0. \end{cases}$$

Solution:

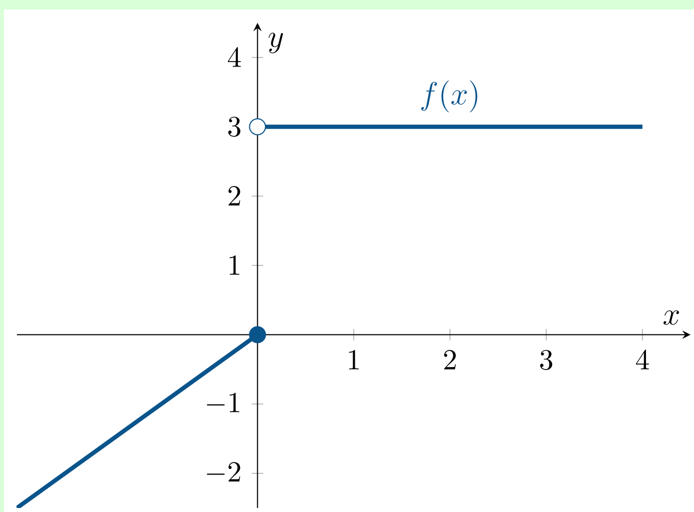
We compute

$$\lim_{x\to 0^+} f(x) = \lim_{x\to 0^+} 3 = 3$$

and

$$\lim_{x\to 0^-} f(x) = \lim_{x\to 0^-} x = 0.$$

Since both one-sided limits exist but do not equal, f has a finite jump discontinuity at $x = 0$ as can be seen from the graph below.



Definition 2.9.5
Infinite Discontinuity

We say that a function $f(x)$ has an **infinite discontinuity** at $x = a$, if

$$\text{one or both of } \lim_{x\to a^+} f(x) \text{ or } \lim_{x\to a^-} f(x) \text{ is } \pm\infty.$$

REMARK

By comparing this definition with Definition 2.7.1 of a vertical asymptote, we see that they are identical. Hence,

$$\begin{aligned} f \text{ has an infinite discontinuity at } x = a &\Leftrightarrow \text{one or both of } \lim_{x\to a^+} f(x) \text{ or } \lim_{x\to a^-} f(x) \text{ is } \pm\infty \\ &\Leftrightarrow f \text{ has a vertical asymptote at } x = a. \end{aligned}$$

Example 2.9.6 Show that the following function has an infinite discontinuity at $x = 0$:

$$f(x) = \frac{1}{x}.$$

Solution:

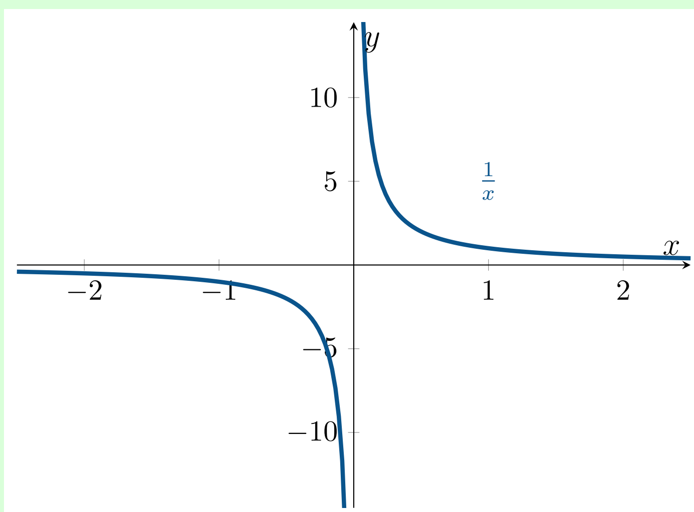
The given function f satisfies

$$\lim_{x\to 0^+} f(x) = \lim_{x\to 0^+} \frac{1}{x} = \infty,$$

hence, has an infinite discontinuity at $x = 0$. Though not necessary to classify $x = 0$ as infinite discontinuity, we also compute the other one-sided limit, as it helps us get a better understanding of how the function f behaves near $x = 0$. We find

$$\lim_{x\to 0^-} f(x) = \lim_{x\to 0^-} \frac{1}{x} = -\infty.$$

As is depicted in the figure below, we can see that as x approaches 0 from the left, $f(x)$ approaches $-\infty$, while $f(x)$ approaches $+\infty$ as x approaches 0 from the right.



Definition 2.9.7

Oscillatory Discontinuity

We say that a function $f(x)$ has an **oscillatory discontinuity** at $x = a$, if $\lim_{x\to a} f(x)$ does not exist, but f is bounded near $x = a$ and is oscillating infinitely often near $x = a$.

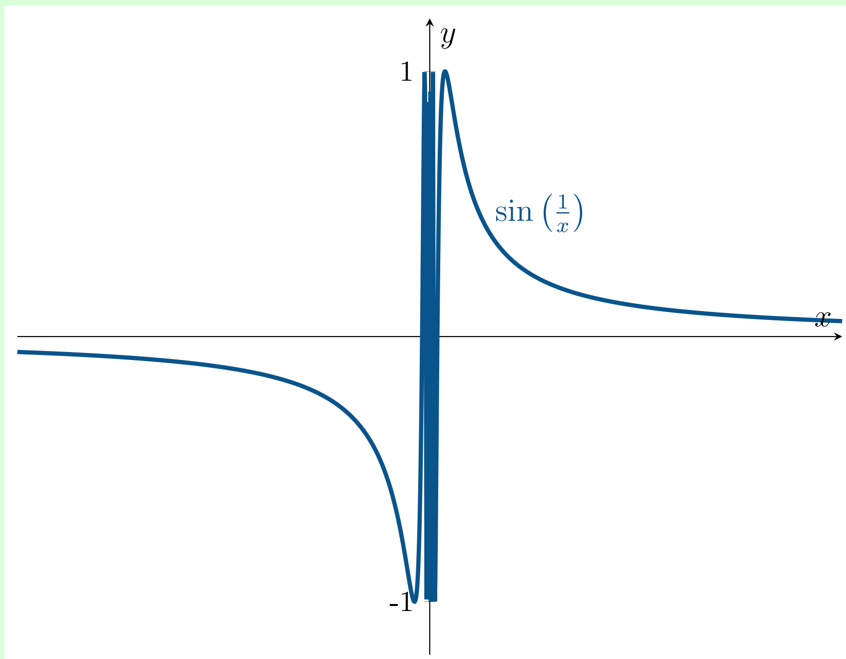
Example 2.9.8

Show that the following function has an oscillatory discontinuity at $x = 0$:

$$f(x) = \sin\left(\frac{1}{x}\right).$$

Solution:

We have seen in Example 2.2.4 that the limit $\lim_{x\to 0} f(x)$ does not exist, but the function is oscillating, as shown in the graph below. Furthermore, since $\sin(x) \leq 1$ for all $x \in \mathbb{R}$, we can see that the function f is bounded (near 0, and everywhere else). Hence, f has an oscillatory discontinuity at $x = 0$.



REMARK

If f has any type of discontinuity, except for the removable discontinuity from Definition 2.9.1, then there is no simple way of removing the discontinuity by redefining the function value $f(a)$. As such, jump discontinuities, infinite discontinuities and oscillatory discontinuities are also referred to as **essential singularities**, or **essential discontinuities.**

Section 2.9 Problems

2.9.1. Consider the function

$$f(x) = \frac{x^2 - x - 2}{x^2 - 5x + 6}.$$

Identify any points at which f is discontinuous, then classify its discontinuities.

2.9.2. Consider the function

$$f(x) = \frac{|x|}{x^2}.$$

Is f continuous at $x = 0$? If not, classify the type of discontinuity.

2.9.3. Consider the function

$$f(x) = \frac{x|1-x|}{x-1}.$$

Is f continuous at $x = 1$? If not, classify the type of discontinuity.

2.9.4. Consider the function

$$f(x) = \begin{cases} \sin^2(x)\sin\left(\frac{1}{x}\right) & \text{if } x \neq 0, \\ 1 & \text{if } x = 0. \end{cases}$$

Is f continuous at $x = 0$? If not, classify the type of discontinuity.

2.9.5. Consider the function

$$f(x) = \begin{cases} \cos(x)\sin\left(\frac{1}{x}\right) & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

Is f continuous at $x = 0$? If not, classify the type of discontinuity.

2.9.6. Suppose that $f : \mathbb{R} \to \mathbb{R}$ has a jump discontinuity at $x = 0$ with

$$\lim_{x\to 0^-} f(x) = -1, \quad \lim_{x\to 0^+} f(x) = 1, \quad \text{and} \quad f(0) = 0.$$

Determine whether the function $\cos(f(x))$ is continuous at $x = 0$, and if not, classify the discontinuity.

2.10 Intermediate Value Theorem

An important theorem that applies to continuous functions is the Intermediate Value Theorem. It essentially states that if f is a continuous function on $[a, b]$, then f will take all value between $f(a)$ and $f(b)$. We will consider two examples before stating the theorem.

Example 2.10.1 Consider a hiker that starts climbing a mountain at 10:00am. Suppose it takes them exactly 7 hours to reach the top of the mountain, where they rest and start their climb down the next morning, exactly at 10:00am again.

The next day, the hiker uses the same route, but being tired from the previous day, they need about an hour longer taking the same route backwards.

It may not be intuitive, but it is a fact that there is a position on the trial that the hiker crosses *exactly at the same time* on both the way up and the way down.

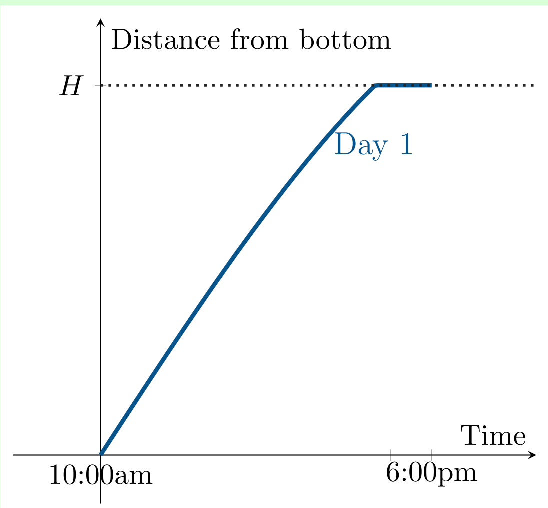
To see this, let $d_1(t)$ represent the hiker's journey up the mountain where $d_1(0) = 0$ and $d_1(7) = H$ (i.e., 10:00am corresponds to $t = 0$ and 5:00pm corresponds to $t = 7$).

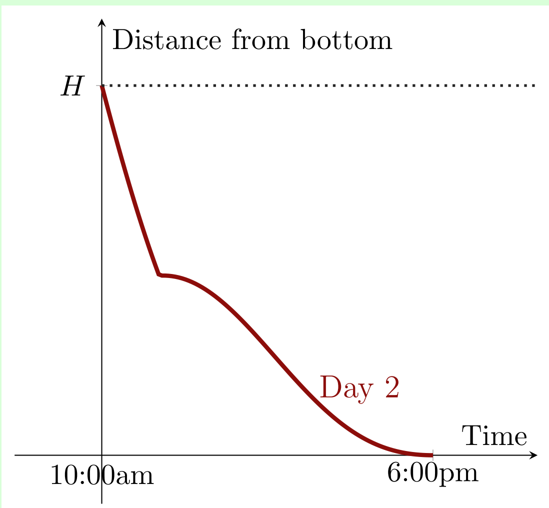
Similarly, let $d_2(t)$ represent the hiker's journey down the mountain where $d_2(0) = H$ and let $d_2(7) = h < H$ (i.e., at 5:00pm, the hiker is still some distance h from the bottom of the mountain.

Clearly, both $d_1(t)$ and $d_2(t)$ must be continuous functions (it's not possible for the hiker to teleport between locations on the mountain!)

Now, define $s(t) = d_2(t) - d_1(t)$. Notice that $s(0) = d_2(0) - d_1(0) = H - 0 = H > 0$ and that $s(7) = d_2(7) - d_1(7) = h - H < 0$. Since $s(t)$ is a continuous function and $s(0) > 0$ and $s(7) < 0$, then it *must* be the case that there is a time t_c such that $s(t_c) = d_2(t_c) - d_1(t_c) = 0 \Rightarrow d_1(t_c) = d_2(t_c)$. So the hiker must be at the same spot at the same time.

This is essentially the Intermediate Value Theorem which we will state below.



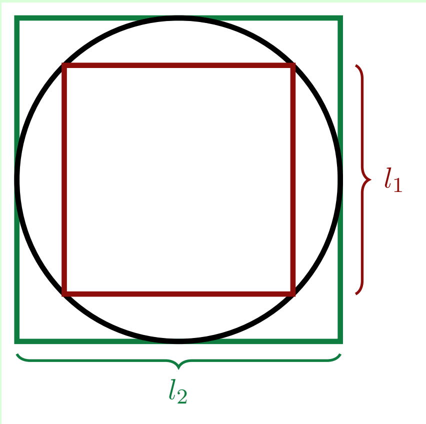


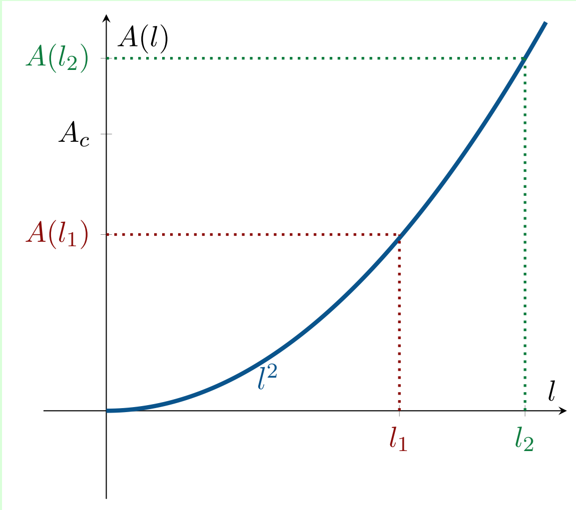
Example 2.10.2 Suppose, for a given circle with area A_c, you inscribe a square with length l_1 and circumscribe a square with length l_2, as shown in the figure below. We can show that there exists a square with side length between l_1 and l_2 that has the same area as circle, A_c, as follows.

The area of a square with side length l is given by $A(l) = l^2$ and, by construction, we have

$$A(l_1) < A_c < A(l_2).$$

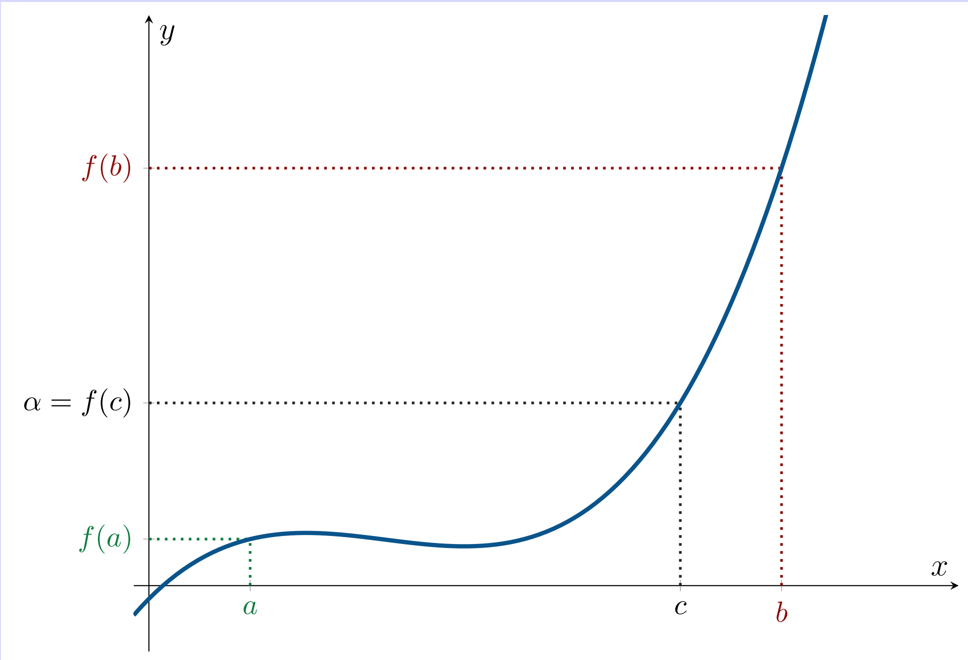
As the area function $A(l)$ is continuous in l, there *must* be a length $l \in (l_1, l_2)$ such that $A(l) = A_c$, i.e., such that the areas of the circle and square match. A version of this example was stated around 500 BCE by Bryson of Heraclea.





Theorem 2.10.3 **(Intermediate Value Theorem (IVT))**

If f is continuous on the closed interval $[a, b]$, and $\alpha \in \mathbb{R}$ is such that either $f(a) < \alpha < f(b)$ or $f(a) > \alpha > f(b)$, then there exists a point $c \in (a, b)$ such that $f(c) = \alpha$.



The proof is beyond the scope of the course, but the result is intuitively clear: If f is above α

at one point, and below α at another, then somewhere in between, we must have $f(x) = \alpha$, as long as f is continuous (i.e., "one line").

Example 2.10.4 Use the IVT to show that the function

$$f(x) = x^5 - 2x^3 - 2$$

has a root between 0 and 2.

Solution:

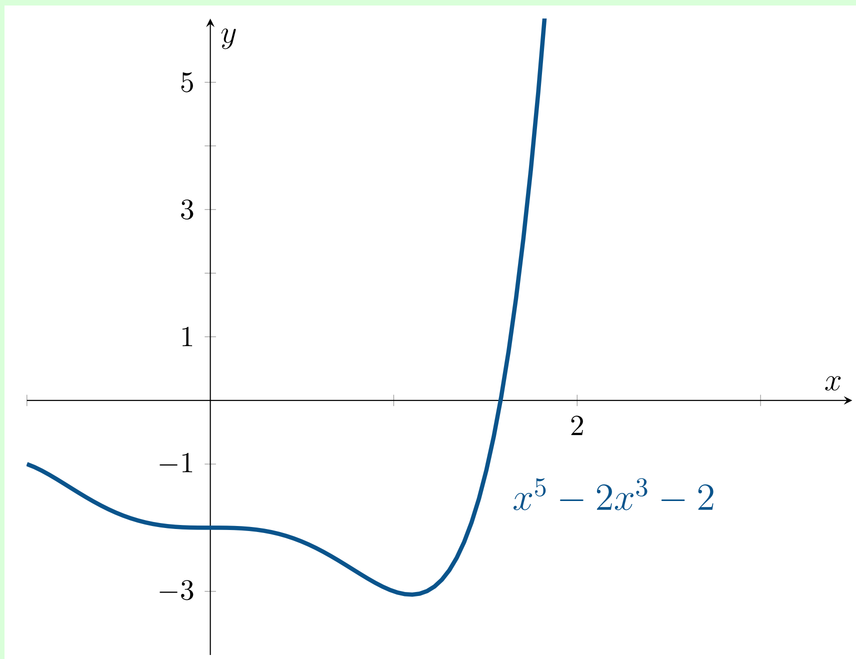
Since f is a polynomial, we know that f is continuous on $\mathbb{R}$, in particular, it is continuous on $[0, 2]$. Next,

$$f(0) = -2 < 0$$

and

$$f(2) = 14 > 0.$$

Hence, by the IVT, there exists a point $c \in (0, 2)$ such that $f(c) = 0$. The graph of f is depicted below.



Example 2.10.5 Show that there exists a point $c \in (0, 1)$ such that

$$\cos(c) = c.$$

Solution:

To prove the statement, we consider the helper function $f(x) = \cos(x) - x$. Note that $\cos(c) = c$ if and only if $\cos(c) - c = 0$. Hence, we prove that $f(c) = 0$ for some $c \in (0, 1)$.

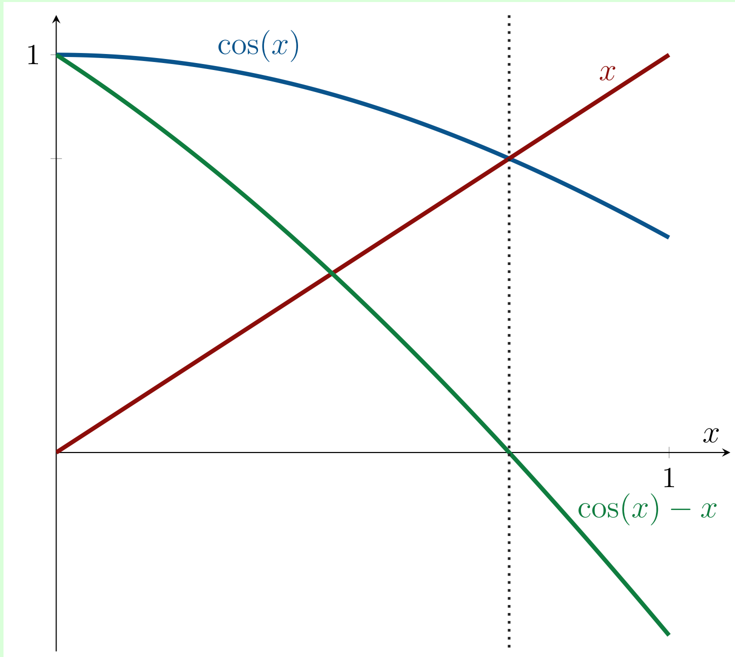
First, note that the function f is continuous as a difference of the continuous functions $\cos(x)$ and x. Furthermore,

$$f(0) = \cos(0) - 0 = 1 > 0$$

and

$$f(1) = \cos(1) - 1 < 0$$

because $\cos(1) \approx 0.54 < 1$. Thus, by the IVT, there is a $c \in (0, 1)$ so that $f(c) = 0$, or equivalently, such that $\cos(c) = c$.



A limitation of the Intermediate Value Theorem is that it only guarantees the existence of a c such that $f(c) = \alpha$; the IVT does not give us any indication of what c is, or if c is unique. However, the IVT can be used to approximate roots of continuous functions.

2.10.1 Approximating Solutions to Equations

We will start by considering polynomials. Our goal is to determine the roots of polynomials of different degrees.

If $p(x)$ is a polynomial of degree 1, how can we solve the equation $p(x) = 0$?
This is easily done! Since $p(x) = ax + b$, we can find $p(x) = 0$ if and only if $x = -b/a$.

Next, if $p(x)$ is a polynomial of degree 2, then we can solve $p(x) = 0$ using the quadratic formula, provided a solution exists. Indeed, let $p(x) = ax^2 + bx + c$. Then $p(x) = 0$ for

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

Note that if $b^2 - 4ac < 0$, then $p(x) = 0$ does not have a solution; if $b^2 - 4ac = 0$, then $p(x) = 0$ has one unique solution and if $b^2 - 4ac > 0$, then $p(x) = 0$ has two solutions.

Next, if $p(x)$ is a polynomial of degrees 3 or 4, then there are (quite ugly) formulas, just as was the case for the quadratic polynomial.

If $p(x)$ is a polynomial of degree 5 or higher, then no general formula exists. But we can still use the IVT to approximate roots!

Example 2.10.6 Let again

$$p(x) = x^5 - 2x^3 - 2.$$

We have shown in Example 2.10.4 that $p(x) = 0$ has a root in $(0, 2)$. But can we narrow down the location of a root further?

Let us investigate the function value at the midpoint of the interval $(0, 2)$, i.e., at $x = \frac{2+0}{2} = 1$. We get

$$p(1) = 1^5 - 2(1)^3 - 2 = -3 < 0$$

while $p(2) > 0$. Hence, by IVT, we know a root is somewhere in the interval $(1, 2)$.

Next, we consider the midpoint of the new interval $(1, 2)$, given by $x = \frac{2+1}{2} = \frac{3}{2}$. We get $p\left(\frac{3}{2}\right) = -\frac{37}{32} < 0$. Since $p(2) > 0$, we now know that a root is in the interval $\left(\frac{3}{2}, 1\right)$.

Next, the new midpoint of $\left(\frac{3}{2}, 1\right)$ is $x = \frac{7}{4}$ with $p(\frac{7}{4}) \approx 3.694 > 0$. As such, a root is in the interval $\left(\frac{3}{2}, \frac{7}{4}\right)$.

We could keep going like this, or even better, use a computer.

This method used in Example 2.10.6, called Bisection Method, is great because each additional step cuts the potential error in half and only requires that the function considered be continuous. Furthermore, since

$$\frac{1}{2^4} = \frac{1}{16} < \frac{1}{10},$$

every four iterations gives us another decimal place of accuracy. Similarly, since

$$\frac{1}{2^{10}} < \frac{1}{1000}$$

every 10 iterations gives as 3 decimal places in accuracy.

Of course, we can use this method on continuous functions beyond polynomials too! Let us generalize the method used in Example 2.10.6.

Method

Bisection Method

Suppose we want to approximate c such that $F(c) = 0$ with an error of at most $\varepsilon > 0$ where $F(x)$ is a continuous function.

- Step 1: Find two points a_0 and b_0 with $a_0 < b_0$ such that $F(a_0)$ and $F(b_0)$ have different signs. Then, since $F(x)$ is continuous, the IVT guarantees the existence of a point $c \in (a_0, b_0)$ with $F(c) = 0$.
- Step 2: Find the midpoint of (a_0, b_0), given by $d = \frac{a_0+b_0}{2}$, and compute $F(d)$.
- Step 3: If $F(a_0)$ and $F(d)$ have the same sign, let $a_1 = d$ and $b_1 = b_0$. Otherwise, let $a_1 = a_0$ and $b_1 = d$. In either case, we get a new interval $[a_1, b_1]$ that contains a solution, and the size of the new interval is $\frac{1}{2}(b_0 - a_0)$, which is half the size of the old interval.

- Step 4: Repeat Steps 2 and 3 to get a sequence of intervals $[a_1, b_1], [a_2, b_2], \ldots$ each containing a solution, with the kth interval having length $\frac{1}{2^k}(b_0 - a_0)$. Stop when $\frac{1}{2^{k+1}}(b_0 - a_0) < \varepsilon$. Then, if $d = \frac{a_k + b_k}{2}$, we know that the maximum distance from d to a solution c is less than ε, i.e., $|d - c| < \varepsilon$.

REMARK

If we are trying to approximate a solution to $f(x) = g(x)$, we can simply define $F(x) = f(x) - g(x)$ and our problem reverts to $F(x) = 0$. We can then apply the Bisection Method as shown above.

Example 2.10.7 Find a root of $f(x) = 3x^4 + 15x^3 - 125x - 1500$ in $[4, 5]$, with error at most 0.0025.

Solution:

The function f is continuous. Since $f(4) = -272 < 0$ and $f(5) = 1625 > 0$, the IVT implies that there is a root in $[a_0, b_0] = [4, 5]$.

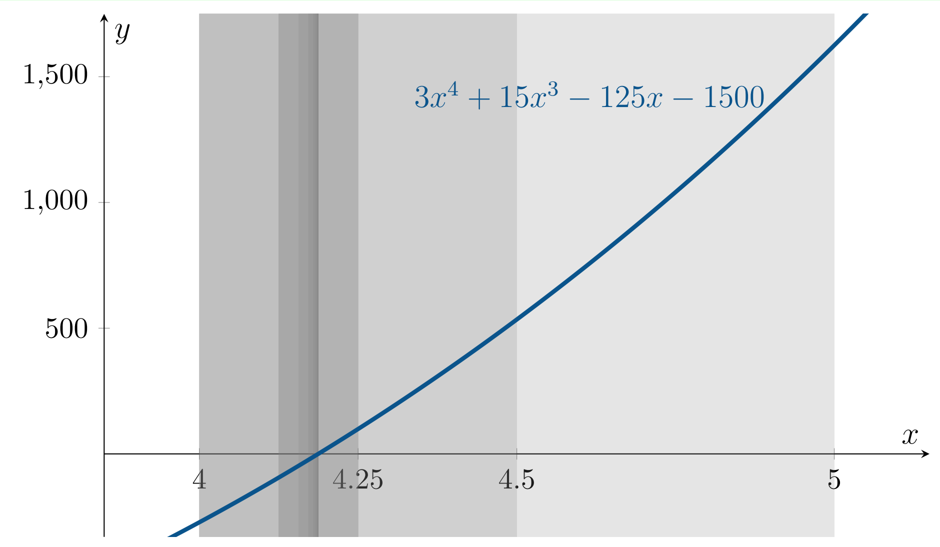
The midpoint is $d = (5 + 4)/2 = 4.5$ with $f(d) \approx 534.5625 > 0$. The next interval is thus $[a_1, b_1] = [4, 4.5]$.

The new midpoint is $d = 4.25$ with $f(d) \approx 98.9961 > 0$, so the next interval is $[a_2, b_2] = [4, 4.25]$. The results are summarized in the table below.

n	a	b	$f(a)$	$f(b)$	d	$f(d)$	Update
0	4	5	-272	1625	4.5	534.5625	$b = d$
1	4	4.5	-272	534.5625	4.25	98.9961	$b = d$
2	4	4.25	-272	98.9961	4.125	-94.1887	$a = d$
3	4.125	4.25	-94.1887	98.9961	4.1875	0.4346	$b = d$
4	4.125	4.1875	-94.1887	0.4346	4.15625	-47.3634	$a = d$
5	4.15625	4.1875	-47.3633	0.4346	4.171875	-23.5867	$a = d$
6	4.171875	4.1875	-23.5867	0.4346	4.1796875	-11.6067	$a = d$
7	4.1796875	4.1875	-11.6067	0.4346	4.18359375	-5.5937	$a = d$
8	4.18359375	4.1875	-5.5937	0.4346	4.185546875	-2.5815	$a = d$
9	4.185546875	4.1875	-2.5815	0.4346			

After 9 iterations, we end up with the interval $[4.1855, 4.1875]$ containing the root of f, which has length $|4.1875 - 4.1855| = 0.002 < 0.0025$, so we stop.

The graph below shows the function $f(x)$ along with the intervals $[a_0, b_0]$, $[a_1, b_1], \ldots$ containing the root as shaded areas. Note how the length of these intervals is cut in half after each iteration.



A LOOK AHEAD

The true solution in Example 2.10.7 is $x \approx 4.18722$. We found the interval $[4.1855, 4, 1875]$ after 9 iterations. In Section 3.6, we will learn a more efficient method to find solutions of $f(x) = 0$, called Newton's Method. You will later see that with Newton's Method, we will get an approximate value of 4.18722 after only 5 iterations!

Section 2.10 Problems

2.10.1. Consider the function $f(x) = x^3 - 4x - 1$.

- (a) Prove that f has at least one root in the interval $(0, 4)$.
- (b) Find an interval of length $\frac{1}{2}$ that contains a root of f.

2.10.2. Prove that $f(x) = x^3 - 12x + 10$ has at least two roots in $(0, 3)$.

2.10.3. Prove that there exists a real number c such that $2^c = c^5$.

2.10.4. Prove that the equation

$$x \sin(x) = 1$$

has infinitely many solutions.

2.10.5. Consider the function $f(x) = \dfrac{x^2 + x + 3}{x^3 + 1}$.

- (a) Verify that $f(-2) < 0$ and $f(1) > 0$.
- (b) Show that there are no real solutions to the equation $f(x) = 0$.
- (c) Do your findings from (a) and (b) contradict the Intermediate Value Theorem? Why or why not?

Chapter 3

Derivatives

Having built a solid foundation with our understanding of limits and continuity, we are ready for the next big step in our calculus journey. It's now time to turn to the namesake of differential calculus: the derivative.

The reader is strongly encouraged to spend the time in the early sections of this chapter ensuring they understand exactly what a derivative *is* and to hold onto that knowledge throughout the more computational aspects of this chapter and the next.

By the end of this chapter, you will be able to

- define the derivative and explain the mathematics behind it;
- calculate the derivative of a variety of functions;
- conduct approximations and find roots via the derivative.

3.1 Average and Instantaneous Velocity

At the heart of the derivative is the concept of rates of change. To ground ourselves in some sense of reality as we begin, we conduct an investigation on a car ride.

Note that when you look at a speed limit sign, you see that the posted limit is displayed in units of km/h, or kilometres *per* hour. Here, we see the fraction line and the word *per* indicate a *rate*. We also note that the units tell us we are looking at the rate of *change* of kilometres with respect to each hour.

Example 3.1.1 Suppose you are driving to campus from out-of-town to attend your favourite lecture (MATH 137, of course). Data for your journey is shown in the table below:

Time t in min	0	30	60	90	120	150	180
Position s in km	0	15	25	50	90	120	135

(a) What was your average velocity over the entire duration of the trip?

(b) What was your average velocity over the first half hour?

(c) What was your average velocity between 60 to 120 minutes into your journey?

Solution: It is convention to use km/h for driving speeds, so we ensure we use these units to answer these problems.

Further, we understand that in general (and from analyzing the units of our rate of change):

$$\text{average velocity} = \frac{\text{change in position}}{\text{change in time}} = \frac{\Delta s}{\Delta t}.$$

(a) We need to take into account our initial and final data points here, being wary of units. This gives us that the average velocity $= \dfrac{135 - 0 \text{ km}}{3 - 0 \text{ h}} = 45$ km/h.

(b) Adjusting to the relevant data points and correct units, we find that the average velocity $= \dfrac{15 - 0 \text{ km}}{0.5 - 0 \text{ h}} = 30$ km/h.

(c) Adjusting to the relevant data points and correct units, we find that the average velocity $= \dfrac{90 - 25 \text{ km}}{2 - 1 \text{ h}} = 65$ km/h.

We see that we can generalize our example to arrive at a formula for average velocity.

Definition 3.1.2
Average Velocity

The **average velocity**, v_{avg}, between times t_0 and t_1 is given by

$$v_{avg} = \frac{s(t_1) - s(t_0)}{t_1 - t_0}$$

where here $s(t)$ is the position at time t.

Now, it is beneficial to gain a graphical understanding of average velocity. Consider drawing the **secant line** (a line joining two distinct points of a given function) between the two data points used in part (c) of our example.

From Figure 3.1.1, we can see that the rate of change we calculated, $\dfrac{\Delta s}{\Delta t}$, is the slope of the secant line between the relevant data points.

REMARK

The slope of the secant line represents the average rate of change.

Now, imagine you are pulled over by a police officer at some point on your journey to campus, and they inform you that you were above the sign-posted limit of 60 km/h. Telling the officer that your average velocity up to that point was below the speed limit won't get you out of a ticket!

In this situation, as in many others in reality, we care about an *instantaneous* rate of change, rather than an average. In this unfortunate circumstance, your speedometer, which reads

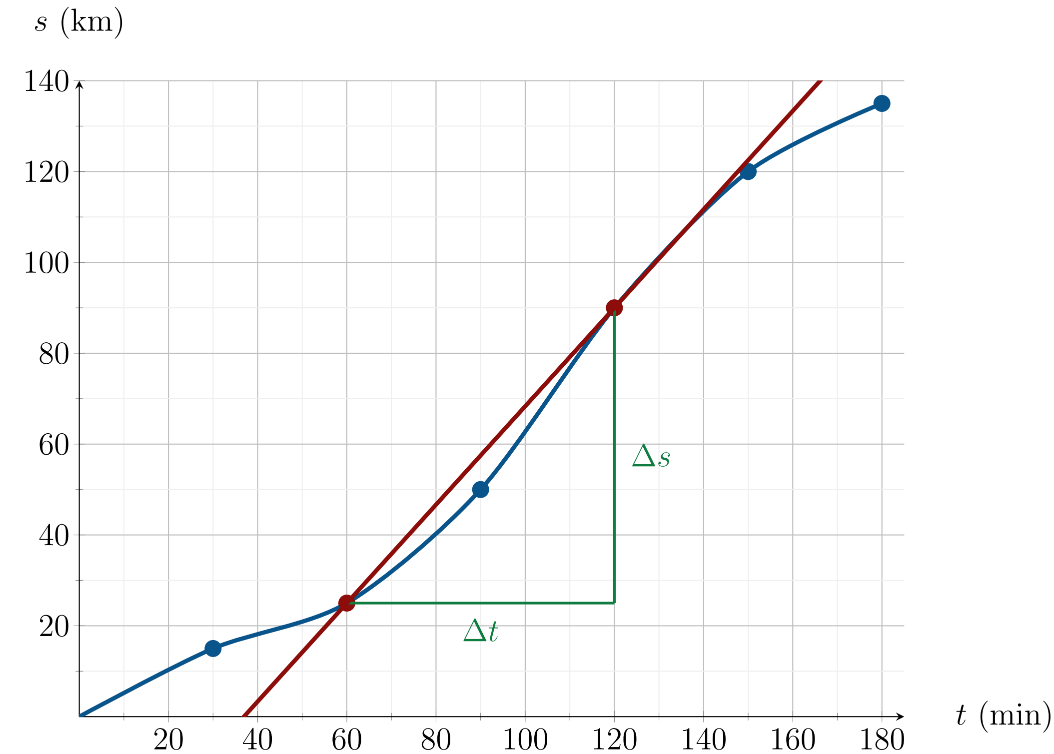


Figure 3.1.1: The car journey and the slope of the secant line in relation to finding the average velocity in Example 3.1.1 (c).

out your instantaneous velocity, indicated that you were speeding at that given instant in time.

But how do we find our velocity at an instant in time? To calculate our average velocity, or the slope of our secant line if we're thinking graphically, we need *two* data points. In an instant, we only have *one* data point.

The key idea is as follows: we will find the average velocity, or slope of the secant line, between our instant t_0 and a nearby time t and use this as an approximation of the instantaneous velocity at t_0.

We repeatedly do this, bringing t closer towards t_0, and we expect that our approximation will get closer to the true instantaneous velocity. Note that we can approach t_0 from the right or left, as demonstrated in Figure 3.1.2.

This idea of having t approach t_0 and seeing where the average rate of change tends towards should be setting off alarm bells in the reader's mind. This is exactly the concept of a *limit*.

We can formalize this as follows:

Definition 3.1.3

Instantaneous Velocity: Version 1

The **instantaneous velocity**, v_{inst} or $v(t_0)$, at time t_0 is given by

$$v_{inst} = v(t_0) = \lim_{t \to t_0} \frac{s(t) - s(t_0)}{t - t_0}$$

where s is the position function.

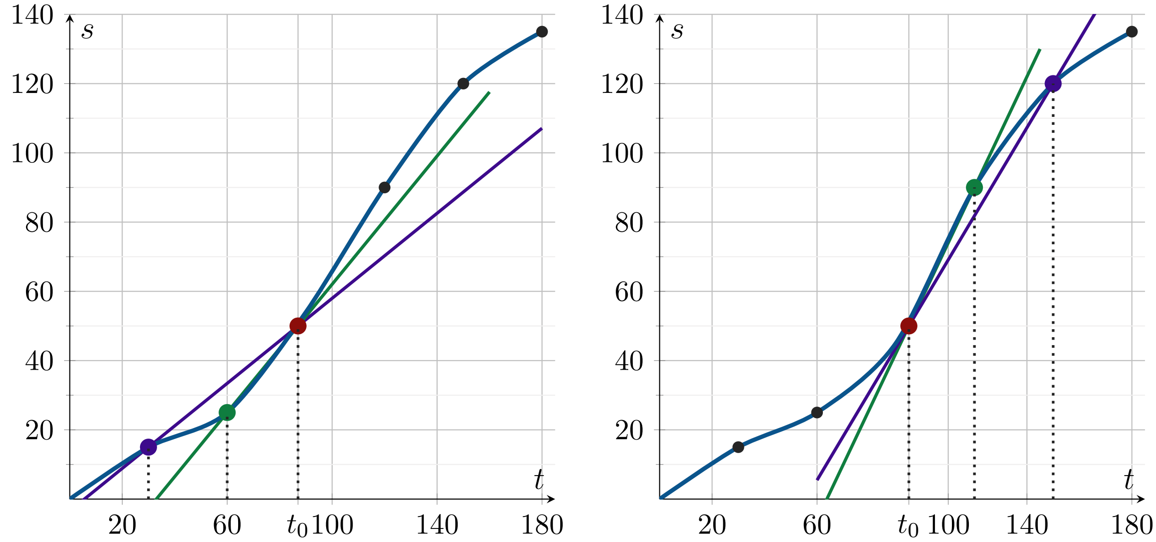


Figure 3.1.2: In the panel on the left, we bring t closer to t_0 from the left, and the slope of the secant line better represents the instantaneous velocity at t_0. In the panel on the right, we bring t closer to t_0 from the right, and the slope of the secant line better represents the instantaneous velocity at t_0.

Now, as mathematicians, we value the ability to look at the same problem from different perspectives. In particular, we will soon see that it can be helpful to reframe our limit definition for an instantaneous rate of change.

If we have that $t \neq t_0$, then we can write $t = t_0 + h$ ($h \neq 0$). Here, h is the offset that t has from t_0 (positive if to the right of t_0 or negative if to the left of t_0). In this frame of reference, when we take $t \to t_0$, this is equivalent to taking $h \to 0$.

We can thus write an equivalent limit definition:

Definition 3.1.4

Instantaneous Velocity: Version 2

The **instantaneous velocity**, v_{inst} or $v(t_0)$, at time t_0 is also given by

$$v_{inst} = v(t_0) = \lim_{h \to 0} \frac{s(t_0+h) - s(t_0)}{(t_0+h) - t_0} = \lim_{h \to 0} \frac{s(t_0+h) - s(t_0)}{h}$$

where s is the position function.

We again seek to have some graphical understanding of the situation. We note that as $t \to t_0$ (or $h \to 0$), our second point for our secant line gets infinitely closer to the first. We call the line that goes through these infinitely close points the **tangent line** - the line just touches the graph of the function at t_0. Now, this is a rather hand-wavy explanation of the tangent line, visualized in Figure 3.1.3, so you must eagerly await when we formalize its definition in Section 3.2.

REMARK

The slope of the tangent line represents the instantaneous rate of change.

We can now determine instantaneous velocity when given a position function.

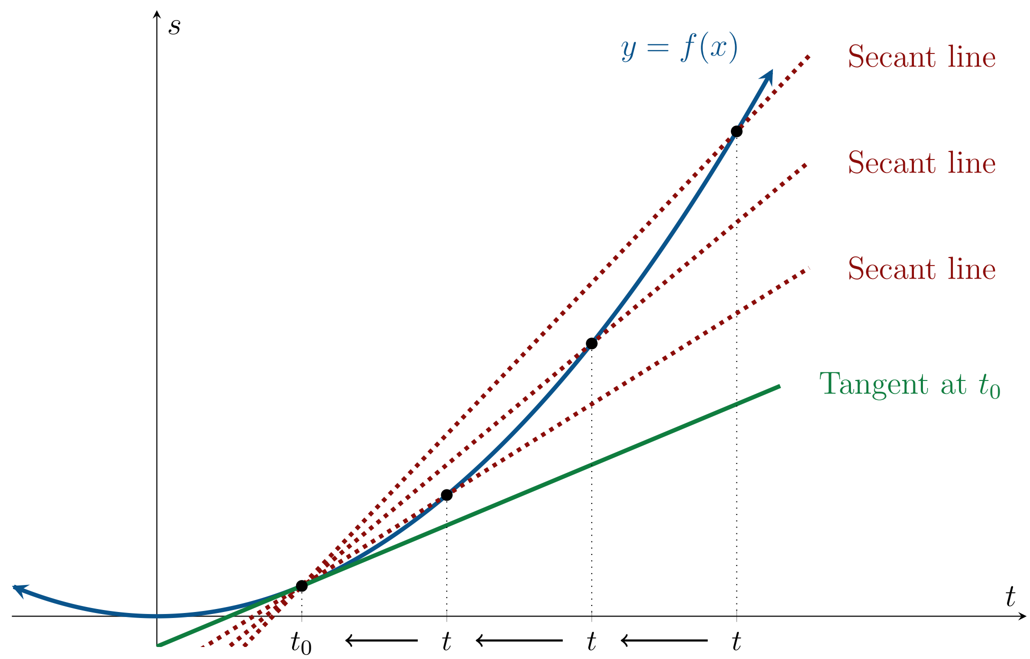


Figure 3.1.3: As we move t towards t_0 we can see that the secant line between t_0 and t better replicates the tangent line at t_0.

Example 3.1.5 Find the instantaneous velocity of $s(t) = t^2 + 3t$ at:

(a) $t = 2$;

(b) $t = t_0$ (that is, an arbitrary time.)

Solution:

(a) Being cautious through our use of brackets, we have

$$\begin{aligned} v(2) &= \lim_{h \to 0} \frac{s(2+h) - s(2)}{h} \\ &= \lim_{h \to 0} \frac{(2+h)^2 + 3(2+h) - \left(2^2 + 3(2)\right)}{h} \\ &= \lim_{h \to 0} \frac{4 + 4h + h^2 + 6 + 3h - 4 - 6}{h} \\ &= \lim_{h \to 0} \frac{7h + h^2}{h} = \lim_{h \to 0} 7 + h = 7. \end{aligned}$$

Thus, the instantaneous velocity at $t = 2$ is 7.

(b) The process here is nearly identical to part (a), giving

$$\begin{aligned} v(t_0) &= \lim_{h \to 0} \frac{s(t_0+h) - s(t_0)}{h} \\ &= \lim_{h \to 0} \frac{(t_0+h)^2 + 3\left(t_0+h\right) - \left(t_0^2 + 3t_0\right)}{h} \end{aligned}$$

$$
\begin{aligned}
&= \lim_{h\to 0} \frac{t_0^2 + 2t_0 h + h^2 + 3t_0 + 3h - t_0^2 - 3t_0}{h} \\
&= \lim_{h\to 0} \frac{(2t_0+3)\,h + h^2}{h} = \lim_{h\to 0} 2t_0 + 3 + h = 2t_0 + 3.
\end{aligned}
$$

Thus, the instantaneous velocity at $t = t_0$ is $2t_0 + 3$.

The advantage to this formula is that it is generic for the given position function. All we need to do is plug in any time t_0 and we immediately find the instantaneous velocity. For example, plugging in $t_0 = 2$ returns 7 as in part (a).

Section 3.1 Problems

3.1.1. For the displacement function, $s(t) = -9.8t^2 + 30t + 12$, determine the instantaneous velocity at $t = 1$.

3.1.2. The height of a tree is measured every year in the autumn. The data is shown below.

Year	0	1	2	3	4	5	6	7	8	9	10
Height (m)	9.91	11.07	12.23	13.28	15.08	16.53	18.40	20.16	22.30	24.62	27.51

(a) Estimate the rate of change over each year.

(b) Tabulate the estimated rate of change divided by the average height of the tree over each year.

(c) Propose an equation linking the instantaneous rate of change $\dfrac{dh}{dt}$ to the height of the tree $h(t)$.

3.1.3. Suppose the speed limit on a road is 50 km/h – when you approach a set of traffic lights, how long should the yellow light stay on to give the motorists sufficient time to stop? What about if the speed limit is 70 km/h? In this question, you will analyze a simple model for the the physics of traffic stops.

(a) When the light turns yellow, a motorist has two choices: either slam on the breaks and stop (call the time it takes to stop T_{stop}), or continue through the intersection (call the time it takes to clear the intersection T_{run}). There is a reaction time associated with the decision (call this decision time T_{reac}). The length of time for the yellow light (T_{yellow}) should be at least as long as,

$$T_{\text{yellow}} = T_{\text{reac}} + \max(T_{\text{stop}}, T_{\text{run}}), \tag{3.1}$$

to ensure that the intersection is clear when the opposing light turns green. Provide an explanation for this equation.

(b) Assuming a typical car length of $L = 5$ m, and a typical intersection width to be $I = 10$ m, a car travelling at v_0 km/h will take $T_{\text{run}} = (I + L)/v_0$ to clear the intersection. On the other hand, assuming uniform braking, a car will take $T_{\text{stop}} = v_0/(2fg)$ to come to a stop, where $g = 9.81$ m/s^2 is the acceleration of gravity and f is the coefficient of friction between the tires and pavement. For a dry, flat stretch of road, a typical value of the friction coefficient is $f = 0.2$.

When the speed limit v_0 is very slow, $T_{\text{run}} > T_{\text{stop}}$; whereas for faster speed limits v_0 the opposite is true, $T_{\text{run}} < T_{\text{stop}}$. Using the typical values above, at what speed limit v_0^{crit} does $T_{\text{run}} = T_{\text{stop}}$?

(c) Most traffic lights are on roads where the speed limit exceeds the critical speed found in part b, $v_0 > v_0^{\text{crit}}$, so that $\max(T_{\text{stop}}, T_{\text{run}}) = T_{\text{stop}}$. Ontario cities can lawfully time the yellow light at $T_{\text{yellow}} = 3.7$ seconds on flat roads with a speed limit of 60 km/h. Is this consistent with Eq. 3.1 above? Assume that $T_{\text{reac}} = 1$ second.

(d) For roads on a hill, the situation is more complicated: the effective weight of the car is reduced (leading to a reduction in the friction between the tires and road), and gravity is working against braking on a down slope (or acting with braking on an up-slope). The US Institute of Transport Engineers (USITE) recommends the following estimator for T_{yellow},

$$T_{\text{yellow}} = T_{\text{reac}} + \frac{v_0}{2g(f+G)},$$

where G is the slope of the road in the direction of approach. Calculate the recommended stopping times for three different speed limits $v_0 = 50$ km/h, 70 km/h and 100 km/h under each of three different road conditions: $G = 0$ (flat road), $G = 6/100$ (recommended maximum slope for a highway, approaching uphill), and $G = -6/100$ (recommended maximum slope for a highway, approaching downhill). What happens if you try to compute the stopping time traveling 50 km/h downhill along North America's steepest road ($G = -37/100$; Canton Ave. in Pittsburgh, PA)?

3.2 Definition of the Derivative

We can generalize our work with position and velocity to other functions. We now consider instead a generic function $f(x)$.

In this context, our average velocity is comparable to:

Definition 3.2.1
Average Rate of Change

The **average rate of change**, f_{av}, of a function $f(x)$ between times $x = a$ and $x = b$ is given by

$$\frac{f(b) - f(a)}{b - a}$$

Our instantaneous velocity corresponds to the term we've been been waiting to use:

Definition 3.2.2
Instantaneous Rate of Change or The Derivative of $f(x)$ at $x = a$

The **instantaneous rate of change** or **derivative of $f(x)$ at $x = a$**, $f'(a)$,is given by

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

In particular, if this limit exists, we say that $f(x)$ is **differentiable** at $x = a$.

REMARKS

- The quantity $\dfrac{f(a + h) - f(a)}{h}$ is known as the Newton quotient or the difference quotient.
- It is crucial to understand that the derivative is just a *limit*. Whether or not a function is differentiable at a point boils down to whether or not this limit exists.
- In the context of our previous discussion on position and velocity, we note that velocity is the derivative of position.
- In Example 2.2.9, we sneakily had you compute your very first derivative - in that case for $f(x) = x^2$ at $a = 2$.

We are now able to formally define the aforementioned tangent line:

Definition 3.2.3
Tangent Line, Point of Tangency

If $f(x)$ is differentiable at $x = a$, then the **tangent line** to $f(x)$ at $x = a$ is the line passing through $(a, f(a))$ with slope $f'(a)$.

The equation of the tangent line is thus

$$y = f'(a)(x - a) + f(a)$$

and the point $(a, f(a))$ is known as the **point of tangency**.

REMARKS

- The equation of the tangent line comes from the point-slope form of the equation of a line with slope m through the coordinate point (x_1, y_1):

$$y - y_1 = m(x - x_1) \implies y = m(x - x_1) + y_1$$

  In calculus, we tend to deal with points of tangency. This form is more useful to us than the slope-intercept form of the equation of a line that is likely burned into your core memories, $y = mx + b$, which demands we always find the y-intercept.

- It is important to realize that we cannot *define* the derivative as the slope of the tangent line. If $f(x)$ is not differentiable at $x = a$ the formal definition of a tangent line does not even exist - as the tangent line is defined based on the derivative! However, it is fair to say that when the formal definition of the tangent line holds, its slope is the derivative.

Example 3.2.4 Find the equation of the tangent line to $f(x) = x^2 + x + 1$ at $x = 3$.

Solution: First, we should compute $f'(3)$:

$$\begin{aligned} f'(3) &= \lim_{h \to 0} \frac{f(3+h) - f(3)}{h} \\ &= \lim_{h \to 0} \frac{(3+h)^2 + (3+h) + 1 - \big(3^2 + 3 + 1\big)}{h} \\ &= \lim_{h \to 0} \frac{9 + 6h + h^2 + 3 + h + 1 - 9 - 3 - 1}{h} \\ &= \lim_{h \to 0} \frac{7h + h^2}{h} = \lim_{h \to 0} 7 + h = 7. \end{aligned}$$

We still need to find the point of tangency to complete the equation. The x-coordinate of 3 is given to us, and to find the y-coordinate we compute $f(3) = 3^2 + 3 + 1 = 13$.

So the equation of the tangent line is $y = 7(x - 3) + 13$.

(If you dearly miss slope-intercept form, you could expand this to find $y = 7x - 8$.)

A LOOK AHEAD

Finding the equation of a tangent line may seem rather pointless right now, but we will see in Section 3.5 how we can use tangent lines to approximate functions. We further extend this idea in Section 3.6 when we discuss another root-finding algorithm.

Having established the definition of the derivative, we now look to make a key connection with continuity. This is a result so nice, we prove it twice.

Theorem 3.2.5 **(Differentiability Implies Continuity)**

If $f(x)$ is differentiable at $x = a$ then it is continuous at $x = a$.

Equivalently, by the contrapositive, if $f(x)$ is discontinuous at $x = a$ then it is not differentiable at $x = a$.

Proof 1:

Let $f(x)$ be differentiable at $x = a$.

Then, we know that $f'(a) = \lim\limits_{x \to a} \dfrac{f(x) - f(a)}{x - a}$ exists.

Now, since we know the limit of this quotient exists, and that $\lim\limits_{x \to a} x - a = 0$, by a result from Theorem 2.2.5 we know that $\lim\limits_{x \to a} f(x) - f(a) = 0$.

By the formal definition of a limit, this means that for all $\epsilon > 0$ there exists a $\delta > 0$ such that for $0 < |x - a| < \delta$ we have that $|(f(x) - f(a)) - 0| < \epsilon$.

That is, for all $\epsilon > 0$ there exists a $\delta > 0$ such that for $0 < |x - a| < \delta$ we have that $|f(x) - f(a)| < \epsilon$.

The previous statement is the formal definition of the limit

$$\lim_{x \to a} f(x) = f(a)$$

which is the definition of continuity for $f(x)$ at $x = a$. $\square$

Proof 2:

Let $f(x)$ be differentiable at $x = a$.

Then, we know that $f'(a) = \lim\limits_{x \to a} \dfrac{f(x) - f(a)}{x - a}$ exists.

Now, examine

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

Now, being a little sneaky, we note

$$\begin{aligned} \lim_{x \to a} f(x) &= \lim_{x \to a}(f(x) - f(a) + f(a)) \\ &= \lim_{x \to a}(f(x) - f(a)) + \lim_{x \to a} f(a) \\ &= 0 + f(a) \\ \lim_{x \to a} f(x) &= f(a) \end{aligned}$$

which is the definition of continuity for $f(x)$ at $x = a$. $\square$

Now, a natural question might be to ask if continuity implies differentiability. The astute reader will notice there is no theorem box stating this result and correctly conclude that the answer to this question is a resounding no!

We show this via the use of a common counterexample.

Example 3.2.6 Show that $f(x) = |x|$ is continuous at $x = 0$ but not differentiable there.

Solution: We first examine $\lim_{x\to 0} f(x)$ by noting that $\lim_{x\to 0^+} f(x) = \lim_{x\to 0^+} x = 0$ and $\lim_{x\to 0^-} f(x) = \lim_{x\to 0^-} -x = 0$. Thus, we have that

$$\lim_{x\to 0^+} f(x) = \lim_{x\to 0^-} f(x) = 0 \implies \lim_{x\to 0} f(x) = 0.$$

Now, we note that $f(0) = |0| = 0$.

So, we have shown that

$$\lim_{x\to 0} f(x) = 0 = f(0)$$

and thus that $f(x)$ is continuous at $x = 0$.

Now, we examine $f'(0)$:

$$\begin{aligned} f'(0) &= \lim_{h\to 0} \frac{|0+h| - |0|}{h} \\ &= \lim_{h\to 0} \frac{|h|}{h}. \end{aligned}$$

Here, we note that $\lim_{h\to 0^+} \frac{|h|}{h} = \lim_{h\to 0^+} \frac{h}{h} = 1$ but $\lim_{h\to 0^-} \frac{|h|}{h} = \lim_{h\to 0^-} \frac{-h}{h} = -1$.

Thus, $f'(0) = \lim_{h\to 0} \frac{|h|}{h}$ DNE, and $f(x)$ is not differentiable at $x = 0$.

REMARKS

- Here, the derivative ceases to exist because approaching from either side of zero we are on different branches of the absolute value function, which have different slopes. The *right-sided derivative* and *left-sided derivative* exist but do not match. When this occurs, we have a **corner**.
- Another common occurrence for a function to be continuous but not differentiable is when the limit definition of the derivative evaluates to $\pm\infty$. Consider for example $f(x) = x^{1/3}$ at $x = 0$. This is typically called the situation of a **vertical tangent line**.
- Yet another situation for a continuous but not differentiable function is that of a **cusp**, which occurs for example with $f(x) = x^{2/3}$ at $x = 0$. A cusp occurs at $x = a$ where the function is continuous when the left-sided derivative and right-sided derivative are $\pm\infty$ and do not match.

Now, we understand that Definition 3.2.2 gives us the derivative of a function *at a point*. However, what is really useful to us in practice is to know the derivative of a function at *every* point in its domain. Ideally, we would like to do this in some way that does not involve re-computing $f'(a)$ for every single a in the domain.

You may recall that in Example 3.1.5 we managed to find a formula for the instantaneous velocity at an arbitrary time given a position function. We can employ the same strategy we applied in that scenario, in general:

Definition 3.2.7

The Derivative Function

We say that $f(x)$ is **differentiable** on an interval I if $f'(a)$ exists for all $a \in I$. On such an interval, the **derivative function**, $f'(x)$, is defined as

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

REMARKS

- The only difference in the definition above and Definition 3.2.2 is that we have replaced each occurrence of a (a specific x-value) with x (an arbitrary x-value).
- The derivative function is the derivative of $f(x)$ at each $x \in I$.
- Do not let it escape your notice that evaluating the limit in the definition above leads you to a *function*, not a *value* as you are used to from the earlier chapters of this course.

Example 3.2.8

Given $f(x) = x^3 + 1$, find $f'(x)$. Then, determine where $f(x)$ has a horizontal tangent line.

Solution:

$$\begin{aligned}
f'(x) &= \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} \\
&= \lim_{h\to 0} \frac{(x+h)^3 + 1 - (x^3+1)}{h} \\
&= \lim_{h\to 0} \frac{x^3 + 3x^2h + 3xh^2 + h^3 + 1 - x^3 - 1}{h} \\
&= \lim_{h\to 0} \frac{3x^2h + 3xh^2 + h^3}{h} \\
f'(x) &= \lim_{h\to 0} 3x^2 + 3xh + h^2 = 3x^2
\end{aligned}$$

We note that a horizontal tangent line has a slope of zero. Thus, we seek to find where $f'(x) = 0$. We have

$$f'(x) = 0 \implies 3x^2 = 0 \implies x = 0.$$

So, $f(x)$ has a horizontal tangent line at $x = 0$.

A LOOK AHEAD

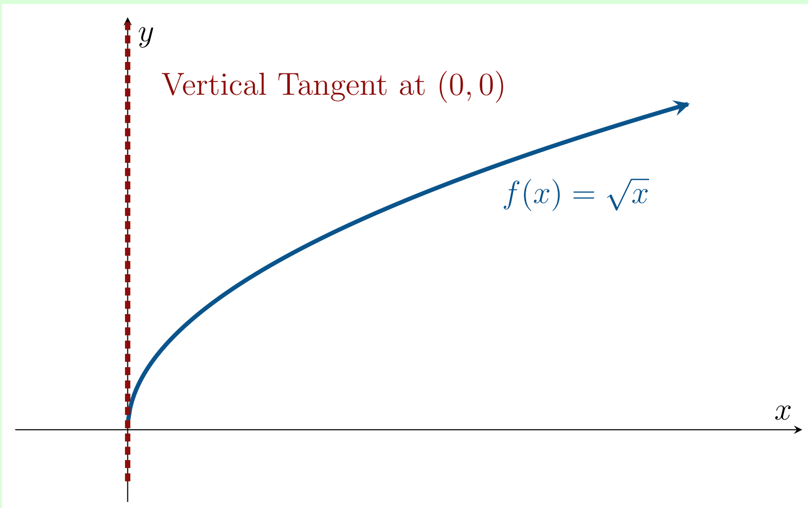
The idea of searching for the locations of horizontal tangent lines will become one of our central tools when we start looking for extrema in Section 4.2.

Example 3.2.9 Given $f(x) = \sqrt{x}$, find $f'(x)$. Then, determine where $f(x)$ is not differentiable.

Solution:

$$\begin{aligned} f'(x) &= \lim_{h\to 0} \frac{\sqrt{x+h}-\sqrt{x}}{h} \cdot \frac{\sqrt{x+h}+\sqrt{x}}{\sqrt{x+h}+\sqrt{x}} \\ &= \lim_{h\to 0} \frac{x+h-x}{h(\sqrt{x+h}+\sqrt{x})} \\ &= \lim_{h\to 0} \frac{h}{h(\sqrt{x+h}+\sqrt{x})} \\ f'(x) &= \lim_{h\to 0} \frac{1}{\sqrt{x+h}+\sqrt{x}} = \frac{1}{2\sqrt{x}} \end{aligned}$$

We see that $f'(x)$ is not defined when $x = 0$, so $f(x)$ is not differentiable there. Graphically, we can see the tangent line to $f(x)$ at $x = 0$ would be vertical. Vertical lines have undefined slope, which matches with our understanding of the derivative being undefined here.



You may also notice that $f'(x)$ is not defined when $x < 0$, but this is outside of the domain of the original function, so the concept of differentiability does not apply there.

It is worthwile to now to make a note on notation. The *primed* notation (f') we have been using so far to denote derivatives comes from Joseph-Louis Lagrange, whose notation closely resembles the original *dot* notation ($\dot{f}$) of Sir Isaac Newton, one of the fathers of calculus. This is the Newton of the fabled falling apple. Given that the reader is unlikely to be naturally inspired by another falling fruit, your study of calculus must continue.

As much of Newton's work in mathematics was grounded in physics, his notation was generally based on functions which were dependent on time, t. As such, his notation has fallen out of practice in most mathematics texts (mathematicians naturally refuse to be bound by such foolish concepts as time) and is now more commonly found in use in the fields of engineering and physics.

While the ease of writing the notation of Lagrange and Newton is appealing due to the mathematician's tendency to be intelligently lazy, there is another common notation of which we must be aware.

The history of calculus, like any good story, was fraught with tension. At the same time as Newton was building the foundations of calculus in Britain in the 17th century, Gottfried Wilhelm Leibniz of Germany was also working on the invention of the subject. By 1711, the Leibniz-Newton controversy was in full force, with each mathematician accusing the other of plagiarism. This controversy outlived Leibniz, who died in 1716. Today, both men are given credit for *independently* inventing the foundations of calculus at roughly the same time.

Given that Leibniz worked on his theories wholly independently of Newton, he had his own notation for derivatives. This notation is much more verbose than that of Newton (and Lagrange). While it means we must write more, it is more communicative of what it is we are doing.

Definition 3.2.10

Notations for the Derivative

We write the derivative function of $y = f(x)$ with the following notations:

$$f'(x) = \frac{df}{dx} = \frac{d}{dx}(f) = y' = \frac{dy}{dx} = \frac{d}{dx}(y)$$

Furthermore, the derivative of $y = f(x)$ at a point $x = a$ is written with the following notations:

$$f'(a) = \left.\frac{df}{dx}\right|_{x=a} = \left.\frac{d}{dx}(f)\right|_{x=a} = \left.\frac{dy}{dx}\right|_{x=a} = \left.\frac{d}{dx}(y)\right|_{x=a}$$

Here, the primed notation is defined to be **Lagrange notation** and the $\frac{d}{dx}$ notation is defined to be **Leibniz notation**.

REMARKS

- Leibniz notation does *not* indicate a fraction.
- $\frac{d}{dx}$ is called the *differential operator*. Leibniz notation is literally telling us to apply the operation of the derivative to the function $y = f(x)$ with respect to the variable x.
- The reader should feel comfortable with both Lagrange and Leibniz notation, as they will be used interchangeably throughout the rest of these notes.
- Leibniz notation will typically become your default notation in further courses when you will have to deal with multivariable functions.

A LOOK AHEAD

While it has already been mentioned that Leibniz notation is more informative than Lagrange notation, we will see just how this longer notation can be helpful when we talk about the Chain Rule in Section 3.4.

Before we proceed to find the derivatives of common functions, we recognize that there is nothing stopping us from taking the derivative of the derivative function. Or indeed, from taking the derivative of the derivative of the derivative function. And so forth.

Definition 3.2.11

The nth Derivative

If $f(x)$ is n-times differentiable, we denote the n**th derivative** as

$$f^{(n)}(x) = \frac{d^n}{dx^n}(f) = \frac{d}{dx}\left(f^{(n-1)}(x)\right)$$

REMARKS

- Be wary of the placement of n's in the Leibniz Notation for the nth derivative.
- It is typical that we denote the first, second, and third derivatives in Lagrange notation as f', f'', and f'''. For the fourth derivative and beyond, we typically write $f^{(4)}$, $f^{(5)}$, etc.
- In the context of position and velocity, we say that acceleration is the first derivative of velocity and the second derivative of position.

Example 3.2.12

If $f(x)$ is 96 times differentiable, then the 96th derivative of $f(x)$ is

$$f^{(96)}(x) = \frac{d^{96}}{dx^{96}}(f) = \frac{d}{dx}\left(f^{(95)}\right).$$

A LOOK AHEAD

While the 96th derivative is generally not of much use, we will see that both the first and *second* derivative provide us useful information about the shape of functions in Section 4.6.

Before moving on, it is worth having a little bit of fun. If we consider s to be position, then:

- s' is velocity
- s'' is acceleration
- s''' is jerk
- $s^{(4)}$ is snap (or jounce)
- $s^{(5)}$ is crackle
- $s^{(6)}$ is pop
- $s^{(7)}$ is lock
- $s^{(8)}$ is drop

Why do we have such strange (read: silly) names? It is likely because after around the fourth derivative of position, these quantities aren't as meaningful in a physical sense.

Section 3.2 Problems

3.2.1. Use the definition of the derivative $f'(x) = \lim_{h\to 0} \dfrac{f(x+h)-f(x)}{h}$ to determine the derivative of the following functions.

$$\text{i) } f(x) = x^2, \qquad \text{ii) } f(x) = \sqrt{x}, \qquad \text{iii) } f(x) = \ln x$$

Hint: For iii), recall that $\lim_{\alpha\to 0} (1+\alpha)^{1/\alpha} = e$, and try a change of variables $\alpha = \frac{h}{x}$.

3.2.2. For each of the following, determine the value of $f'(a)$ using the definition of the derivative.

(a) $f(x) = x^4$, $a = 2$

(b) $f(x) = \sqrt{x-2}$, $a = 9$

(c) $f(x) = \dfrac{3}{x^2+7}$, $a = -3$

(d) (**Challenge Problem!**) $f(x) = \dfrac{1}{\sqrt{x-3}+\sqrt{x+2}}$, $a = 7$

3.2.3. For $f(x) = \dfrac{x+1}{x-1}$, find $f'(x)$ using the limit definition.

3.2.4. If f is differentiable at x and $f(x) \neq 0$, use the definition of the derivative to prove that

$$\frac{d}{dx}\left(\frac{1}{f(x)}\right) = \frac{-f'(x)}{[f(x)]^2}.$$

3.2.5. If f is differentiable and $f(x) > 0$, use the definition of the derivative to prove that

$$\frac{d}{dx}\left(\sqrt{f(x)}\right) = \frac{f'(x)}{2\sqrt{f(x)}}.$$

3.3 Derivatives of Some Common Functions

In this section, we seek to find the derivatives of common functions by using the first-principles Definition 3.2.7.

Example 3.3.1

Find the derivative of the constant function $f(x) = c$ where $c \in \mathbb{R}$.

Solution: We begin by understanding that $f(x) = c$ is a function which, no matter the input, will always output the constant c. In particular, this means that $f(x+h) = c$. Using the definition of the derivative, we have

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

So, given $f(x) = c$, we have that $f'(x) = 0$.

REMARK

Given our understanding that fundamentally the derivative is the instantaneous rate of change at any arbitrary x-value, it was rather painful to proceed from the first principles in this example. A constant function is a horizontal line, which has a slope, or instantaneous rate of change, of zero everywhere. Nevertheless, it is good to see that calculus matches our intuition and is not broken.

Example 3.3.2

Find the derivative of the linear function $f(x) = mx + b$ where $m, b \in \mathbb{R}$.

Solution:

$$\begin{aligned} f'(x) &= \lim_{h\to 0} \frac{m(x+h) + b - (mx+b)}{h} \\ &= \lim_{h\to 0} \frac{mx + mh + b - mx - b}{h} \\ &= \lim_{h\to 0} \frac{mh}{h} \\ f'(x) &= \lim_{h\to 0} m = m \end{aligned}$$

So, given $f(x) = mx + b$, we have that $f'(x) = m$.

REMARK

Again, the calculus matches our intuition that the slope everywhere of a line with equation $f(x) = mx + b$ is m.

Example 3.3.3

Find the derivative of the quadratic function $f(x) = ax^2 + bx + c$ where $a, b, c \in \mathbb{R}$.

Solution:

$$\begin{aligned} f'(x) &= \lim_{h \to 0} \frac{a(x+h)^2 + b(x+h) + c - (ax^2 + bx + c)}{h} \\ &= \lim_{h \to 0} \frac{ax^2 + 2axh + ah^2 + bx + bh + c - ax^2 - bx - c}{h} \\ &= \lim_{h \to 0} \frac{h(2ax + ah + b)}{h} \\ f'(x) &= \lim_{h \to 0} 2ax + ah + b = 2ax + b \end{aligned}$$

So, given $f(x) = ax^2 + bx + c$, we have that $f'(x) = 2ax + b$.

Example 3.3.4

Find the derivative of the sine function $f(x) = \sin(x)$.

Solution:

$$\begin{aligned} f'(x) &= \lim_{h \to 0} \frac{\sin(x+h) - \sin(x)}{h} \\ &= \lim_{h \to 0} \frac{\sin(x)\cos(h) + \cos(x)\sin(h) - \sin(x)}{h} && \text{(via trig identity for } \sin(x+h)) \\ f'(x) &= \lim_{h \to 0} \left(\sin(x)\frac{\cos(h) - 1}{h} + \cos(x)\frac{\sin(h)}{h} \right) && \text{(via some regrouping)} \end{aligned}$$

At this point, a sense of panic is natural. However, the reader should notice the waving red flag in the second term of the addition that is the Fundamental Trig Limit from Section 2.4. This should, for a moment, provide a sense of calm.

We now take a detour to evaluate a useful limit that is part of the first term of our addition.

$$\begin{aligned} \lim_{h \to 0} \frac{\cos(h) - 1}{h} &= \lim_{h \to 0} \frac{\cos(h) - 1}{h} \cdot \frac{\cos(h) + 1}{\cos(h) + 1} && \text{(multiplying by a smart choice of 1)} \\ &= \lim_{h \to 0} \frac{\cos^2(h) - 1}{h(\cos(h) + 1)} \\ &= \lim_{h \to 0} \frac{(1 - \sin^2(h)) - 1}{h(\cos(h) + 1)} && \text{(via Pythagorean identity)} \\ &= \lim_{h \to 0} \frac{-\sin^2(h)}{h(\cos(h) + 1)} \\ &= \lim_{h \to 0} \frac{\sin(h)}{h} \cdot \frac{-\sin(h)}{\cos(h) + 1} \\ \lim_{h \to 0} \frac{\cos(h) - 1}{h} &= 1 \cdot \frac{0}{2} = 0 && \text{(via Fundamental Trig Limit)} \end{aligned}$$

We can now return to finish our original problem:

$$\begin{aligned} f'(x) &= \lim_{h \to 0} \left(\sin(x)\frac{\cos(h) - 1}{h} + \cos(x)\frac{\sin(h)}{h} \right) \\ &= \sin(x)(0) + \cos(x)(1) && \text{(via the limit above and Fundamental Trig Limit)} \\ f'(x) &= \cos(x) \end{aligned}$$

So, given $f(x) = \sin(x)$, we have that $f'(x) = \cos(x)$.

EXERCISE

Show that the derivative of the cosine function $f(x) = \cos(x)$ is $f'(x) = -\sin(x)$.

Before we can compute our last derivative of this section, we first need to take a step back. We are familiar with Euler's number, e, which naturally was not discovered by Leonhard Euler but rather by Jacob Bernoulli.

However, we are perhaps not as *comfortable* with e as we are with another irrational number: π. Besides having a day dedicated to it as an excuse to eat pie, it is brought to our attention early in our mathematical education that the value of π comes from a circle. Specifically, π is the ratio of a circle's circumference to its diameter.

On the other hand, the value of e is commonly defined in several different contexts. Bernoulli's original definition of the value of e came from his study of compound interest:

$$e = \lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x,$$

with another common limit definition for the number being

$$e = \lim_{x\to 0}(1+x)^{1/x}.$$

Later, in MATH 138, you will find that

$$e = \frac{1}{0!}+\frac{1}{1!}+\frac{1}{2!}+\frac{1}{3!}+\ldots$$

However, in our context, we will define the value of e in relation to exponential growth and decay functions, $f(x) = a^x$ $(a > 0)$, which are common in modeling.

Notice that all such functions, $f(x) = a^x$ $(a > 0)$, go through the point $(0, 1)$ since $f(0) = a^0 = 1$. However, depending on the value of the base, a, the slope of the tangent line to the function through $(0, 1)$ will differ. This can be seen in Figure 3.3.4.

Definition 3.3.5
Euler's number, e

We define e to be the unique value of $a \in \mathbb{R}$ such that the slope of the tangent line to $f(x) = a^x$ through $(0, 1)$ is 1.

This definition tells us that for $f(x) = e^x$ we have

$$\begin{aligned} f'(0) &= 1 \\ \lim_{h\to 0}\frac{e^{0+h}-e^0}{h} &= 1 \\ \lim_{h\to 0}\frac{e^h-1}{h} &= 1. \end{aligned}$$

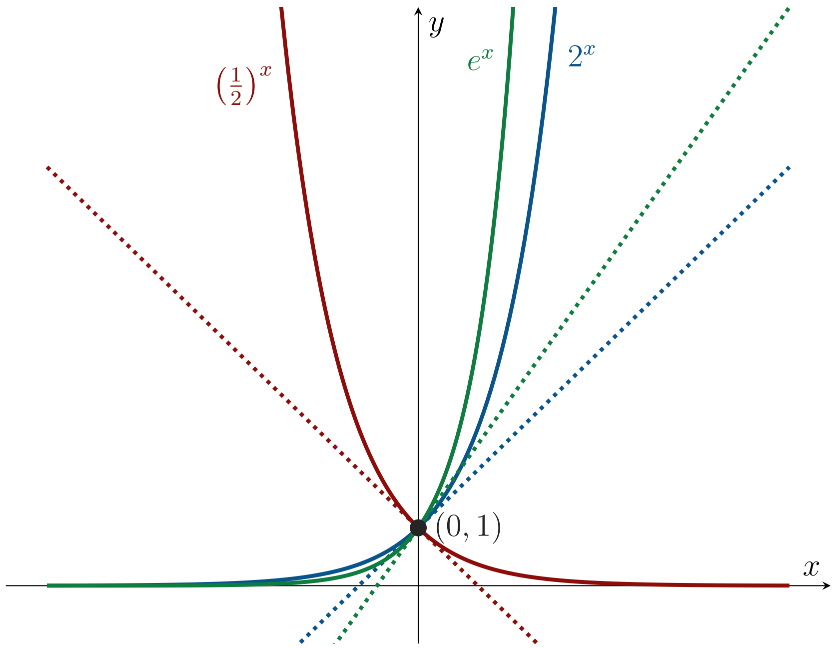


Figure 3.3.4: All curves of the form $f(x) = a^x$ $(a > 0)$ go through the point $(0, 1)$ but have varying slopes of their tangent line at that point.

Example 3.3.6

Find the derivative of the natural exponential function $f(x) = e^x$.

Solution:

$$\begin{aligned} f'(x) &= \lim_{h \to 0} \frac{e^{x+h} - e^x}{h} \\ &= \lim_{h \to 0} e^x \left(\frac{e^h - 1}{h} \right) \\ &= e^x(1) && \text{(from the above limit)} \\ f'(x) &= e^x \end{aligned}$$

So, given $f(x) = e^x$, we have that $f'(x) = e^x$.

A LOOK AHEAD

We need to do a little more work before we can discuss the derivative of the general exponential function $f(x) = a^x$ $(a > 0)$. This will be something for you to look forward to in Section 3.4!

Section 3.3 Problems

3.3.1. Show that $f(x) = |\sin x|$ is continuous but not differentiable for any $x \in \{k\pi, k \in \mathbb{Z}\}$.

3.3.2. The trigonometric functions $\sin(\omega x)$ and $\cos(\omega x)$ exhibit useful symmetries upon differentiation.

(a) Use the definition of the derivative to show that for $\omega \in \mathbb{R}$,

$$\frac{d}{dx}\sin(\omega x) = \omega\cos(\omega x), \quad \text{and} \quad \frac{d}{dx}\cos(\omega x) = -\omega\sin(\omega x).$$

(b) Use part (a) to show that

$$\frac{d^2}{dx^2}\sin(\omega x) = -\omega^2\sin(\omega x), \quad \text{and} \quad \frac{d^2}{dx^2}\cos(\omega x) = -\omega^2\cos(\omega x)$$

(c) An equation that involves a function $y(x)$ and its derivatives is called a differential equation. Show that $y(x) = A\cos(\omega x) + B\sin(\omega x)$ satisfies the differential equation,

$$y''(x) + \omega^2 y(x) = 0, \tag{3.2}$$

for arbitrary constants $A, B \in \mathbb{R}$.

(d) If, in addition to Eq. 3.2, we are told that $y(0) = 1$ and $y'(0) = 0$, determine the function $y(x)$.

3.3.3. The exponential function has the unique property that derivatives are transformed to scalar multiplication.

(a) Use the definition of the derivative to show that for $\lambda \in \mathbb{R}$,

$$\frac{d}{dx}e^{\lambda x} = \lambda e^{\lambda x}$$

(b) Show that $y(x) = Ae^{\lambda x}$ obeys the differential equation,

$$y'(x) = \lambda y(x), \tag{3.3}$$

for arbitrary constant A.

(c) From Ex. 2c you should obtain an equation that looks like Eq. 3.3. Determine a value of λ and A that best fits the data.

3.4 Derivative Rules

In practice, we do not tend to compute derivatives from first-principles. Rather, just as with limits, we have rules which we apply. The idea here is to take our base set of functions for which we found the derivatives in the previous section, and build out our toolkit of functions we can differentiate.

Theorem 3.4.1 **(Constant Multiple Rule)**

Assume that $f(x)$ is differentiable at $x = a$. Let $h(x) = cf(x)$ where $c \in \mathbb{R}$. Then $h(x)$ is differentiable at $x = a$ and

$$h'(a) = cf'(a).$$

Proof:

Let $f(x)$ be differentiable at $x = a$ and let $c \in \mathbb{R}$.

Then,

$$\begin{aligned}
(cf)'(a) &= \lim_{h\to 0} \frac{cf(a+h) - cf(a)}{h} \\
&= \lim_{h\to 0} c \cdot \frac{f(a+h) - f(a)}{h} \\
&= c \lim_{h\to 0} \frac{f(a+h) - f(a)}{h} \\
&= cf'(a). && \text{(via the definition of the derivative at a point)}
\end{aligned}$$

□

Theorem 3.4.2 **(Sum Rule)**

Assume that $f(x)$ and $g(x)$ are differentiable at $x = a$. Let $h(x) = f(x) + g(x)$. Then $h(x)$ is differentiable at $x = a$ and

$$h'(a) = f'(a) + g'(a).$$

Proof:

Let $f(x)$ and $g(x)$ be differentiable at $x = a$.

Then,

$$\begin{aligned}
(f+g)'(a) &= \lim_{h\to 0} \frac{(f(a+h) + g(a+h)) - (f(a) + g(a))}{h} \\
&= \lim_{h\to 0} \left(\frac{f(a+h) - f(a)}{h} + \frac{g(a+h) - g(a)}{h} \right) \\
&= \lim_{h\to 0} \frac{f(a+h) - f(a)}{h} + \lim_{h\to 0} \frac{g(a+h) - g(a)}{h} \\
& && \text{(because we know both limits exist)} \\
&= f'(a) + g'(a). && \text{(via the definition of the derivative at a point)}
\end{aligned}$$

□

The previous two rules are likely not all that surprising, and would match a naive guess at the derivative of the given function. We saw their proofs fell in to place neatly with direct application of limit laws. However, the following two rules might be unexpected for a first-time viewer. Additionally, their proofs contain subtle but important justifications, which the reader is encouraged to focus on.

Theorem 3.4.3 **(Product Rule)**

Assume that $f(x)$ and $g(x)$ are differentiable at $x = a$. Let $h(x) = f(x)g(x)$. Then $h(x)$ is differentiable at $x = a$ and

$$h'(a) = f'(a)g(a) + f(a)g'(a).$$

Proof:

Assume that $f(x)$ and $g(x)$ are differentiable at $x = a$.

Then,

$$\begin{aligned}
(fg)'(a) &= \lim_{h\to 0} \frac{f(a+h)g(a+h) - f(a)g(a)}{h} \\
&= \lim_{h\to 0} \frac{f(a+h)g(a+h) - f(a+h)g(a) + f(a+h)g(a) - f(a)g(a)}{h} \\
&& \text{(via adding 0)} \\
(fg)'(a) &= \lim_{h\to 0} \left(f(a+h)\left[\frac{g(a+h)-g(a)}{h}\right] + g(a)\left[\frac{f(a+h)-f(a)}{h}\right]\right). \\
&& \text{(via regrouping)}
\end{aligned}$$

Now, since $f(x)$ is differentiable at $x = a$, we know by Theorem 3.2.5 that it is continuous at $x = a$.

Then, by definition of continuity, we have $\lim\limits_{h\to 0} f(a+h) = f(a)$.

Further, by definition of differentiability at a point, we have $\lim\limits_{h\to 0} \dfrac{f(a+h)-f(a)}{h} = f'(a)$ and $\lim\limits_{h\to 0} \dfrac{g(a+h)-g(a)}{h} = g'(a)$.

Returning to the matter at hand, we have

$$\begin{aligned}
(fg)'(a) &= \lim_{h\to 0} \left(f(a+h)\left[\frac{g(a+h)-g(a)}{h}\right] + g(a)\left[\frac{f(a+h)-f(a)}{h}\right]\right) \\
&= \lim_{h\to 0} f(a+h)\left[\frac{g(a+h)-g(a)}{h}\right] + g(a)\lim_{h\to 0}\frac{f(a+h)-f(a)}{h} \\
&= f(a)g'(a) + g(a)f'(a) && \text{(via our work above)} \\
(fg)'(a) &= f'(a)g(a) + f(a)g'(a).
\end{aligned}$$

$\square$

REMARK

A common mnemonic for remembering product rule is

$$(uv)' = vdu + udv$$

where the d indicates taking a derivative of the function it precedes.

Theorem 3.4.4 **(Quotient Rule)**

Assume that $f(x)$ and $g(x)$ are differentiable at $x = a$ and that $g(a) \neq 0$. Let $h(x) = \dfrac{f(x)}{g(x)}$. Then $h(x)$ is differentiable at $x = a$ and

$$h'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{[g(a)]^2}.$$

Proof:

Assume that $f(x)$ and $g(x)$ are differentiable at $x = a$ and that $g(a) \neq 0$.

Note that

$$\left(\frac{f}{g}\right)'(a) = \left(f \cdot \frac{1}{g}\right)'(a)$$

Applying product rule, we then have

$$\left(\frac{f}{g}\right)'(a) = f'(a)\left(\frac{1}{g}\right)(a) + f(a)\left(\frac{1}{g}\right)'(a).$$

Moreover,

$$\begin{aligned}
\left(\frac{1}{g}\right)'(a) &= \lim_{h\to 0} \frac{\frac{1}{g(a+h)} - \frac{1}{g(a)}}{h} \\
&= \lim_{h\to 0} \frac{g(a) - g(a+h)}{hg(a+h)g(a)} \\
&= \lim_{h\to 0} \left(-\frac{g(a+h) - g(a)}{h} \cdot \frac{1}{g(a+h)g(a)}\right) \\
\left(\frac{1}{g}\right)'(a) &= -\frac{g'(a)}{[g(a)]^2}
\end{aligned}$$

where the final step above comes via the definition of the derivative at a point and the fact that $g(x)$ being differentiable at $x = a$ implies it is continuous there, giving $\lim_{h\to 0} g(a+h) = g(a)$.

Returning to the problem at hand, we have

$$\begin{aligned}
\left(\frac{f}{g}\right)'(a) &= f'(a)\left(\frac{1}{g}\right)(a) + f(a)\left(\frac{1}{g}\right)'(a) \\
&= \frac{f'(a)}{g(a)} + f(a)\left(-\frac{g'(a)}{[g(a)]^2}\right) \\
&= \frac{f'(a)g(a) - f(a)g'(a)}{[g(a)]^2}.
\end{aligned}$$

$\square$

REMARK

A common mnemonic for remembering quotient rule is

$$\left(\frac{u}{v}\right)' = \frac{vdu - udv}{v^2}$$

where the d indicates taking a derivative of the function it precedes. Be mindful that unlike product rule, there is a subtraction here, so order is crucial.

Theorem 3.4.5 **(Power Rule)**

Assume that $f(x) = x^\alpha$ where $\alpha \in \mathbb{R} \setminus \{0\}$. Then $f(x)$ is differentiable and

$$f'(x) = \alpha x^{\alpha-1}$$

wherever $x^{\alpha-1}$ is defined.

A LOOK AHEAD

We will see a proof of the power rule for when $x \neq 0$ in Section 3.8.

REMARK

Through the use of the power, sum, and constant multiple rules, along with the derivative of a constant function, we can find the derivative of any polynomial.

We have for polynomial $p(x) = a_0 + a_1x^1 + a_2x^2 + ... + a_nx^n$ that

$$\begin{aligned} p'(x) &= 0 + (1)a_1x^{1-1} + (2)a_2x^{2-1} + ... + (n)a_nx^{n-1} \\ p'(x) &= a_1 + 2a_2x + ... + na_nx^{n-1}. \end{aligned}$$

We note in particular that $p'(x)$ exists for all $x \in \mathbb{R}$, so polynomials $p(x)$ are differentiable for all $x \in \mathbb{R}$.

Theorem 3.4.6 **(Chain Rule)**

Assume that $y = f(x)$ is differentiable at $x = a$ and that $z = g(y)$ is differentiable at $y = f(a)$. Then $h(x) = g \circ f(x) = g(f(x))$ is differentiable at $x = a$ and

$$h'(a) = g'(f(a))f'(a).$$

The proof of the chain rule is arduous. Should you wish examine the proof, it is left as an exercise to the reader (ha!) to refer to another resource.

REMARK

Leibniz notation is particularly well-suited to the chain rule.

If we have $z = g(y)$ then $g'(y) = \dfrac{dz}{dy}$ and if we have $y = f(x)$ then $f'(x) = \dfrac{dy}{dx}$.

Then, according to chain rule above, we have

$$\begin{aligned}\frac{dz}{dx} &= g'(f(x))f'(x)\\ &= \left.\frac{dz}{dy}\right|_{f(x)} \left.\frac{dy}{dx}\right|_{x}\end{aligned}$$

That is,

$$\frac{dz}{dx} = \frac{dz}{dy}\frac{dy}{dx}$$

which is appealing to us visually, in the sense that if we were to *pretend* these are fractions, the dy's would cancel and both sides would be equal. This is a *notational* feature only.

With all of our derivative rules in play, we proceed to find the derivatives of a few more common functions.

Example 3.4.7 Find the derivative of the tangent function $f(x) = \tan(x)$.

Solution:

$$\begin{aligned}\frac{d}{dx}(\tan(x)) &= \frac{d}{dx}\left(\frac{\sin(x)}{\cos(x)}\right)\\ &= \frac{\frac{d}{dx}(\sin(x))\cos(x) - \sin(x)\frac{d}{dx}(\cos(x))}{[\cos(x)]^2} && \text{(via quotient rule)}\\ &= \frac{\cos(x)\cos(x) - \sin(x)(-\sin(x))}{\cos^2(x)}\\ &= \frac{\cos^2(x) + \sin^2(x)}{\cos^2(x)}\\ &= \frac{1}{\cos^2(x)}\\ \frac{d}{dx}(\tan(x)) &= \sec^2(x)\end{aligned}$$

Example 3.4.8 Find the derivative of the secant function $f(x) = \sec(x)$.

Solution:

$$\begin{aligned}\frac{d}{dx}(\sec(x)) &= \frac{d}{dx}\left(\frac{1}{\cos(x)}\right)\\ &= \frac{\frac{d}{dx}(1)(\cos(x)) - (1)\frac{d}{dx}(\cos(x))}{[\cos(x)]^2} && \text{(via quotient rule)}\end{aligned}$$

$$
\begin{aligned}
&= \frac{(0)(\cos(x)) - (1)(-\sin(x))}{\cos^2(x)} \\
&= \frac{\sin(x)}{\cos^2(x)} \\
&= \frac{\sin(x)}{\cos(x)} \cdot \frac{1}{\cos(x)} \\
\frac{d}{dx}(\sec(x)) &= \tan(x)\sec(x)
\end{aligned}
$$

EXERCISE

Show that the derivative of the cosecant function $f(x) = \csc(x)$ is $f'(x) = -\cot(x)\csc(x)$.

EXERCISE

Show that the derivative of the cotangent function $f(x) = \cot(x)$ is $f'(x) = -\csc^2(x)$.

Example 3.4.9 Find the derivative of the general exponential function $f(x) = a^x$ $(a > 0)$.

Solution:

Note that $a^x = e^{\ln(a^x)} = e^{x\ln(a)}$. Then,

$$
\begin{aligned}
\frac{d}{dx}(a^x) &= \frac{d}{dx}\left(e^{x\ln(a)}\right) \\
&= \frac{d}{dx}(e^x)\bigg|_{x\ln(a)} \cdot \frac{d}{dx}(x\ln(a))\bigg|_x && \text{(via chain rule)} \\
&= e^x\big|_{x\ln(a)} \cdot \ln(a)\big|_x \\
&= e^{x\ln(a)}\ln(a) \\
\frac{d}{dx}(a^x) &= a^x\ln(a).
\end{aligned}
$$

We now try our hand at several examples, some of which require multiple rules.

Example 3.4.10 Given $f(x) = 2^x\tan(x)$, find $f'(x)$.

Solution:

$$
\begin{aligned}
f'(x) &= \frac{d}{dx}(2^x)\tan(x) + 2^x\frac{d}{dx}\tan(x) && \text{(via product rule)} \\
f'(x) &= 2^x\ln(2)\tan(x) + 2^x\sec^2(x)
\end{aligned}
$$

Example 3.4.11 Given $f(x) = \dfrac{\sec(x)}{x+1}$, find the equation of the tangent line to $f(x)$ at $x = 0$.

Solution:

We first find the slope of the tangent line at $x = 0$, which would be given by $f'(0)$.

$$f'(x) = \frac{\frac{d}{dx}(\sec(x))(x+1) - \sec(x)\frac{d}{dx}(x+1)}{(x+1)^2} \qquad \text{(via quotient rule)}$$

$$f'(x) = \frac{\tan(x)\sec(x)(x+1) - \sec(x)(1)}{(x+1)^2}$$

Then, we have

$$\begin{aligned} f'(0) &= \frac{\tan(0)\sec(0)(0+1) - \sec(0)}{(0+1)^2} \\ &= \frac{0-1}{1} \\ f'(0) &= -1. \end{aligned}$$

Now, we need the y-coordinate of the point of tangency, given by

$$f(0) = \frac{\sec(0)}{0+1} = \frac{1}{1} = 1.$$

Finally, the equation of the tangent line is thus

$$y = -1(x-0) + 1 \implies y = 1 - x.$$

Example 3.4.12 Given $f(x) = \dfrac{x^2\sin(x)}{x^4+1}$, find $f'(x)$.

Solution:

$$\begin{aligned} f'(x) &= \frac{\frac{d}{dx}(x^2\sin(x))(x^4+1) - (x^2\sin(x))\frac{d}{dx}(x^4+1)}{(x^4+1)^2} && \text{(via quotient rule)} \\ &= \frac{(\frac{d}{dx}(x^2)\sin(x) + x^2\frac{d}{dx}\sin(x))(x^4+1) - (x^2\sin(x))(4x^3+0)}{(x^4+1)^2} && \text{(via product rule)} \\ f'(x) &= \frac{(2x\sin(x) + x^2\cos(x))(x^4+1) - 4x^5\sin(x)}{(x^4+1)^2} \end{aligned}$$

Example 3.4.13 Given $f(x) = (3x^2+2x+7)^{19}$, find $f'(x)$.

Solution:

$$\begin{aligned} f'(x) &= 19(3x^2+2x+7)^{19-1}\cdot\frac{d}{dx}(3x^2+2x+7) && \text{(via chain and power rules)} \\ f'(x) &= 19(3x^2+2x+7)^{18}(6x+2) \end{aligned}$$

Example 3.4.14 Given $f(x) = 2^{3x} + 5^{\cos(x)}$, find $f'(x)$.

Solution:

$$\begin{aligned} f'(x) &= 2^{3x}\ln(2)\frac{d}{dx}(3x) + 5^{\cos(x)}\ln(5)\frac{d}{dx}(\cos(x)) && \text{(via chain rule)} \\ &= 2^{3x}\ln(2)(3) + 5^{\cos(x)}\ln(5)(-\sin(x)) \\ f'(x) &= 2^{3x}3\ln(2) - 5^{\cos(x)}\ln(5)\sin(x) \end{aligned}$$

Example 3.4.15 Given $f(x) = \sin(3^x x^e + \sqrt{x})$, find $f'(x)$.

Solution:

$$\begin{aligned} f'(x) &= \cos(3^x x^e + \sqrt{x})\frac{d}{dx}(3^x x^e + \sqrt{x}) && \text{(via chain rule)} \\ &= \cos(3^x x^e + \sqrt{x})\left(\frac{d}{dx}(3^x)x^e + 3^x\frac{d}{dx}(x^e) + \frac{d}{dx}x^{1/2}\right) && \text{(via product \& sum rules)} \\ &= \cos(3^x x^e + \sqrt{x})\left(3^x\ln(3)x^e + 3^x e x^{e-1} + \frac{1}{2}x^{1/2-1}\right) && \text{(via power rule)} \\ f'(x) &= \cos(3^x x^e + \sqrt{x})\left(3^x\ln(3)x^e + 3^x e x^{e-1} + \frac{1}{2\sqrt{x}}\right) \end{aligned}$$

Example 3.4.16 Given $f(x) = e^{\sin(x^2)}$, find $f'(x)$.

Solution:

$$\begin{aligned} f'(x) &= e^{\sin\left(x^2\right)}\frac{d}{dx}\left(\sin\left(x^2\right)\right) && \text{(via chain rule)} \\ &= e^{\sin\left(x^2\right)}\left(\cos\left(x^2\right)\right)\frac{d}{dx}\left(x^2\right) && \text{(via chain rule again)} \\ f'(x) &= e^{\sin\left(x^2\right)}\left(\cos\left(x^2\right)\right)(2x) \end{aligned}$$

Section 3.4 Problems

3.4.1. Let $f(x) = \dfrac{ax+b}{ax-b}$ where $a \neq 0, b \neq 0$.

(a) Find $f'(x)$ using any method.

(b) Show that given $x \neq \frac{b}{a}$, $abf'(x) < 0$.

3.4.2. In each case, find $f'(x)$ using any method.

(a) $f(x) = 5^x \sin x + (x^3 + x^2)\cos x$.

(b) $f(x) = \dfrac{x^2 + x - 2}{x^3 + 6}$.

(c) $f(x) = \sqrt{2\tan^2 x + 3}$.

3.4.3. Compute $\dfrac{dy}{dx}$ and $\dfrac{d^2y}{dx^2}$ in each case.

(a) $y = \cos(x^2)$.

(b) $y = \cos^2 x$.

3.4.4. Using ONLY the Chain Rule and the Product Rule (and not the Reciprocal/Quotient rules), prove the following.

(a) The derivative of an even function is odd.

(b) The Quotient Rule. [Hint: $\dfrac{f(x)}{g(x)} = f(x)\,(g(x))^{-1}$].

3.4.5. If $y = f(u)$ and $u = g(x)$ where f and g are twice differentiable functions, prove that

$$\frac{d^2y}{dx^2} = \frac{d^2y}{du^2}\left(\frac{du}{dx}\right)^2 + \frac{dy}{du}\frac{d^2u}{dx^2}.$$

3.5 Linear Approximation

In this section, we explore the idea of approximating potentially complicated functions with a much simpler family of functions: linear functions.

Let's first remind ourselves of the following version of the definition of the derivative at a point:

$$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}.$$

In defining the derivative with a limit, what we are saying is that for x-values *very close* to a we have:

$$f'(a) \approx \frac{f(x) - f(a)}{x - a}.$$

To be clear, we are saying we can *approximate* the value of the derivative at a point with the above expression for x-values *very close* to a.

We can then rearrange the above approximation to read

$$f(x) \approx f'(a)(x - a) + f(a)$$

for x-values *very close* to a.

Now wait just a minute - this looks strikingly familiar. This is nothing but the equation of the tangent line to $f(x)$ at $x = a$ from Definition 3.2.3.

We can understand our approximation graphically. Imagine we have a more complicated function, $f(x)$, and we know the value $f(a)$. If we want to approximate the value of $f(x)$ for x *very close* to a, rather than moving along the more complicated function, $f(x)$, we instead move along the tangent line to the function at $x = a$, as shown in Figure 3.5.5.

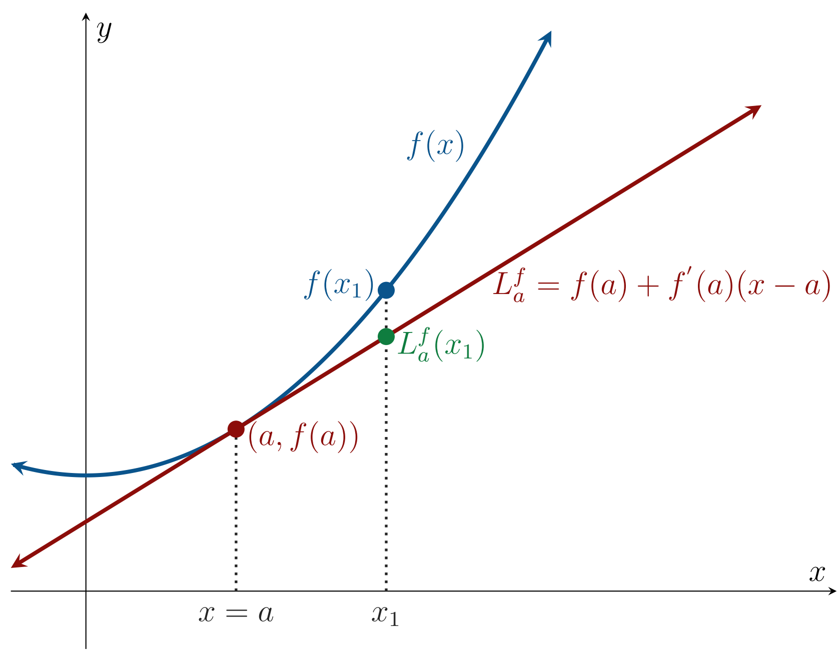


Figure 3.5.5: We can move along the tangent line to approximate the value of $f(x_1)$ for x_1 near a, rather than move along the more complicated function. Here, $L_a^f(x_1) \approx f(x_1)$.

We formalize this idea with the following definition.

Definition 3.5.1

Linear Approximation, Linearization, Tangent Line Approximation, $L_a^f(x)$

Let $f(x)$ be differentiable at $x = a$. Then the **linear approximation**, **linearization**, or **tangent line approximation**, $L_a^f(x)$, to $f(x)$ at $x = a$ is the function

$$L_a^f(x) = f'(a)(x - a) + f(a)$$

where if $f(x)$ is clear from context, we write instead $L_a(x)$.

REMARKS

Notice the following:

- $L_a^f(a) = f'(a)(a - a) + f(a) = 0 + f(a) = f(a)$, as we would expect, since the tangent line by definition matches the function at the point of tangency.
- $(L_a^f)'(a) = f'(a)(1 - 0) + 0 = f'(a)$, as we would expect, since the tangent line by definition matches the function's derivative at the point of tangency.
- That is, the linear approximation matches key features of the original function at a and would thus reasonably make a fair approximation *very close* to a.
- The linear approximation is the only linear function such that $L(a) = f(a)$ and $L'(a) = f'(a)$.

Additionally, keep in mind that a should be a value at which we can easily find $f(a)$ and should be *very close* to the x-value we want to approximate $f(x)$ at.

Example 3.5.2

Let $f(x) = \sqrt{x}$. Use an appropriate linearization to approximate $\sqrt{3.98}$ and $\sqrt{4.05}$.

Solution:

First, we must choose an appropriate value of a. An intelligent choice would be $a = 4$ because we can quickly evaluate $f(a) = f(4) = \sqrt{4} = 2$ and because 4 is *very close* to 3.98 and 4.05.

Since we have already determined $f(4) = 2$, all that remains to write down our equation for the linearization is to find $f'(4)$. We have that $f'(x) = \dfrac{1}{2\sqrt{x}}$, so $f'(4) = \dfrac{1}{2\sqrt{4}} = \dfrac{1}{4}$.

This gives us

$$L_4(x) = \frac{1}{4}(x - 4) + 2.$$

We can then find

$$\sqrt{3.98} \approx L_4(3.98) = \frac{1}{4}(3.98 - 4) + 2 = \frac{1}{4}\left(-\frac{1}{50}\right) + 2 = \frac{399}{200}$$

and

$$\sqrt{4.05} \approx L_4(4.05) = \frac{1}{4}(4.05 - 4) + 2 = \frac{1}{4}\left(\frac{1}{20}\right) + 2 = \frac{161}{80}.$$

If you were to break out your calculator (covered in dust from lack of use) you would find that our approximations are

$$\sqrt{3.98} \approx \frac{399}{200} = 1.995$$

and

$$\sqrt{4.05} \approx \frac{161}{80} = 2.0125$$

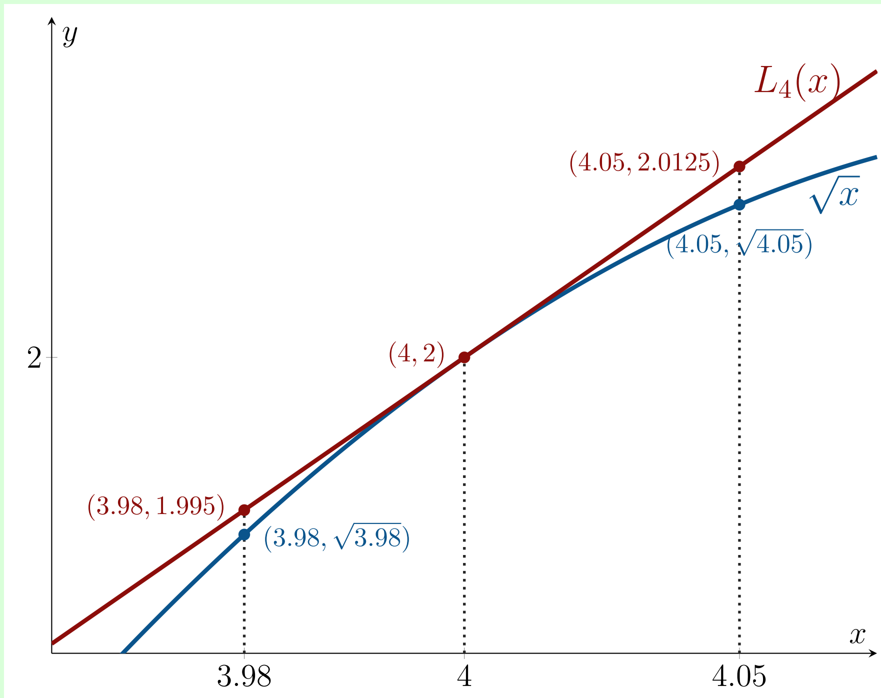
whereas the true values are

$$\sqrt{3.98} = 1.99499...$$

and

$$\sqrt{4.05} = 2.01246...$$

Overall, our approximations were pretty good! We can see we *overestimated* the true values of both $\sqrt{3.98}$ and $\sqrt{4.05}$, which we can see graphically by noting that the tangent line at $x = 4$ lies *above* $f(x)$.



Now, what if we attempted to use our linearization in the previous example to approximate the value of $\sqrt{1}$ or $\sqrt{16}$? We would find that $\sqrt{1} \approx L_4(1) = \frac{1}{4}(1-4)+2 = 1.25$ and that $\sqrt{16} \approx L_4(16) = \frac{1}{4}(16-4)+2 = 5$.

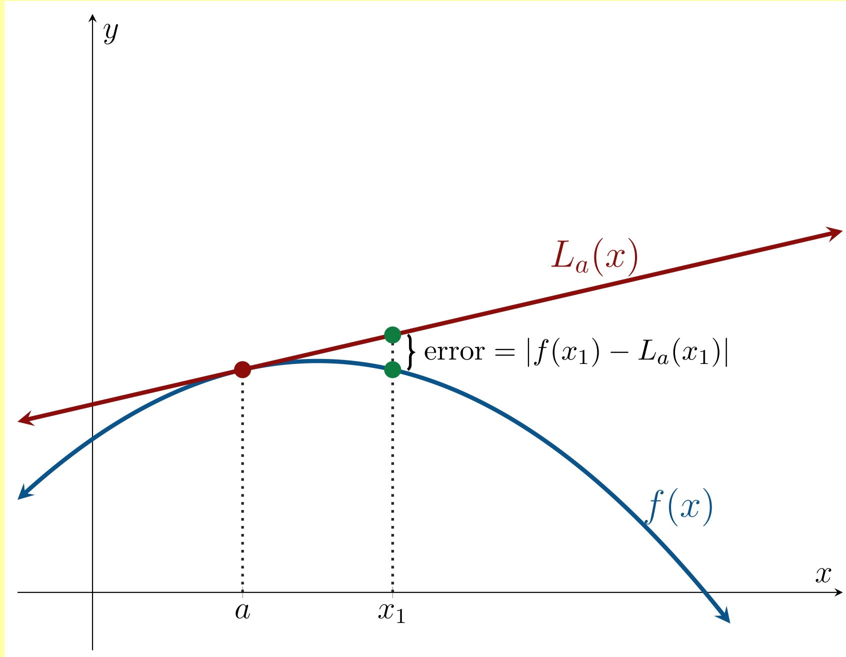
Both of these are much worse approximations than those we found in the example. There is no point in using an approximation if it is *bad*. What constitutes *bad* exactly would depend on the context: to an astrophysicist it might be totally acceptable to be off by around $\pm 10^{28}$ kilograms when measuring the mass of a star, but a nurse might only have an acceptable error of around ± 0.05 grams in providing a dose of morphine to a patient.

Here, we are talking about the *error* of our linear approximation.

Definition 3.5.3
Error in the Linear Approximation

Let $f(x)$ be differentiable at $x = a$. Then, the **error in using the linear approximation** to approximate the value of $f(x)$ is given by

$$\text{error} = |f(x) - L_a(x)|.$$



Regardless of context, it would be useful to have an *upper bound* on the error of our linear approximation, so we would know how *bad* the approximation would be in the worst possible case.

So, what affects how good or bad a linear approximation might be? Well, one factor you may have already guessed from our emphasis on the words 'very close' is that in general, the further we move away from the point of tangency, the worse our approximation gets. It should be noted that this does not have to be the case, as demonstrated in Figure 3.5.6.

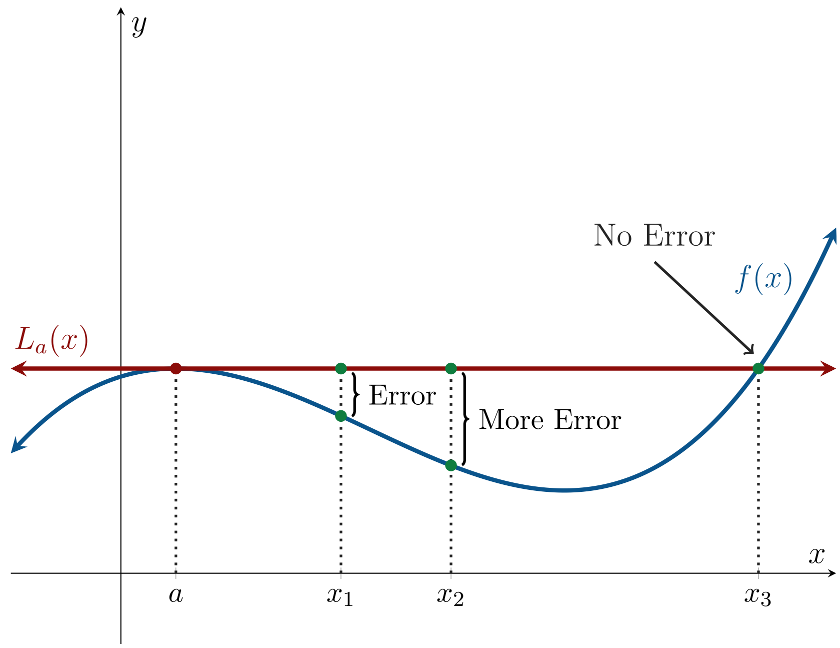


Figure 3.5.6: Generally, the further we get from the point of tangency, the worse the linear approximation is. However, this does not *have* to be the case. Here, we see that the approximation gets worse, and then gets better to the point of being exactly correct (and then gets worse again).

Another feature of a function that will affect the error is how *curved* it is, as demonstrated in Figure 3.5.7. In general, the more curved the function is near $x = a$, the greater the potential error.

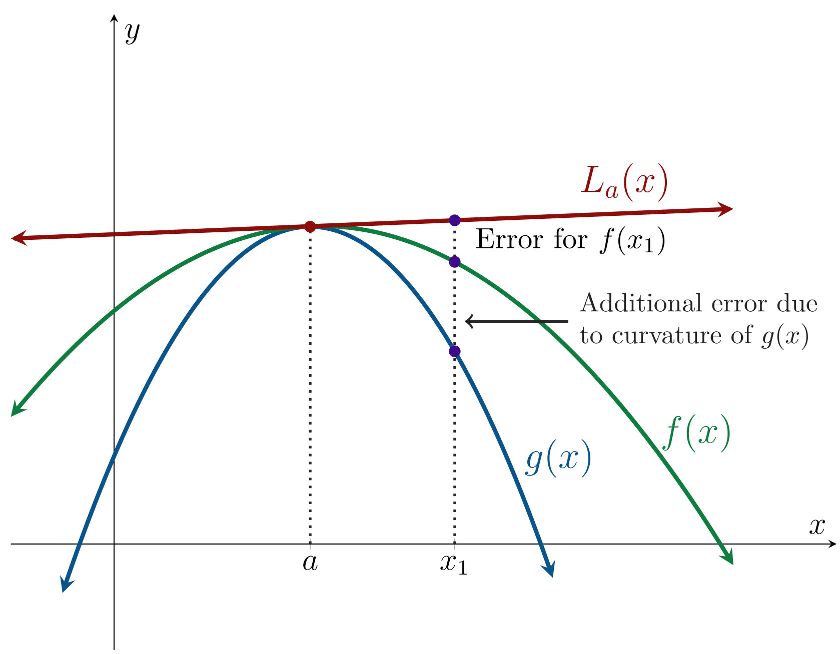


Figure 3.5.7: Here, since $g(x)$ is more 'curved' than $f(x)$, there is more error in using $L_a(x_1)$ to approximate $g(x_1)$ than $f(x_1)$.

Now, the question is: how do we describe the idea of curvature? Note that the value of the slopes changes more gradually for the less curved function $f(x)$, and the value of the slopes changes more rapidly for the more curved function $g(x)$.

What we are discussing is the *rate of change* of the slopes of the tangent lines. That is: the rate of change of the rate of change. We recognize this to be the second derivative. In particular, the greater the magnitude of the rate of change of the first derivative, the more curved the function will be. This means we are concerned with $|f''(x)|$ when discussing how curved the function is.

Theorem 3.5.4 **(Upper Bound on Error in the Linear Approximation)**

Assume that $f(x)$ is such that $|f''(x)| \leq M$ for each x in an interval I containing $x = a$. Then,

$$\text{error} = |f(x) - L_a(x)| \leq \frac{M}{2}(x-a)^2$$

for each $x \in I$.

A LOOK AHEAD

You will see in MATH 138 that the theorem above is a special case of *Taylor's Inequality.*

We will return to the relationship of the second derivative and the curvature of a function in Section 4.6.

Example 3.5.5 Find an upper bound on the error in using $L_9(x)$ to approximate $f(x) = \sqrt{x}$ on $[4, 15]$.

Solution:

We first deal with finding M, which is the upper bound of $|f''(x)|$ on the interval $[4, 15]$.

We have that $f'(x) = \frac{1}{2}x^{-1/2}$, so $f''(x) = -\frac{1}{4}x^{-3/2} = -\dfrac{1}{4x^{3/2}}$.

Now, we examine $|f''(x)| = \left|-\dfrac{1}{4x^{3/2}}\right| = \dfrac{1}{4x^{3/2}}$ on the interval $[4, 15]$. We note that as the denominator becomes larger, the quantity because smaller. Thus, this quantity takes on its largest value on $[4, 15]$ when $x = 4$.

So, $M = \dfrac{1}{4(4)^{3/2}} = \dfrac{1}{32}$.

Then, we have

$$\text{error} = |\sqrt{x} - L_9(x)| \leq \frac{\frac{1}{32}}{2}(x-9)^2 = \frac{1}{64}(x-9)^2.$$

Recall that we are trying to find an upper bound on the error. So we seek to maximize the right-hand side of the above inequality.

Now, on the interval $[4, 15]$ we note that $(x-9)^2$ is largest for $x = 15$.

Thus, we have

$$\text{error} = |\sqrt{x} - L_9(x)| \leq \frac{1}{64}(15-9)^2 = \frac{1}{64}(36) = \frac{36}{64} = \frac{9}{16}$$

So, an upper bound on the error in using $L_9(x)$ to approximate $f(x) = \sqrt{x}$ on $[4, 15]$ is $\dfrac{9}{16}$.

REMARK

In the example above, we found *one* valid choice of M. Any larger value would also suffice.

Consider the linear approximation of $f(x) = \sin(x)$ at $x = 0$ as an approximation of $\sin(1)$. If we calculate M as the maximum of $|f''(x)|$ on $[0, 1]$, you will get $M = \sin(1)$; but using this is silly since $\sin(1)$ is what you are approximating! Therefore, $M = 1$ is a better choice.

One application of the linear approximation is in estimating change.

Here, we assume that we know the value of $f(a)$ and would like to know how $f(x)$ changes as we move to an x-value, x_1, close to $x = a$.

That is, we would like to know $\Delta y = f(x_1) - f(a)$ if $\Delta x = x_1 - a$.

Now, using the fact that $f(x_1) \approx L_a(x_1)$, we have

$$\begin{aligned}\Delta y &\approx L_a(x_1) - f(a)\\ &= f'(a)(x_1 - a) + f(a) - f(a)\\ &= f'(a)\Delta x.\end{aligned}$$

Thus, we have that

$$\Delta y \approx f'(a)\Delta x.$$

We can understand this graphically as demonstrated in Figure 3.5.8.

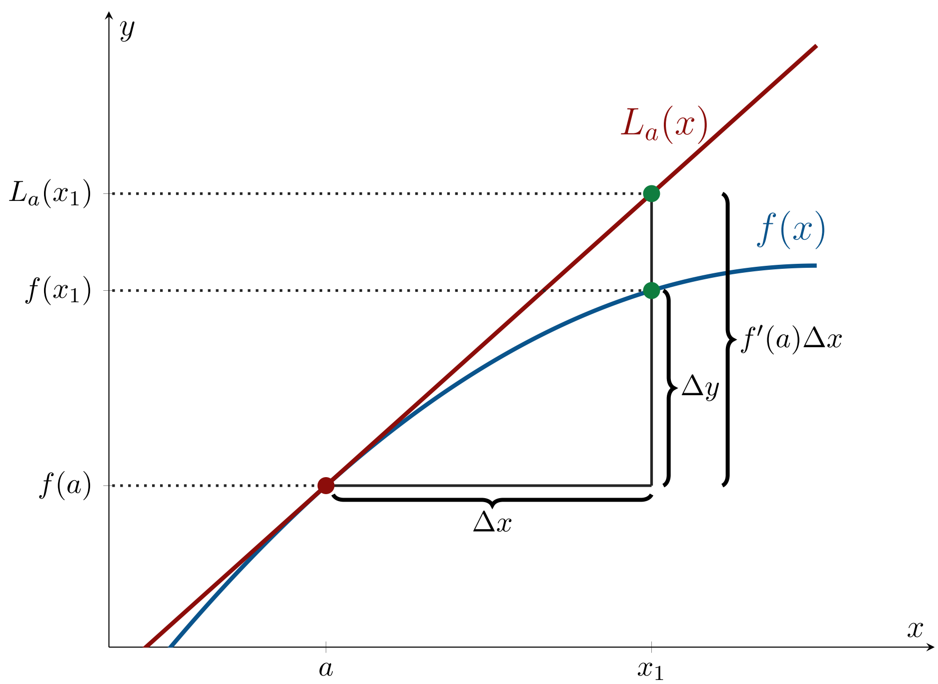


Figure 3.5.8: Here, we would like to estimate how $f(x)$ changes, Δy, for a small change in x, Δx. We can approximate Δy as $f'(a)\Delta x$.

Example 3.5.6

The Minions are trying to help Gru steal the Moon*, which we assume to be spherical, and are using their shrinking ray in order to have it fit in their pocket. By the time you notice, the radius of the Moon is 1000 kilometres. Two seconds later, the radius has decreased by another 500 metres. Estimate the change in volume of the Moon over these two seconds.

Solution:

According to our work above, we have

$$\Delta V \approx V'(1000)\Delta r.$$

Here, we have a sphere, so

$$V(r) = \frac{4}{3}\pi r^3 \implies V'(r) = 4\pi r^2 \implies V'(1000) = 4\pi(1000)^2 = 4000000\pi$$

and we are told the radius decreases by 500 metres, which we translate to kilometres to give

$$\Delta r = -0.5.$$

Thus, we have that $\Delta V \approx (4000000\pi)(-0.5) = -2000000\pi$ km^3.

For reference, the actual change in volume is

$$\Delta V = V(999.5) - V(1000) = \frac{4}{3}\pi(999.5)^3 - \frac{4}{3}\pi(1000)^3 = -1999000.1\bar{6}\pi \text{ km}^3.$$

**Despicable Me* is a work of art - go watch it instead of doing calculus.

Section 3.5 Problems

3.5.1. In each case, determine the equation of the tangent to $y = f(x)$ at the point where $x = a$.

(a) $f(x) = x^2, a = 3$

(b) $f(x) = \cos(x), a = -\frac{3\pi}{4}$.

(c) $f(x) = e^x, a = \ln\pi$.

(d) $f(x) = 4^x, a = -2$.

3.5.2. Let $f(x) = \sin(x)$.

(a) Determine the equation for the linear approximation to $f(x)$ at $x = a$ for any $a \in \mathbb{R}$.

(b) Determine the equation for the linear approximation to $f(x)$ at $x = \frac{\pi}{3}$.

(c) Use your answer to (b) to approximate $\sin(1)$.

(d) Since $f''(x) = -\sin(x)$, we can safely say that $|f''(x)| \leq 1$. Given this fact, what is the maximum error (worst-case scenario) for your answer in (c)?

3.5.3. Rich bought a yoga ball that is made of material which, when the ball is properly inflated, is a sphere with outer radius R. The manufacturer has determined that the material can tolerate a 4% "stretch" beyond specifications, meaning that if the ball is inflated in such a way that the surface area increases by more than 4% of the actual size, the material will rupture and the ball will deflate rather suddenly.
In the instructions for inflation, consumers are told to inflate the ball so that one side touches a wall and the other side touches a box placed $2R$ units away from the wall. Rich uses a ruler to measure $2R$ units from a wall, and then inflates the ball according to the instructions. If Rich's ruler and Rich's measurement skills create an error of 3% in excess of what he thinks he is measuring (that is, instead of inflating to a diameter of $2R$ it is inflated to a diameter of $1.03(2R)$), should we expect the ball to survive this initial inflation?

3.6 Newton's Method

Recall that in Section 2.10, we learned a root-finding algorithm called the Bisection Method. This method utilized the Intermediate Value Theorem.

In this section, we make use of the linear approximation to motivate another, more efficient, root-finding algorithm: Newton's Method.

The idea of the iterative process is as follows:

1. Make an initial guess, x_1, of where the root of $f(x)$ is. Intermediate Value Theorem may be useful here.
2. Take the linear approximation, $L^f_{x_1}(x)$ and find its root. Call this value x_2.
3. Repeat Step 2 at x_2 to find x_3 and so on.

This procedure is demonstrated in Figure 3.6.9.

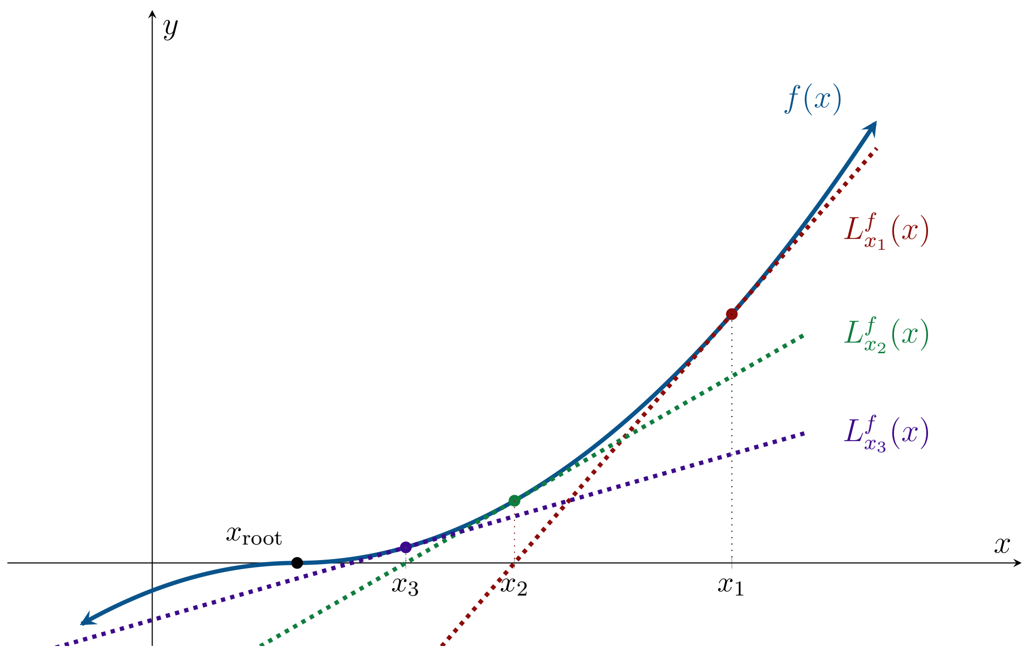


Figure 3.6.9: After making an initial guess, x_1, as to the value of x_{root}, we use the root of the linear approximation $L^f_{x_1}(x)$ as our next guess, x_2. We then iterate.

Now, focus on Step 2 of the process outlined above. During the first iteration, in that step, we are seeking to find x_2 where

$$L^f_{x_1}(x_2) = 0$$

We can then use of our knowledge of the formula for the linear approximation to find

$$\begin{aligned} f'(x_1)(x_2 - x_1) + f(x_1) &= 0 \\ f'(x_1)x_2 &= f'(x_1)x_1 - f(x_1) \\ x_2 &= x_1 - \frac{f(x_1)}{f'(x_1)} \quad \text{if } f'(x_1) \neq 0 \end{aligned}$$

We can generalize this to state:

Method

Newton's Method

Newton's Method is a root-finding algorithm based upon the recursive formula

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

REMARKS

- In general, Newton's Method converges much faster than Bisection Method.
- Notice that Newton's Method requires f to be differentiable at x_n whereas Bisection Method only requires continuity.
- Newton's Method does not always converge, whereas Bisection Method always converges.
  - Consider where $f'(x_n) = 0$. The formula is not computable, and this is the case of a horizontal tangent line - which would not have a root for the next iteration.
  - Consider $f(x) = \sqrt[5]{x}$ for which Newton's Method will not work for any guess of root $x_1 \neq 0$, as the formula gives $x_{n+1} = -6x_n$
  - Consider $f(x) = x^3 + x^2 - 3x + 3$ for which a guess of $x_1 = 0$ will lead to $x_2 = 1$ and $x_3 = 0$ a thus a cycle which will not converge to the root.
  - Some choices of x_1 for a given $f(x)$ can lead to convergence to a different root than desired.
- A common stopping point for Newton's Method is when two successive iterations match to a certain number of decimal places.

Example 3.6.1

Find the root of $f(x) = 3x^4 + 15x^3 - 125x - 1500$ on $[4, 5]$, with error at most 10^{-5}. Use $x_1 = 4$.

Solution:

First, for the sake of sanity, we note that $f(4) < 0$, $f(5) > 0$, and that since $f(x)$ is a polynomial it is continuous everywhere. Thus by IVT, there is a root on $(4, 5)$.

Now, $f'(x) = 12x^3 + 45x^2 - 125$.

Using the Newton's Method formula this gives

$$x_{n+1} = x_n - \frac{3x_n^4 + 15x_n^3 - 125x_n - 1500}{12x_n^3 + 45x_n^2 - 125}.$$

We find that

$$\begin{aligned}
x_2 &= 4 - \frac{3(4)^4 + 15(4)^3 - 125(4) - 1500}{12(4)^3 + 45(4)^2 - 125} \approx 4.1995598,\\
x_3 &\approx 4.1872682,\\
x_4 &\approx 4.1872187,\\
x_5 &\approx 4.1872187.
\end{aligned}$$

We note that x_4 and x_5 match up for the first 5 decimal places (we are looking for an error of 10^{-5}), so we stop here and round to 5 decimal places.

So there is a root of $f(x)$ at $x \approx 4.18722$.

As a quick check, $f(4.18721) < 0$ and $f(4.18723) > 0$, so IVT confirms there to be a root in between.

Example 3.6.2 Examine what happens when applying Newton's Method to find the root of $f(x) = \arctan(x)$ with an initial guess of $x_1 = 1.5$. You can take for granted* that $f'(x) = \dfrac{1}{1+x^2}$.

Solution:

You may know from prior studies that $f(x) = \arctan(x)$ has a root at the origin.

We are told that $f'(x) = \dfrac{1}{1+x^2}$. Newton's Method formula then gives

$$x_{n+1} = x_n - \frac{\arctan(x_n)}{\frac{1}{1+x_n^2}} = x_n - (1+x_n^2)\arctan(x_n).$$

We would find that

$$\begin{aligned} x_2 &\approx -1.694, \\ x_3 &\approx 2.321, \\ x_4 &\approx -5.114, \end{aligned}$$

which is clearly diverging.

That didn't work out for us. Let's try a closer guess of $x_1 = 1$. We would then find that

$$\begin{aligned} x_2 &\approx -0.571, \\ x_3 &\approx 0.117, \\ x_4 &\approx -0.001, \end{aligned}$$

which we can see is converging towards the expected root at $x = 0$.

*You will have the tools to find this derivative yourself by the end of Section 3.7.

Section 3.6 Problems

3.6.1. Newton's method provides a powerful algorithm to extract square-roots.

(a) For $a \in \mathbb{R}$, $a > 0$, show that Newton's Method applied to the function $f(x) = x^2 - a$ generates the recursive sequence,

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) \tag{3.4}$$

(b) Assuming that the sequence $\{x_n\}$ converges, show that the limit is $\sqrt{a}$.

(c) Make the substitution $x_n \mapsto \sqrt{a}\coth(b_n)$, where

$$\coth u = \frac{e^u + e^{-u}}{e^u - e^{-u}}$$

is the hyperbolic cotangent, to convert the original sequence to

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) \qquad \mapsto \qquad b_{n+1} = 2b_n.$$

Solve this sequence for b_n (*i.e.* for some initial term b_0, find b_n as an explicit function of n). Use your solution to prove that the original sequence $\{x_n\}$ converges to $\sqrt{a}$, and comment on the speed of this convergence with increasing n.

3.6.2. Use Newton's Method (Ex. 1) to approximate each of the following, correct to 5 decimal places.

(a) $\sqrt{7}$ (use initial guess $x_1 = 3$).

(b) $\sqrt{\pi}$ (use initial guess $x_1 = 2$).

3.6.3. Consider the function $f(x) = \dfrac{6x+1}{3x+5}$.

(a) What is the domain of f?

(b) Find $f'(x)$.

(c) There is only one point $c \in \mathbb{R}$ where $f(c) = 0$, find it directly.

(d) Now, starting with $x_1 = 5$ (a particularly foolish choice), perform 3 iterations of Newton's Method. Use 5 decimal places.

(e) It is clear that starting with $x_1 = 5$ will not lead us to the root. In fact, the sequence generated by Newton's method in this case diverges. Prove that using $x_1 = 5$ as a starting value, Newton's Method will not converge to the root of this function. (Hint: show the recursive sequence you get from Newton's Method is strictly decreasing).

3.6.4. Find the root of $\sin x = 1 - x$ to 5 decimal places. Use a sketch to estimate the initial point x_0.

3.6.5. How many roots does the equation $\tan x = x$ have? Find the one between $\frac{\pi}{2}$ and $\frac{3\pi}{2}$ to 5 decimal places.

3.7 Derivatives of Inverse Functions

Despite our extensive toolkit of known derivatives, there are still a few common functions whose derivatives we have not yet discussed.

In this section, the central idea is that if we already know the derivative of a given function $f(x)$, we would like to be able to find the derivative of its inverse $f^{-1}(x)$ from our knowledge of $f'(x)$, rather than starting from square one.

We will use our recent work with the linear approximation to attack this problem.

Recall from prior studies that we have two approaches for finding the inverse of a function:

1. Graphically, whereby we reflect our function across the line $y = x$, or equivalently, swap the x- and y-coordinates of the points on the function.
2. Algebraically, whereby we swap x and y in the equation of our function, and then solve for y.

Now, let us assume that we have an invertible function, $f(x)$, with a tangent line to the function at co-ordinate point (a, b), which we recognize is $L_a^f(x)$. If we take the graphical approach and reflect both of these across $y = x$, we will have depicted the inverse function, $f^{-1}(x)$, and a tangent line to the inverse function at co-ordinate point (b, a), which we recognize is $L_b^{f^{-1}}(x)$. We note we can also denote this new tangent line as $(L_a^f)^{-1}(x)$. This is demonstrated in Figure 3.7.10.

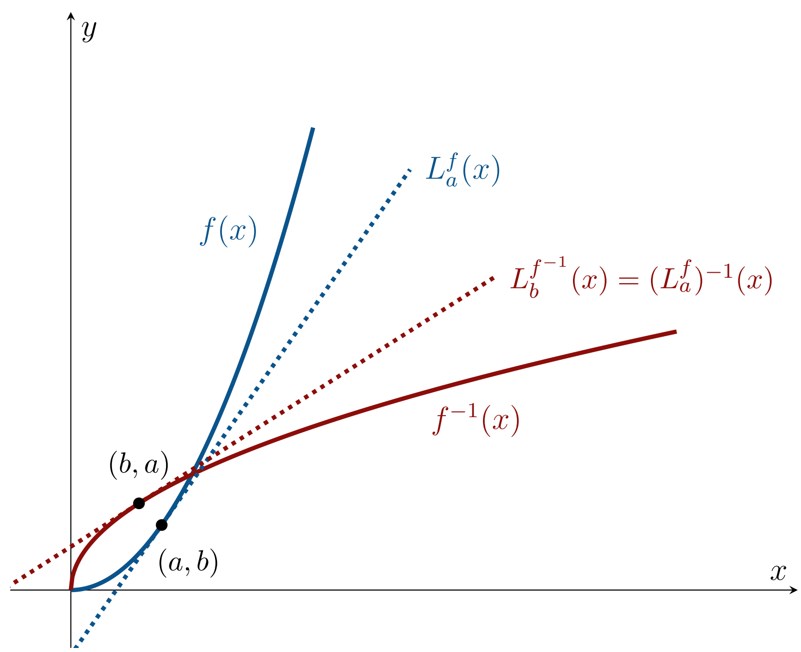


Figure 3.7.10: A function $f(x)$ with its linear approximation at $x = a$, $L_a^f(x)$, and their inverses.

Now, let's determine an equation for $L_b^{f^{-1}}(x)$ in two ways.

First, we directly use the point-slope form of the equation of a line. We know that this line goes through the point (b, a), and has slope $(f^{-1})'(b)$ from its construction as a tangent line. This gives

$$L_b^{f^{-1}}(x) = (f^{-1})'(b)(x - b) + a. \qquad (*)$$

Second, we take the equation of $L_a^f(x)$ and use it to algebraically find the inverse. We know by definition that

$$L_a^f(x) = y = f'(a)(x-a) + f(a)$$

but since $f(a) = b$, we have

$$L_a^f(x) = y = f'(a)(x-a) + b.$$

Swapping x and y, we have

$$x = f'(a)(y-a) + b.$$

We then solve for y, obtaining

$$y = \frac{1}{f'(a)}(x-b) + a$$

under the condition that $f'(a) \neq 0$.

So we have found that

$$L_b^{f^{-1}}(x) = \frac{1}{f'(a)}(x-b) + a. \qquad (**)$$

Comparing $(*)$ and $(**)$ which are two equations of the same line, we see that

$$(f^{-1})'(b) = \frac{1}{f'(a)} = \frac{1}{f'(f^{-1}(b))},$$

where the second equality comes from the fact that $f^{-1}(b) = a$ as (b, a) is a point on the graph of $f^{-1}(x)$.

We have arrived at:

Theorem 3.7.1 **(Inverse Function Theorem (IFT))**

Assume that $f(x)$ is continuous and invertible with inverse $f^{-1}(x)$ on $[c, d]$ and differentiable at $a \in (c, d)$, where $f'(a) \neq 0$. Then, $f^{-1}(x)$ is differentiable at $b = f(a)$ and

$$(f^{-1})'(b) = \frac{1}{f'(a)} = \frac{1}{f'(f^{-1}(b))}.$$

Furthermore, $L_a^f(x)$ is also invertible and

$$(L_a^f)^{-1}(x) = L_b^{f^{-1}}(x) = L_{f(a)}^{f^{-1}}(x).$$

REMARKS

- The IFT has enough notation to cause a migraine. Let it not be lost on the reader what the *point* of the IFT is. The IFT says that if we already know the derivative of the invertible function $f(x)$, we can use this to find the derivative of the inverse function $f^{-1}(x)$ without starting from scratch. We even have the potential to find $(f^{-1})'(b)$ without knowing what $f^{-1}(x)$ is, as we will see in Example 3.7.3 below.

- Consider what it means if $f'(a) = 0$. In this case, there would be a horizontal tangent line to $f(x)$ at $x = a$. When a horizontal line is reflected across $y = x$, it becomes a vertical line. That would mean that $L_b^{f^{-1}}(x)$ is a vertical tangent line, and as discussed in Section 3.2, this means that $f^{-1}(x)$ is not differentiable at $x = b$.

Example 3.7.2 Let $f(x) = x^5$. Find $(f^{-1})'(3)$.

Solution 1 (without IFT):

We can directly find (via the algebraic method) that if $f(x) = x^5$ then $f^{-1}(x) = \sqrt[5]{x}$.

Then, we have that

$$\begin{aligned}(f^{-1})'(x) &= \frac{d}{dx}x^{1/5} \\ &= \frac{1}{5}x^{-4/5},\end{aligned}$$

which gives us that $(f^{-1})'(3) = \frac{1}{5}(3)^{-4/5} = \frac{1}{5(3)^{4/5}}$.

Solution 2 (with IFT):

IFT tells us that

$$(f^{-1})'(3) = \frac{1}{f'(f^{-1}(3))}.$$

Now, we have that $f^{-1}(3) = \sqrt[5]{3}$, since we know $f^{-1}(x)$ from Solution 1.

And, for $f(x) = x^5$, we have that $f'(x) = 5x^4$.

Thus, we have $(f^{-1})'(3) = \frac{1}{f'(\sqrt[5]{3})} = \frac{1}{5(\sqrt[5]{3})^4} = \frac{1}{5(3)^{4/5}}$.

Now, the preceding example is a little bit goofy, since we were able to find $f^{-1}(x)$ and thus its derivative. However, this will not always be the case, which is when IFT can really shine.

Example 3.7.3 Let $f(x) = 3x^3 + x^2 + 2x + 4$. Find $(f^{-1})'(0)$.

Solution:

Here, we are not able to find the inverse function directly through rudimentary algebra. Even if the reader were to use computational software to find the inverse function, finding the derivative of the inverse function would be dreadful.

We avoid finding the inverse function altogether by making use of IFT, which tells us that

$$(f^{-1})'(0) = \frac{1}{f'(f^{-1}(0))}.$$

We now need to find $f^{-1}(0)$. Recall that $f^{-1}(0) = a \iff f(a) = 0$.

Trialling a few values, we notice that $f(-1) = 0$, so we have that $f^{-1}(0) = -1$.

Next, we note that $f'(x) = 9x^2 + 2x + 2$.

Thus, we have $(f^{-1})'(0) = \dfrac{1}{f'(-1)} = \dfrac{1}{9(-1)^2 + 2(-1) + 2} = \dfrac{1}{9}$.

We can also arrive at the IFT via chain rule.

Assume that $f(x)$ has inverse $f^{-1}(x)$ and both functions are differentiable. Now, by definition

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

Taking the derivative of both sides we have

$$\frac{d}{dx} f(f^{-1}(x)) = \frac{d}{dx} x.$$

Taking care to apply chain rule to the left-hand side, this gives us

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

Provided that $f'(f^{-1}(x)) \neq 0$, we arrive again at the IFT

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

Let us now use IFT to find the derivatives of common inverse functions.

Example 3.7.4 Given $f(x) = e^x$ and $f^{-1}(x) = \ln(x)$ $(x > 0)$, find $(f^{-1})'(x)$.

Solution:

IFT tells us that

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

Noting that $f'(x) = e^x$, this gives

$$\begin{aligned} (\ln(x))' &= \frac{1}{f'(\ln(x))} \\ &= \frac{1}{e^{\ln(x)}} \\ \frac{d}{dx}(\ln(x)) &= \frac{1}{x}. \end{aligned}$$

Now, if we consider our three most common trig functions $\sin(x)$, $\cos(x)$, and $\tan(x)$, the reader will notice that reflecting any of these across $y = x$ will not result in a function, due to the periodic nature of the original function.

As such, we consider each of these three functions on a specified restricted domain, such that they are invertible, with inverses $\arcsin(x)$, $\arccos(x)$, and $\arctan(x)$.

We restrict $f(x) = \sin(x)$ to the domain $x \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ with range $y \in [-1, 1]$, which results in $f^{-1}(x) = \arcsin(x)$ on the domain $x \in [-1, 1]$ with range $y \in [-\frac{\pi}{2}, \frac{\pi}{2}]$, as depicted in Figure 3.7.11.

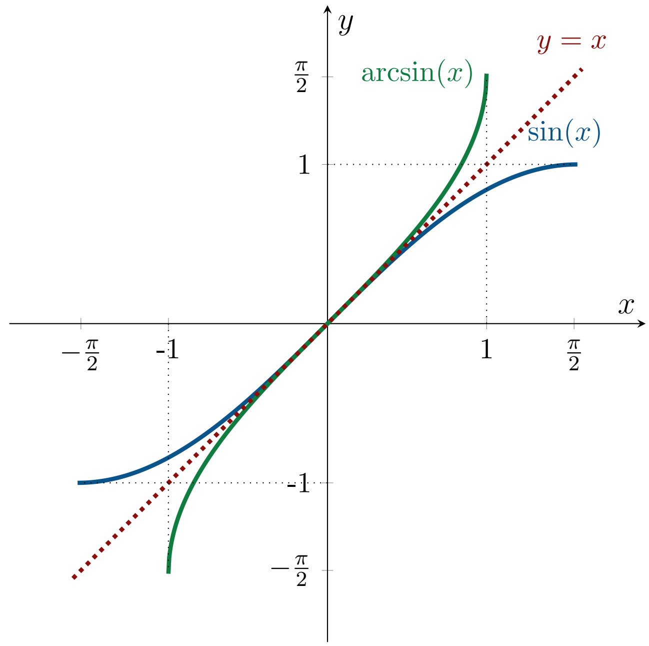


Figure 3.7.11: Graphs of restricted sin and arcsin, side by side.

Example 3.7.5

Find the derivative of $\arcsin(x)$.

Solution:

Here, we will set $f(x) = \sin(x)$ ($x \in [-\frac{\pi}{2}, \frac{\pi}{2}]$) and $f^{-1}(x) = \arcsin(x)$. Notice that $\sin(x)$ is continuous and has an inverse on $[-\frac{\pi}{2}, \frac{\pi}{2}]$ and is differentiable on $(-\frac{\pi}{2}, \frac{\pi}{2})$ with $f'(x) \neq 0$ on that interval. That is, the conditions of IFT are satisfied.

IFT tells us that

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

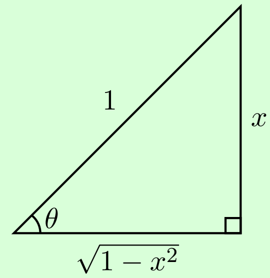
Noting that $f'(x) = \cos(x)$ ($x \in [-\frac{\pi}{2}, \frac{\pi}{2}]$), this gives

$$(\arcsin(x))' = \frac{1}{f'(\arcsin(x))}$$
$$(\arcsin(x))' = \frac{1}{\cos(\arcsin(x))}.$$

While it may seem like we are finished, we can greatly improve how this answer looks.

Recall that the output of $\arcsin(x)$ is an angle - let's call it θ. That is, $\arcsin(x) = \theta$ or equivalently $\frac{x}{1} = \sin(\theta)$ where $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$.

We can construct a right-angled triangle containing θ as depicted below, with the opposite side of length x and the hypotenuse of length 1. The base of the triangle is found via Pythagorean Theorem.



Now, we seek to find $\cos(\theta)$. Since $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$, we know that $\cos(\theta) \geq 0$, so we see from the right-angled triangle that $\cos(\theta) = \dfrac{\sqrt{1-x^2}}{1}$.

That is, $\cos(\arcsin(x)) = \sqrt{1-x^2}$.

Finally, this gives us

$$\frac{d}{dx}(\arcsin(x)) = \frac{1}{\sqrt{1-x^2}}.$$

You might notice that $\displaystyle\lim_{x\to-1^+} \frac{1}{\sqrt{1-x^2}} = \lim_{x\to1^-} \frac{1}{\sqrt{1-x^2}} = \infty$, which lines up with our expectation of the tangent lines to $\arcsin(x)$ becoming nearly vertical approaching the endpoints of its domain.

REMARK

Our work above required a behind-the-scenes assumption that $\arcsin(x)$ was differentiable. This is given to us for free via IFT since we know $\sin(x)$ is differentiable.

We restrict $f(x) = \cos(x)$ to the domain $x \in [0, \pi]$ with range $y \in [-1, 1]$, which results in $f^{-1}(x) = \arccos(x)$ on the domain $x \in [-1, 1]$ with range $y \in [0, \pi]$, as depicted in Figure 3.7.12.

EXERCISE

Show that $\dfrac{d}{dx}(\arccos(x)) = -\dfrac{1}{\sqrt{1-x^2}}$.

We restrict $f(x) = \tan(x)$ to the domain $x \in (-\frac{\pi}{2}, \frac{\pi}{2})$ with range $y \in (-\infty, \infty)$, which results in $f^{-1}(x) = \arctan(x)$ on the domain $x \in (-\infty, \infty)$ with range $y \in (-\frac{\pi}{2}, \frac{\pi}{2})$, as depicted in Figure 3.7.13.

EXERCISE

Show that $\dfrac{d}{dx}(\arctan(x)) = \dfrac{1}{1+x^2}$.

We can now solve derivative problems involving these inverse functions.

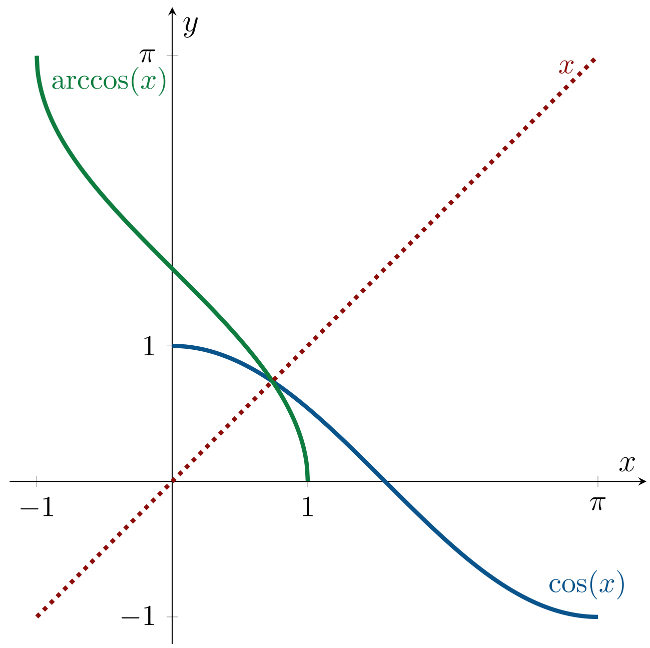


Figure 3.7.12: Graphs of restricted cos and arccos, side by side.

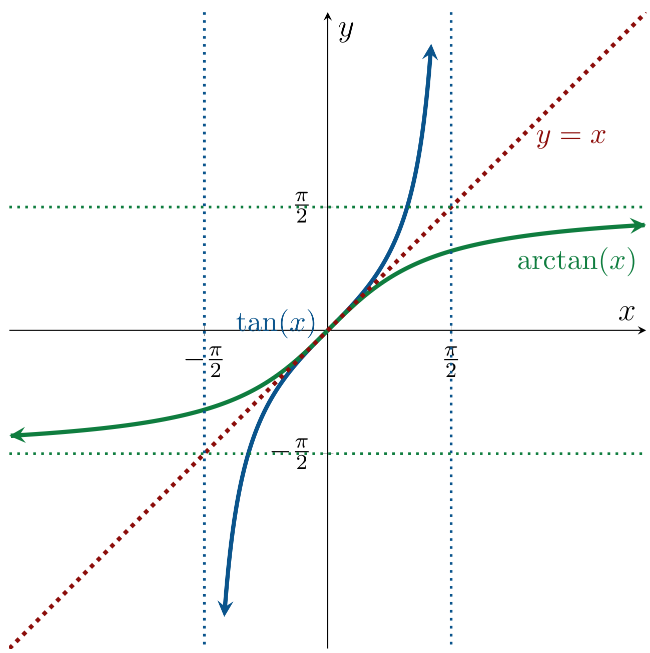


Figure 3.7.13: Graphs of restricted tan and arctan, side by side.

Example 3.7.6

Find the derivative of $f(x) = \arctan\left(5^{\cos(x)}\right)$.

Solution:

Applying chain rule, we have

$$f'(x) = \frac{1}{1 + (5^{\cos(x)})^2} \cdot \frac{d}{dx}(5^{\cos(x)})$$

$$
\begin{aligned}
&= \frac{1}{1+5^{2\cos(x)}} \cdot \left(5^{\cos(x)}\ln(5)\right) \cdot \frac{d}{dx}(\cos(x)) \\
f'(x) &= \frac{-5^{\cos(x)}\ln(5)\sin(x)}{1+5^{2\cos(x)}}.
\end{aligned}
$$

Example 3.7.7 Find the derivative of $f(x) = \ln(\arccos(3x^2))$.

Solution:

Applying chain rule, we have

$$
\begin{aligned}
f'(x) &= \frac{1}{\arccos(3x^2)} \cdot \frac{d}{dx}(\arccos(3x^2)) \\
&= \frac{1}{\arccos(3x^2)} \cdot \frac{-1}{\sqrt{1-(3x^2)^2}} \cdot \frac{d}{dx}(3x^2) \\
f'(x) &= \frac{-6x}{\arccos(3x^2)\sqrt{1-9x^4}}.
\end{aligned}
$$

Section 3.7 Problems

3.7.1. Simplify the following.

$$\text{i) } \sin(\arccos x), \quad \text{ii) } \sin(\arctan x), \quad \text{iii) } \tan(\arcsin x).$$

3.7.2. Determine $y'(x)$ for each of the following,

$$\text{i) } y(x) = x\arcsin x, \quad \text{ii) } y(x) = \arcsin\left(\tfrac{a}{x}\right) \quad (a \in \mathbb{R}), \quad \text{iii) } y(x) = \frac{\arcsin x}{\sin x}.$$

3.7.3. It is difficult to prove the identity

$$\arctan x + \operatorname{arccot} x = \frac{\pi}{2}, \quad (x > 0) \tag{3.5}$$

directly, but we can exploit the properties of the derivative to facilitate the proof.

(a) First, show that the derivative of the inverse cotangent, $\operatorname{arccot} x$, is,

$$\frac{d}{dx}\operatorname{arccot} x = \frac{-1}{1+x^2}.$$

Hint: The cotangent $\cot x = \cos x / \sin x$, and obeys the identity $1+\cot^2 x = \csc^2 x$.

(b) Next, show that the function $f(x) = \arctan x + \operatorname{arccot} x$ has zero derivative, $f'(x) = 0$.

(c) Use part (b) to find a suitable x that proves the identity (3.5).

3.7.4. Prove the identity

$$2\arcsin x = \arccos\left(1-2x^2\right) \qquad (x \geq 0)$$

Hint: Do Ex. 3 first.

3.8 Implicit and Logarithmic Differentiation

Every function we have learned to take the derivative of thus far has been an *explicit* function. That is, a function of the form $y = f(x)$ - where we have isolated for y in terms of an expression soley of x.

In this section, we explore how to take the derivatives of *implicit* functions. Here, this is any function that is not explicitly defined.

Before we learn the method of implicit differentiation, a preliminary point of order. We have seen from our derivative rules that

$$\frac{d}{dx}(x) = 1.$$

However, we must be cautious with the seemingly similar, but fundamentally different fact that

$$\frac{d}{dx}(y) = \frac{dy}{dx} = y'.$$

The difference here is that y is an *unknown function of* x, which we are differentiating with respect to x. This is crucial to understanding how implicit differentiation works.

Example 3.8.1 Given the implicit function $3x^3y^3 + x^2y + 13x = 12$, find y'.

Solution:

We start by differentiating both sides with respect to x, giving

$$\frac{d}{dx}(3x^3y^3 + x^2y + 13x) = \frac{d}{dx}(12).$$

The first two terms of the left-hand side will require product rule. With this in mind, we get

$$9x^2y^3 + 3x^3\frac{d}{dx}(y^3) + 2xy + x^2\frac{d}{dx}(y) + 13 = 0.$$

Now, with our understanding that y is some unknown function of x, this means that the term y^3 will require *chain rule*! This gives us

$$9x^2y^3 + 3x^3 3y^2\frac{d}{dx}(y) + 2xy + x^2\frac{d}{dx}(y) + 13 = 0,$$

and from our note before this example, we have

$$9x^2y^3 + 9x^3y^2y' + 2xy + x^2y' + 13 = 0.$$

At this point, all that remains is algebraic work. We separate the terms with y' from those without to obtain

$$9x^3y^2y' + x^2y' = -13 - 2xy - 9x^2y^3$$

and then we factor out the y' to get

$$y'(9x^3y^2 + x^2) = -13 - 2xy - 9x^2y^3$$

before finally solving to find

$$y' = \frac{-13 - 2xy - 9x^2y^3}{9x^3y^2 + x^2}.$$

Example 3.8.2 Given the implicit function $x^2+y^2=9$, find the points of tangency of any horizontal tangent lines.

Solution:

We start by differentiating both sides with respect to x, giving

$$\frac{d}{dx}(x^2+y^2)=\frac{d}{dx}(9).$$

Keeping in mind the need for chain rule for the y^2 term, this yields

$$2x+2yy'=0.$$

Isolating gives

$$y'=\frac{-2x}{2y}=-\frac{x}{y}.$$

Now, we seek to find where $y'=0$. This occurs when $0=-\dfrac{x}{y} \implies x=0$.

Substituting this into the implicit function $x^2+y^2=9$ yields $0^2+y^2=9 \implies y=\pm 3$.

Thus there are horizontal tangent lines at $(0,3)$ and $(0,-3)$.

The results of this calculation are as we might expect, since $x^2+y^2=9$ describes a circle with the points $(0,\pm 3)$ at the very top and bottom.

REMARK

We must be careful to not just blindly apply the technique of implicit differentiation. If the previous example were modified to begin with the equation $x^2+y^2=-9$, the steps of the implicit differentiation would be identical, leading again to $y'=-\frac{x}{y}$. However, in this context, that result is completely meaningless! There are no real-valued (x,y) co-ordinates that satisfy the equation $x^2+y^2=-9$, so there is no function for which to find the derivative.

We can use implicit differentiation to implement yet another differentiation technique: logarithmic differentiation. We will explore several contexts in which this is a useful tool.

The first classic use case for logarithmic differentiation is when we are dealing with a function of the form

$$h(x)=f(x)^{g(x)} \quad (f(x)>0).$$

It should be noted that we have already seen how to handle functions of the form $h(x)=f(x)^a$ (power and chain rules) and functions of the form $h(x)=a^{f(x)}$ (exponential function and chain rules). The difference here is that there is a function in both the base and the exponent.

Example 3.8.3 Given $y = x^x$ $(x > 0)$, find y'.

Solution:

We begin logarithmic differentiation by taking the natural logarithm of both sides, giving

$$\ln(y) = \ln(x^x).$$

The power of this process comes from the fact that we can now apply log laws to rewrite this as

$$\ln(y) = x\ln(x)$$

where now there is no longer a function raised to a function power but a product instead.

We now have an implicit function, and can proceed via implicit differentiation from this point. This gives

$$\frac{d}{dx}\ln(y) = \frac{d}{dx}(x\ln(x))$$

Being careful with chain rule on the left-hand side and applying product rule to the right-hand side, we obtain

$$\frac{1}{y}\frac{d}{dx}(y) = (1)\ln(x) + x\frac{1}{x}$$

and, upon cleaning this up, we have

$$\frac{y'}{y} = \ln(x) + 1$$

which we isolate to find

$$y' = y(\ln(x) + 1).$$

Finally, with the realization that we were given that $y = x^x$, we have

$$y' = x^x(\ln(x) + 1).$$

REMARK

An inquisitive reader may wonder if instead of logarithmic differentiation for the previous example, one could write $f(x) = e^{\ln(x^x)} = e^{x\ln(x)}$ and then proceed via chain rule. This would indeed work for the previous example, and in general for $f(x)^{g(x)}$ $(f(x) > 0)$, but why go that route when logarithmic differentiation is so much fun?

Before proceeding to the next classic use case for logarithmic differentiation, we examine the derivative of the function

$$h(x) = \ln|f(x)|.$$

When $f(x) > 0$, we have that $h(x) = \ln(f(x))$. This gives, as we saw in the previous example,

$$h'(x) = \frac{f'(x)}{f(x)}$$

On the other hand, when $f(x) < 0$, we have that $h(x) = \ln(-f(x))$. This gives

$$h'(x) = \frac{-f'(x)}{-f(x)} = \frac{f'(x)}{f(x)}.$$

Thus, regardless of the sign of $f(x)$, we have that

$$\frac{d}{dx} \ln|f(x)| = \frac{f'(x)}{f(x)}.$$

Example 3.8.4 Given $f(x) = \dfrac{\sqrt{x^2+4}(x-3)}{(x+5)^4(x^2+1)}$, find $f'(x)$.

Solution:

This derivative *could* be found through the use of quotient, product, and chain rules. However, this would be an incredibly tedious bit of work. Here, logarithmic differentiation greatly simplifies and shortens the process.

Let us take the natural logarithms of the absolute value of both sides (to cover our bases for when $f(x) < 0$ inside the logarithm). This gives

$$\ln|f(x)| = \ln\left(\left|\frac{\sqrt{x^2+4}(x-3)}{(x+5)^4(x^2+1)}\right|\right).$$

Now, we apply log laws to the right-hand side to obtain

$$\ln|f(x)| = \frac{1}{2}\ln|x^2+4| + \ln|x-3| - 4\ln|x+5| - \ln|x^2+1|,$$

which we implicitly differentiate, giving

$$\frac{d}{dx}\ln|f(x)| = \frac{d}{dx}\left(\frac{1}{2}\ln|x^2+4| + \ln|x-3| - 4\ln|x+5| - \ln|x^2+1|\right).$$

Through application of chain rule, and our investigation prior to this example with regards to absolute values inside a logarithm when taking a derivative, we obtain

$$\frac{f'(x)}{f(x)} = \frac{2x}{2(x^2+4)} + \frac{1}{x-3} - \frac{4}{x+5} - \frac{2x}{x^2+1}.$$

Rearranging yields

$$f'(x) = f(x)\left(\frac{x}{x^2+4} + \frac{1}{x-3} - \frac{4}{x+5} - \frac{2x}{x^2+1}\right).$$

Finally, substituting back in for $f(x)$ gives us that

$$f'(x) = \left(\frac{\sqrt{x^2+4}(x-3)}{(x+5)^4(x^2+1)}\right)\left(\frac{x}{x^2+4} + \frac{1}{x-3} - \frac{4}{x+5} - \frac{2x}{x^2+1}\right).$$

REMARK

In the previous example, applying absolute values helps when $f(x) < 0$, but this doesn't fix the fact that the logarithm is undefined for $f(3) = 0$. To calculate $f'(3)$, we need to go the long way of applying various rules.

We can also use logarithmic differentiation to prove the power rule.

Example 3.8.5 Prove that the derivative of $f(x) = x^\alpha$, $x \neq 0$, $\alpha \neq 0$, is $f'(x) = \alpha x^{\alpha-1}$ using logarithmic differentiation.

Solution:

We take the natural logarithms of the absolute value of both sides, giving

$$\ln|f(x)| = \ln|x^\alpha| = \ln|x|^\alpha.$$

Applying log laws, this yields

$$\ln|f(x)| = \alpha \ln|x|.$$

Implicitly differentiating, we have

$$\frac{d}{dx}\ln|f(x)| = \frac{d}{dx}\alpha\ln|x|$$

which becomes

$$\frac{f'(x)}{f(x)} = \alpha\frac{1}{x},$$

and rearranging gives

$$f'(x) = \alpha\frac{f(x)}{x}.$$

Substituting back in for $f(x)$ finally gives us that

$$f'(x) = \alpha\frac{x^\alpha}{x} = \alpha x^{\alpha-1}.$$

EXERCISE

Prove that the derivative of $f(x) = u(x)v(x)$, $f(x) \neq 0$, is $f'(x) = u'(x)v(x) + u(x)v'(x)$ using logarithmic differentiation.

Section 3.8 Problems

3.8.1. For each of the following, determine $\frac{dy}{dx}$, assuming that $f(x)$ and $g(x)$ are positive, differentiable functions.

a) $y = \ln\left(f\left(x^2\right)\right)$ b) $y = [f(x)]^{g(x)}$ c) $y = f\left(\sqrt{x}\,\ln\left(g(x)\right)\right)$

3.8.2. Find $\frac{dy}{dx}$ for $\arcsin(x^2y) + xy = 1$.

3.8.3. If $y = (\arcsin(x))^2$ and $0 < x < 1$, then prove that

$$(1 - x^2)\left(\frac{dy}{dx}\right)^2 = 4y,$$

and thereby deduce that

$$(1 - x^2)\frac{d^2y}{dx^2} - x\frac{dy}{dx} - 2 = 0.$$

Hint: To deduce the second expression, try differentiating the first expression with respect to x.

3.8.4. Find the equation of the tangent line at the point $(0, -1)$ to the curve defined by $xy + y^3 = \arctan(x) - 1$.

3.8.5. Use the logarithmic differentiation to find $\frac{dy}{dx}$

(a) $y = x^{\sin x}$ with $x > 0$.

(b) $y = (2x)^{x^{\frac{1}{3}}}$.

(c) $y = 2^{\sin(\sec x)}$.

Chapter 4

Applications of the Derivative

With a solid grasp of the meaning of the derivative and plenty of practice with its computation, it is now time to see how we can use the derivative. In this chapter, we will explore both theoretical and practical applications.

Given that there will be some practical applications at hand, the reader is warned that the frightening return of word problems to the mathematics classroom is approaching. Fret not however, as many proof still awaits you, and there is still much to explore in differential calculus, including a major link to your next calculus course.

By the end of this chapter, you will be able to

- apply differential calculus knowledge to word problems;
- sketch and optimize functions;
- evaluate limits of indeterminate form;
- bid MATH 137 farewell.

4.1 Related Rates

With the method of implicit differentiation fresh in our minds, we examine a practical application via related rates. Mathematicians have never been very good with creative naming, so all we might desire to know is in the name itself: we will be examining quantities whose *rates of change* are intrinsically *related*. That is, with the knowledge of how one quantity changes, we will be able to determine how the related quantity changes.

Let us explore an introductory example, before discussing how we want to tackle these kinds of problems, in general.

Example 4.1.1 Ryan the Rabbit is helping set up celebrations for Michelle the Moose's birthday party, and is on balloon duty. Ryan has purchased spherical balloons, and is blowing air into a balloon at a rate of 10 cm^3 per minute. Determine the rate at which the radius of the balloon is changing when the diameter of the balloon is 16 cm.

Solution:

In this problem, we are given the rate of change of volume with respect to time (in units of cm^3/min).

We are being asked to find the rate of change of radius with respect to time (which would be most easily found in units of cm/min).

Now, we recognize that both volume and radius are functions of time, which we will call $V(t)$ and $r(t)$ respectively.

These quantities are related via the volume of a sphere formula, $V = \frac{4}{3}\pi r^3$.

We now implicitly differentiate this equation with respect to time:

$$\begin{aligned}\frac{d}{dt}(V) &= \frac{d}{dt}\left(\frac{4}{3}\pi r^3\right)\\ \frac{dV}{dt} &= \frac{4}{3}\pi(3r^2)\frac{dr}{dt}\\ \frac{dV}{dt} &= 4\pi r^2\frac{dr}{dt}\end{aligned}$$

We were told in the problem that $\frac{dV}{dt} = 10$, and we seek to find $\frac{dr}{dt}$ when the diameter is 16 cm, meaning $r = 8$. This gives us

$$\begin{aligned}10 &= 4\pi(8)^2\frac{dr}{dt}\\ \frac{dr}{dt} &= \frac{10}{256\pi} = \frac{5}{128\pi}\end{aligned}$$

Thus, the radius of the balloon is increasing at a rate of $\dfrac{5}{128\pi}$ cm per minute.

The previous example highlights the need to have an organized plan of attack when handling related rates problems.

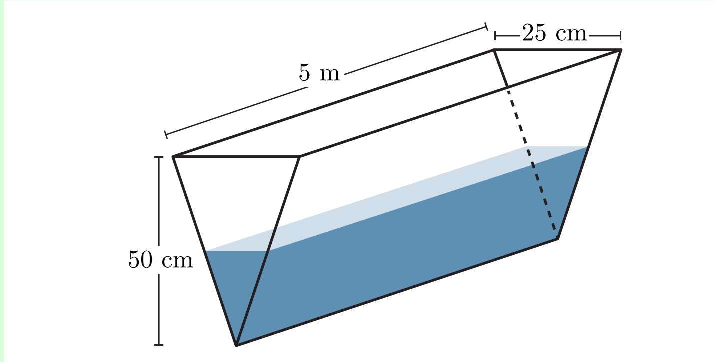
Method

Approach for Related Rates Problems

1. If necessary, draw a picture of the situation.
2. Identify the quantities that are changing.
3. Identify any quantities that are constant.
4. Find a key equation relating the quantities that are changing.
5. Implicitly differentiate the key equation.
6. Solve for the desired rate of change, using known quantities.
7. Conclude, ensuring that you include the appropriate units.

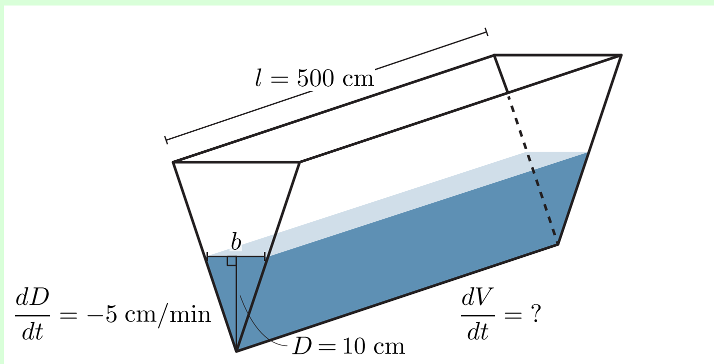
Example 4.1.2

Zack the Zebra is drinking water from Faisal the Farmer's trough. Faisal's trough is a triangular prism as shown below, with depth 50 cm, length 5 m, and width 25 cm. Zack is drinking the water such that the water depth reduces at 5 cm per minute. How fast is the volume of water changing when the water is 10 cm deep?



Solution:

Let's add some more labels to the diagram that will be useful to us as we progress through the problem.

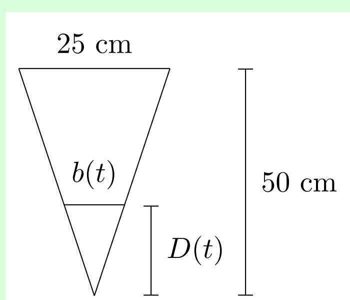


The quantities that are of interest to us are the water depth and water volume, which are changing with respect to time. We call these $D(t)$ and $V(t)$ respectively. Now, the 'base' of the water level triangle is also changing with respect to time, we will call this $b(t)$.

We also note that there is a constant quantity in this problem: the length of the trough, which is 5 metres, which we will convert to centimetres and assign $l = 500$.

Given that the water in the trough takes the form of a triangular prism, the key equation in this scenario is $V = \frac{1}{2}bDl$.

Now, before we implicitly differentiate, we would like to get rid of any variables which we are not interested in at the end of the problem. We know that we can utilize $l = 500$. However, we must figure out how to handle $b(t)$. Examining our diagram above, we realize that there are similar triangles at play, as brought into focus below.



Thus

$$\frac{b}{D} = \frac{25}{50} \implies b = \frac{1}{2}D.$$

Our key equation thus becomes $V = \frac{1}{2}(\frac{1}{2}D)D(500) = 125D^2$.

We now implicitly differentiate:

$$\frac{d}{dt}(V) = \frac{d}{dt}(125D^2)$$
$$\frac{dV}{dt} = 125(2D)\frac{dD}{dt}$$

We are now ready to plug in the known quantities, but we should be cautious. Given that the depth of the water is *decreasing*, to match mathematics to the physical scenario, we should denote that $\frac{dD}{dt} = -5$. We also know that $D = 10$ at the instant of interest.

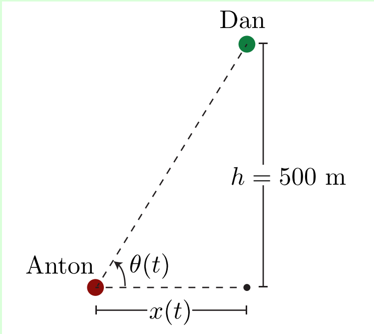
Thus, we have that $\frac{dV}{dt} = 125(2)(10)(-5) = -12500$. Here again, the negative sign conveys to us the quantity is *decreasing*.

So, the volume of water is decreasing at 12500 cm^3/min.

Example 4.1.3 Dan the Dragon is currently hovering in the air 500 m above a spot on the ground that is 250 m away from Anton the Antelope. If Dan starts flying towards Anton (maintaining his same height above the ground) at a speed of 4 m/s, how fast is Dan's angle of elevation changing 25 seconds after he begins moving?

Solution:

We start by drawing a diagram of the situation when Dan is in flight towards Anton.



Now, we would like to find the rate of change of the angle of elevation with respect to time, which we will call $\theta(t)$, utilizing our knowledge of the rate of change of the ground distance between Anton and Dan, which we will call $x(t)$.

We note that Dan's height above the ground is constant in this problem, so we will set $h = 500$.

The key equation relating these quantities would be $\tan(\theta) = \frac{h}{x}$, or substituting in for h, $\tan(\theta) = \frac{500}{x}$.

We now implicitly differentiate:

$$\begin{aligned}
\frac{d}{dt}(\tan(\theta)) &= \frac{d}{dt}\left(\frac{500}{x}\right) \\
\sec^2(\theta)\frac{d\theta}{dt} &= -\frac{500}{x^2}\frac{dx}{dt} \\
\frac{d\theta}{dt} &= -\frac{500}{x^2}\frac{dx}{dt}\cos^2(\theta)
\end{aligned}$$

We might at this point become momentarily excited seeing that we have isolated for the quantity we seek, but this joy should rapidly fade away upon realization that there is yet work to do to determine x and $\cos^2(\theta)$.

We are asked to determine $\frac{d\theta}{dt}$ at 25 seconds after Dan starts flying. In this time, since he flies at 4 m/s, he has flown $25 \times 4 = 100$ m. Thus, at this time, he is $250 - 100 = 150$ m away from Anton. So, we have found that $x = 150$ and (keeping in mind sign conventions) $\dfrac{dx}{dt} = -4$.

Now, at this time of interest, the side adjacent to the angle of elevation is 150 m, the side opposite the angle is 500 m, and using the Pythagorean Theorem, the hypotenuse of the triangle is thus $\sqrt{150^2 + 500^2} = \sqrt{272500}$ m. Then, we have that $\cos(\theta) = \dfrac{150}{\sqrt{272500}}$.

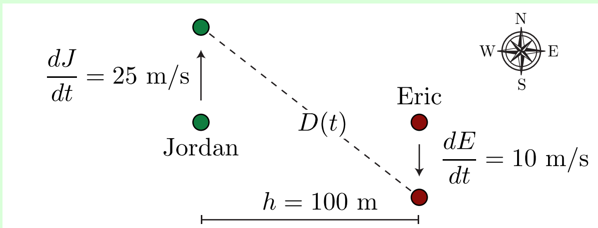
Finally, we can evaluate that $\dfrac{d\theta}{dt} = -\dfrac{500}{150^2}(-4)\left(\dfrac{150}{\sqrt{272500}}\right)^2 = \dfrac{2000}{272500} = \dfrac{4}{545}$. Note here that had we forgotten to ensure that $\dfrac{dx}{dt}$ was negative, our answer for $\dfrac{d\theta}{dt}$ would have been negative, which would make the angle appear to be decreasing as Dan flies closer, which does not make physical sense.

Thus, Dan's angle of elevation is increasing at $\dfrac{4}{545}$ rad/s.

Example 4.1.4 Jordan the Jaguar is standing 100 metres due West of Eric the Elephant. Jordan starts running due North at 25 m/s at the same time as Eric starts running due South at 10 m/s. How fast is the direct distance between Jordan and Eric changing 5 seconds later?

Solution:

We start by drawing a diagram of the situation when Jordan and Eric are running.



We are interested in finding the rate of change of the direct distance between Jordan and Eric, which we will call $D(t)$. We will utilize information about the distance Jordan has run, which we will call $J(t)$ and the distance Eric has run, which we will call $E(t)$. In particular, we will be concerned with their rates of change.

From our diagram, the horizontal distance between Jordan and Eric remains constant, which we will call $h = 100$.

We also see from our diagram that we can create the key equation, via Pythagorean Theorem, of $D^2 = (J+E)^2 + h^2$, or plugging in for h, $D^2 = (J+E)^2 + 10000$.

We now implicitly differentiate:

$$\begin{aligned}
\frac{d}{dt}(D^2) &= \frac{d}{dt}((J+E)^2 + 10000) \\
(2D)\frac{dD}{dt} &= 2(J+E)\frac{d}{dt}(J+E) + 0 \\
(2D)\frac{dD}{dt} &= 2(J+E)\left(\frac{dJ}{dt} + \frac{dE}{dt}\right) \\
\frac{dD}{dt} &= \frac{J+E}{D}\left(\frac{dJ}{dt} + \frac{dE}{dt}\right)
\end{aligned}$$

Now, 5 seconds into the run, Jordan has travelled $25 \times 5 = 125$ metres, while Eric has travelled $10 \times 5 = 50$ metres. That is, at our time of interest, $J = 125$ and $E = 50$. Then, utilizing our key equation, we find that at our time of interest, $D^2 = (125+50)^2 + 10000$ which gives that $D = \sqrt{40625}$.

Since both Jordan's and Eric's distances are increasing, we have that their rates of change are positive, $\frac{dJ}{dt} = 25$ and $\frac{dE}{dt} = 10$.

Putting this all together gives $\frac{dD}{dt} = \frac{125+50}{\sqrt{40625}}(25+10) = \frac{6125}{25\sqrt{65}} = \frac{245}{\sqrt{65}}$.

Thus, the direct distance between Jordan and Eric is increasing at $\frac{245}{\sqrt{65}}$ m/s.

Section 4.1 Problems

4.1.1. A rotating beacon is located 500 metres out in the water. Let A be the point on the shore that is closest to the beacon. As the beacon rotates at 10 rev/min, the beam of light sweeps down the shore once each time it revolves. At a point 10 metres along the shore from point A, how fast is beam of light moving? (Assume that the shore line is straight and runs perpendicular to the line connecting A and the beacon.)

4.1.2. A baseball diamond is a square 90 feet on a side. A player runs from first base to second base at 15 feet/sec. At what rate is the player's distance from third base decreasing when they are half-way from first to second base?

4.1.3. The sun is setting at the rate of $\frac{1}{250}$ rad/min, and appears to be dropping perpendicular to the horizon. How fast is the shadow of a 5 meter wall lengthening at the moment when the shadow is 10 meters long?

4.1.4. A person 1.8 meters tall walks at the rate of 1 meter per second toward a streetlight that is 4 meters above the ground. At what rate is their shadow shortening? At what rate is the tip of their shadow moving?

4.1.5. A picture of a sea horse is stamped on the surface of a latex balloon, and occupies 1/20 of the surface area. The balloon is inflated at a rate of 10 cm^3/sec. When the balloon has a radius of 5 cm, how quickly is the area of the sea horse increasing? (Assume that the balloon is a sphere.)

4.2 Extrema

One of the central applications of differential calculus is *optimization*, which we will end the course with in Section 4.10. To start off, we build out key vocabulary and groundwork that we will need both in optimization and the work we do along the way to get there.

4.2.1 Theoretical Foundations

When optimizing, we are looking to find a maximum or minimum value. You may understand what this means in an everyday sense, but we frame this mathematically with the following definition.

Definition 4.2.1

Global/Absolute Extrema

Suppose $f(x)$ is a function defined on some interval I. Let $c \in I$.

- We say there is a **global minimum** on I at $x = c$ if $f(c) \leq f(x)$ for all $x \in I$.
- We say there is a **global maximum** on I at $x = c$ if $f(c) \geq f(x)$ for all $x \in I$.
- We say there is a **global extremum** on I at $x = c$ if there is either a global minimum or global maximum there.

Here, the word global can be interchanged with **absolute**, and the plural form of extremum is **extrema**.

REMARK

There is some ambiguity in the usage of the term extremum. It may mean the x-value, the y-value, or the (x, y) co-ordinate point. It will always be made clear in the context of any question you are posed which one we would like to see.

Now, it is reasonable to wonder whether or not a function of interest defined on a non-empty interval achieves both a global maximum and a global minimum, or just one or the other, or perhaps even neither. We examine a few illuminating examples.

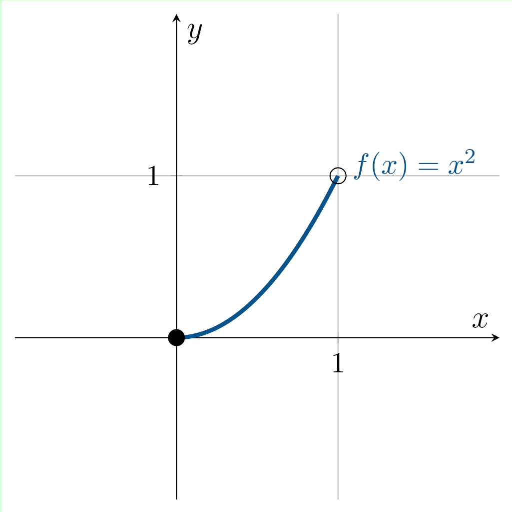
Example 4.2.2

Determine the co-ordinates of any global extrema of $f(x) = x^2$ on the interval $[0, 1)$.

Solution:

We note that since the right endpoint of the interval is exclusive, the function $f(x) = x^2$ keeps growing larger in value as x approaches 1. So, there is no global maximum.

We can see however that $f(x)$ obtains its lowest y-value at the left endpoint, so there is a global minimum at $(0, 0)$.

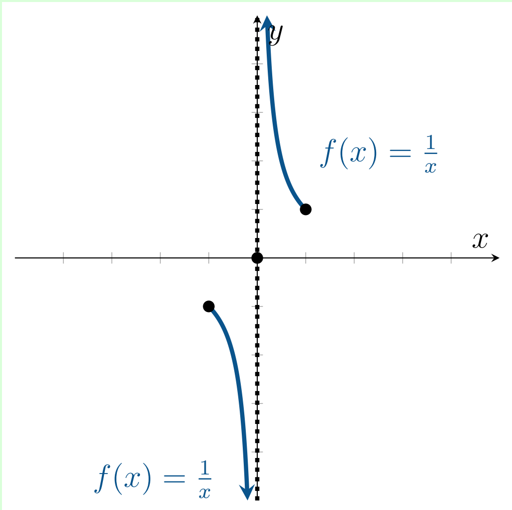


From the above example, it seems to be the case that the inclusivity of the endpoints on the interval which we examine our function of interest is important in obtaining global extrema.

Example 4.2.3 Determine the co-ordinates of any global extrema of $f(x) = \begin{cases} \frac{1}{x} & x \neq 0 \\ 0 & x = 0 \end{cases}$ on the interval $[-1, 1]$.

Solution:

We note that due to the vertical asymptote at $x = 0$, $f(x)$ zooms to ever-higher values to the right of 0 and ever-lower values to the left of 0. So, there is no global maximum and no global minimum.



While the above example remedied the inclusivity of endpoints, this was still not enough to guarantee that our function attained a global maximum and global minimum. It seems that continuity is also of importance.

We collect our observations and state the following result:

Theorem 4.2.4 **(Extreme Value Theorem (EVT))**

Assume that $f(x)$ is continuous on the closed interval $[a, b]$. Then, there exist two numbers c_1 and c_2 in the closed interval $[a, b]$ such that $f(c_1) \leq f(x) \leq f(c_2)$ for all $x \in [a, b]$. That is, there exists a global minimum at $x = c_1$ and a global maximum at $x = c_2$.

REMARKS

It is important to understand what EVT is and what it is not.

- The conditions of inclusivity of endpoints and continuity are *sufficient* but not *necessary* conditions to guarantee the existence of a global maximum and global minimum of a function on an interval. Consider for example the function $f(x) = \begin{cases} 1 & x \neq 0 \\ -1 & x = 0 \end{cases}$ on the interval $(-2, 2)$. This function on this interval meets neither the conditions of continuity nor inclusivity of endpoints, yet has a global minimum at $x = 0$ and global maxima at every point on the interval except at $x = 0$.
- The usefulness of EVT is then that if we *know* we are dealing with a continuous function on a closed interval, then we are *guaranteed* a global maximum and global minimum on the interval.
- EVT is an *existence* theorem: it does not tell us where the global extrema are or how to find them.

Alas, the entire point of optimization is to find these extrema! It seems there is more work to be done, beyond simply knowing EVT. Could it be, perhaps, that the entire preceding chapter on derivatives was more than just a creative outlet for a mathematician with too much time on their hands? That instead, derivatives might actually be of some use to us here? While this is certainly an exciting prospect, let us first introduce the other kind of extrema that we will encounter.

Definition 4.2.5
Local Extrema

Suppose $f(x)$ is a function.

- We say there is a **local minimum** at $x = c$ for $f(x)$ if there exists an open interval (a, b) containing c such that $f(c) \leq f(x)$ for all $x \in (a, b)$.
- We say there is a **local maximum** at $x = c$ for $f(x)$ if there exists an open interval (a, b) containing c such that $f(c) \geq f(x)$ for all $x \in (a, b)$.
- We say there is a **local extremum** at $x = c$ for $f(x)$ if there is either a local minimum or local maximum there.

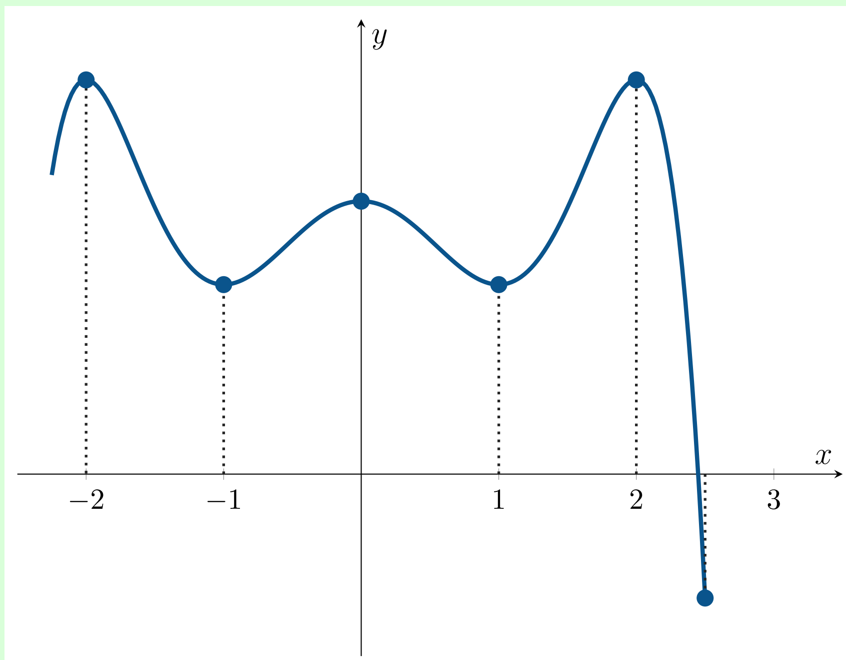
REMARKS

- We sometimes refer to the open interval (a, b) as a *neighbourhood* around $x = c$.

- Our definition above means that endpoints *cannot* be local extrema, as we cannot build a neighbourhood around them.
- You may find that some texts state that local extrema *can* occur at endpoints. In these texts, the definition for local extrema would be different than ours.
- Recall that global extrema can occur at endpoints. However, any global extrema that are not at endpoints are, for free, also local extrema.
- Any local extrema we find are candidates for (though not guaranteed to be) global extrema. Additionally, endpoints are candidates for (though not guaranteed to be) global extrema

Example 4.2.6

Find the x-values of all extrema for the function depicted on the interval $[-2.25, 2.5]$ below.



Solution:

We see that the global maxima occur at $x = -2$ and $x = 2$, while the global minimum occurs at $x = 2.5$.

There are local maxima at $x = -2$, $x = 0$, and $x = 2$, and local minima at $x = -1$ and $x = 1$.

Given that local extrema are then candidates for global extrema, it would be very useful to be able to locate local extrema. The next theorem provides some insight into this process.

Theorem 4.2.7 **(Fermat's Theorem)**

If there is a local extremum for $f(x)$ at $x = c$ and $f'(c)$ exists, then $f'(c) = 0$.

Proof:

Assume without loss of generality that there is a local minimum for $f(x)$ at $x = c$ and that $f'(c)$ exists.

Since $f'(c)$ exists, then by the definition of the derivative at a point we have that

$$f'(c) = \lim_{h\to 0} \frac{f(c+h)-f(c)}{h} = \lim_{h\to 0^+} \frac{f(c+h)-f(c)}{h} = \lim_{h\to 0^-} \frac{f(c+h)-f(c)}{h}.$$

Now, since there is a local minimum at $x = c$, by definition there exists an open interval (a, b) where $c \in (a, b)$ such that $f(c) \leq f(x)$ for all $x \in (a, b)$.

Then, for $h > 0$ small enough that $c < c + h < b$ it must be the case that $f(c) \leq f(c + h)$.

This means that

$$f'(c) = \lim_{h\to 0^+} \frac{f(c+h)-f(c)}{h} \geq 0$$

since $f(c + h) - f(c) \geq 0$ and $h > 0$.

On the other hand, for $h < 0$ small enough that $a < c + h < c$ it must be the case that $f(c) \leq f(c + h)$.

This means that

$$f'(c) = \lim_{h\to 0^-} \frac{f(c+h)-f(c)}{h} \leq 0$$

since $f(c + h) - f(c) \geq 0$ and $h < 0$.

That is, we have that

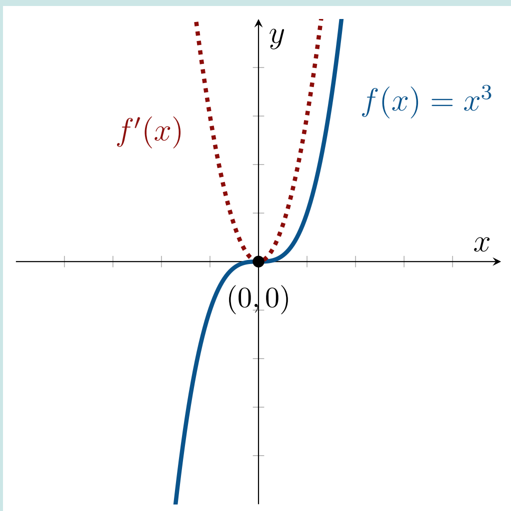
$$0 \leq f'(c) \leq 0$$

which means we must have that $f'(c) = 0$. The argument is similar for a local maximum. □

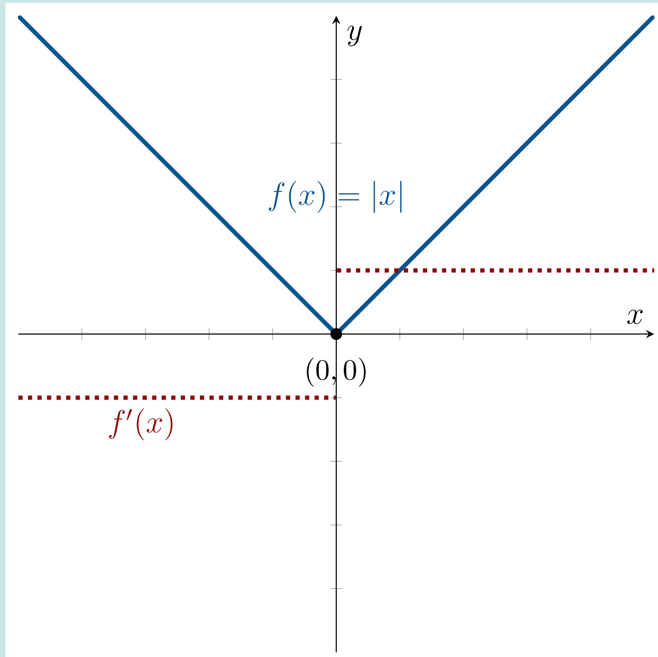
REMARKS

As with any theorem, it is vital to realize there is precision in the way that Fermat's Theorem is stated.

- The converse is false. That is, $f'(c) = 0$ does *not* imply that there is a local extremum at $x = c$. Consider for example $f(x) = x^3$, for which $f'(0) = 3(0)^2 = 0$, but has no local extremum at $x = 0$. So, places where $f'(c) = 0$ are merely *candidates* for local extrema.



- The entirety of the hypothesis is required. That is, there being a local extremum at $x = c$ does *not* imply that $f'(c) = 0$. Consider for example $f(x) = |x|$ which has a local minimum at $x = 0$ but for which $f'(0)$ does not exist. So, places where $f'(c)$ does not exist are also *candidates* for local extrema.



Notice that the result of Fermat's Theorem is that for a local extremum where the derivative exists, there is a horizontal tangent line. This was foreshadowed back in Section 3.2.

We have a special name for locations we will want to investigate for local extrema.

Definition 4.2.8
Critical Points

We say that there is a **critical point** for a function $f(x)$ at $x = c$ if c is in the domain of $f(x)$ and either $f'(c) = 0$ or $f'(c)$ does not exist. Critical points are candidates for local extrema.

REMARK

Just as with extrema, there is some ambiguity in the usage of the term critical point. It will always be clear what is desired in the context of any question you are asked.

Example 4.2.9

Find the x-values of all critical points of $f(x) = (x - 2)x^{1/3}$.

Solution:

First, note that the domain of $f(x)$ is $x \in (-\infty, \infty)$.

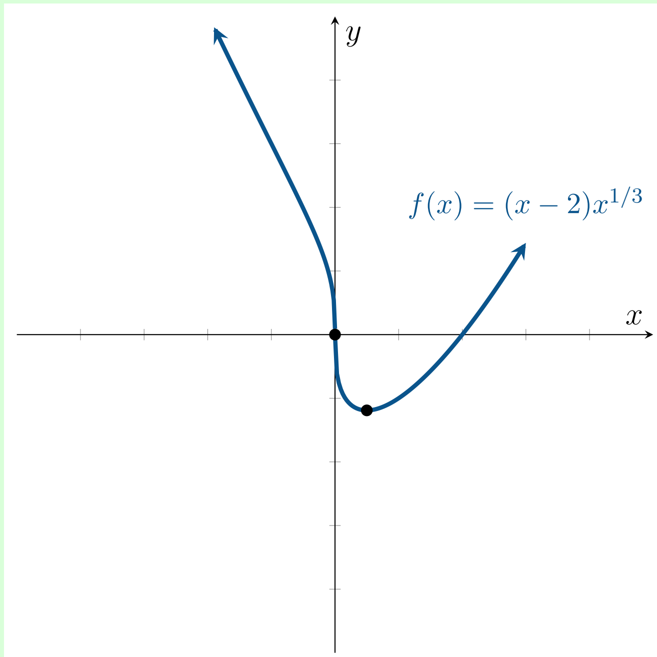
Now, via product rule, $f'(x) = (1)x^{1/3} + (x - 2)\frac{1}{3}x^{-2/3} = \frac{3x + (x - 2)}{3x^{2/3}} = \frac{4x - 2}{3x^{2/3}}$.

We can see that $f'(x) = 0$ when the numerator is zero, giving $4x - 2 = 0 \implies x = \frac{1}{2}$.

We can also see that $f'(x)$ does not exist when the denominator is zero, giving $3x^{2/3} = 0 \implies x = 0$.

So there are critical points at $x = \frac{1}{2}$ and $x = 0$.

We can see in the graph below that the former ends up being a local (and global) minimum, and the latter ends up being the location of a vertical tangent line.



We now recap several pieces of knowledge we have gained. We note that for a continuous function on a closed interval $[a, b]$, EVT guarantees that we will obtain a global minimum and a global maximum. These global extrema will either occur at endpoints, or they will occur in the open interval (a, b). If they occur in the open interval, they are also local extrema. Thus, local extrema are candidates for global extrema. Further, we know that critical points are candidates for local extrema.

This leads us to a method for finding global extrema for continuous functions on closed intervals.

4.2.2 Closed Interval Method

Method

Closed Interval Method

Given a continuous function $f(x)$ on a closed interval $[a, b]$, we can use the **closed interval method** to locate global extrema:

1. Calculate $f(a)$ and $f(b)$.
2. Find $f'(x)$.

3. Find the locations of all critical points on (a, b). That is, find points in the domain of $f(x)$ such that $f'(c) = 0$ or $f'(c)$ does not exist.
4. Calculate the function value at all points found in Step 3.
5. The global maximum occurs at the highest value from Steps 1 and 4, and the global minimum occurs at the lowest value from Steps 1 and 4.

Example 4.2.10 Find the co-ordinates of the global extrema of $f(x) = \frac{1}{3}x^3 - 3\sqrt[3]{x}$ on $[-8, 1]$.

Solution:

We first note that $f(x)$ is the sum of two continuous functions and is thus itself continuous. Then, since $f(x)$ is continuous on a closed interval, EVT is applicable and guarantees that we will find global maxima and minima.

Next, we find that $f(-8) = \frac{1}{3}(-8)^3 - 3\sqrt[3]{-8} = -\frac{494}{3}$ and $f(1) = \frac{1}{3}1^3 - 3\sqrt[3]{1} = -\frac{8}{3}$.

Taking the derivative of $f(x)$ we find that $f'(x) = x^2 - \dfrac{1}{x^{2/3}}$.

We notice that $f'(0)$ does not exist and since $x = 0$ is in the domain of $f(x)$, there is a critical point at $x = 0$. We calculate that $f(0) = \frac{1}{3}0^3 - 3\sqrt[3]{0} = 0$.

We also look to find where $f'(x) = 0$, which gives

$$0 = x^2 - \frac{1}{x^{2/3}} \implies x^2 = \frac{1}{x^{2/3}} \implies x^{8/3} = 1 \implies x = \pm 1.$$

Now, we have already considered $x = 1$, so we calculate $f(-1) = \frac{1}{3}(-1)^3 - 3\sqrt[3]{-1} = \frac{8}{3}$.

Finally, comparing the values of $f(-8)$, $f(-1)$, $f(0)$, and $f(1)$, we find that the global minimum is at $(-8, -\frac{494}{3})$ and the global maximum is at $(-1, \frac{8}{3})$.

A LOOK AHEAD

While the reader may be thrilled that they can now find global extrema on closed intervals, it is sometimes required in optimization problems to work on open intervals. In such situations, the closed interval method cannot be applied to find global extrema due to the lack of EVT's applicability. We may also on occasion in applications be interested in local extrema. In both of these circumstances, it would be necessary to be able to classify critical points as maxima, minima, or neither - a technique which we will arrive at in Section 4.7.

Section 4.2 Problems

4.2.1. Find all critical points for the following functions.

(a) $f(x) = x + \dfrac{1}{x}$.

(b) $f(x) = \dfrac{(x-1)^3}{(x+1)^4}$.

4.2.2. Find the global maximum and minimum of $f(x) = 2\cos x + \sin 2x$ for $x \in \left[0, \frac{\pi}{2}\right]$.

4.2.3. Organic waste deposited in a lake at $t = 0$ decreases the oxygen content of the water. Suppose the oxygen content is $C(t) = t^3 - 30t^2 + 6000$ ppm for $0 \le t \le 25$ days. Find the maximum and minimum oxygen content during that time.

4.2.4. A bacterial population grows according to

$$N(t) = 5000 + \frac{3000t}{100 + t^2},$$

with time t measured in hours. Determine the maximum population size for $t \ge 0$.

4.3 Mean Value Theorem

In this section, we will arrive at and prove one of the key theorems in differential calculus. This theorem is crucial not necessarily because of what it says, but because it is used in proving many further theorems which will have great use to us.

Let us start by considering a seemingly random example.

Example 4.3.1

The Evil Mathematician lives 7 km away from work. If you know that they completed their drive in to work today in 6 minutes, and that the speed limit on their route to work was 50 km/h, how can you guarantee that they sped at some point and get them in trouble?

Solution:

We start by calculating the Evil Mathematician's average velocity over the course of their trip:

$$v_{avg} = \frac{7 \text{ km}}{\frac{1}{10} \text{ h}} = 70 \text{ km/h}$$

Now, there are two possibilities for how the Evil Mathematician drove to work and averaged a velocity of 70 km/h.

The first, and perhaps more boring, possibility is that the Evil Mathematician drove at a constant 70 km/h their entire trip - thereby speeding the entire journey. Got 'em!

The alternative is that the Evil Mathematician may well have been respecting the speed limit and driving at (or below) 50 km/h for some portion of their trip. But, in order for the average velocity of the journey to reach 70 km/h, there *must* have been some portion of the journey spent at more than 70 km/h too (otherwise the average would be less than 70 km/h). It is in this latter portion where they are guaranteed to have been speeding.

In the famous words of Dora the Explorer: we did it!

Now, there is a subtle point in the previous example. As of the writing of these course notes, cars do not 'jump' around in position or velocity. Your car cannot teleport from your home garage to down the street, and it takes some time for a car to speed up from rest.

In the case of the Evil Mathematician, we already know that one option is that they travelled at the average velocity of 70 km/h for the duration of their journey to work. But in the alternative scenario, we also know that in order for the Evil Mathematician to change from driving at 50 km/h (or below) to more than 70 km/h, they must at some point have been going at *exactly* 70 km/h.

That is, at some point along the journey, the Evil Mathematician was *guaranteed to have an instantaneous velocity of exactly the average velocity*. Graphically, we see in the next diagram that this means that we are guaranteed that at some point the slope of the tangent line will equal the slope of the secant lane.

This is, in essence, what we will aim to establish mathematically. But, before we can get there, we first need to prove an intermediary theorem.

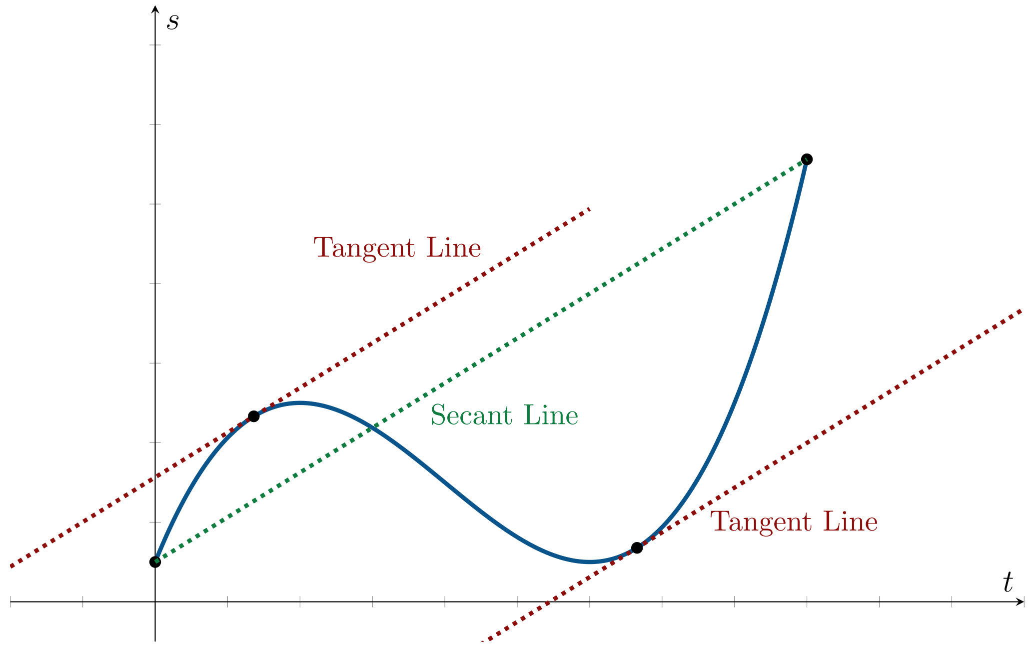


Figure 4.3.1: The average velocity is given by the slope of the secant line between the start and end of the journey depicted above. We see in this example that there are two times where the instantaneous velocity, given by the slope of the tangent line, matches the average velocity.

Theorem 4.3.2 **(Rolle's Theorem)**

Assume that $f(x)$ is continuous on the closed interval $[a, b]$ and differentiable on the open interval (a, b), and that $f(a) = f(b) = k \in \mathbb{R}$. Then, there exists a $c \in (a, b)$ such that $f'(c) = 0$.

Proof:

Assume that $f(x)$ is continuous on the closed interval $[a, b]$ and differentiable on the open interval (a, b), and that $f(a) = f(b) = k \in \mathbb{R}$. We consider three cases.

Case 1: $f(x) = k$

Here, $f'(x) = 0$. Thus, $f'(c) = 0$ for any $c \in (a, b)$. That is, we have infinitely many choices for c that satisfies the conclusion of Rolle's Theorem.

Case 2: $f(x_0) > k$ for some $x_0 \in (a, b)$

Since $f(x)$ is continuous on the closed interval $[a, b]$, we can apply the Extreme Value Theorem to $f(x)$ on $[a, b]$. EVT guarantees that $f(x)$ attains a global maximum at some $c \in [a, b]$.

But, since we know that $f(a) = f(b) = k$ and $f(x_0) > k$, the global maximum thus does not occur at the endpoints, and rather $c \in (a, b)$.

Since the global maximum occurs on the open interval (a, b) it is also a local maximum.

Finally, since $f(x)$ is differentiable on the open interval (a, b) where the local maximum at c resides, we must have that $f'(c) = 0$ by Fermat's Theorem.

Case 3: $f(x_0) < k$ for some $x_0 \in (a, b)$

The work here is identical to Case 2, except working with a global minimum instead. $\square$

REMARK

- When $k = 0$, Rolle's Theorem tells us that for a function $f(x)$ that meets its hypothesis, between the zeros of $f(x)$, there is a zero of $f'(x)$.

We are now ready to state and prove the main result of this section.

Theorem 4.3.3 **(Mean Value Theorem (MVT))**

Assume that $f(x)$ is continuous on the closed interval $[a, b]$ and differentiable on the open interval (a, b). Then there exists a $c \in (a, b)$ such that

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Proof:

Assume that $f(x)$ is continuous on the closed interval $[a, b]$ and differentiable on the open interval (a, b).

Consider the function $h(x) = f(x) - \left[f(a) + \dfrac{f(b) - f(a)}{b - a}(x - a)\right]$.

We note that $h(x)$ is continuous on the closed interval $[a, b]$ and differentiable on the open interval (a, b) since it is the difference of $f(x)$ which is continuous and differentiable on those intervals and a linear function which is continuous and differentiable everywhere.

We further note that $h(a) = f(a) - \left[f(a) + \dfrac{f(b) - f(a)}{b - a}(a - a)\right] = f(a) - f(a) = 0$ and that $h(b) = f(b) - \left[f(a) + \dfrac{f(b) - f(a)}{b - a}(b - a)\right] = f(b) - (f(a) + f(b) - f(a)) = 0$.

Thus, we can apply Rolle's Theorem to $h(x)$ on the interval $[a, b]$, which tells us that there exists a $c \in (a, b)$ such that $h'(c) = 0$.

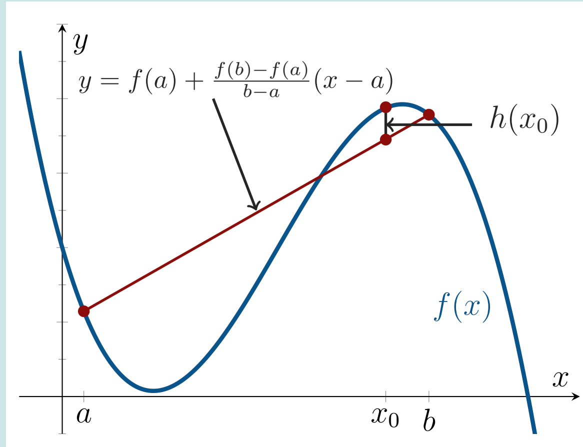
Now, we have that $h'(x) = f'(x) - \dfrac{f(b) - f(a)}{b - a}$.

So, we have that $h'(c) = f'(c) - \dfrac{f(b) - f(a)}{b - a} = 0$.

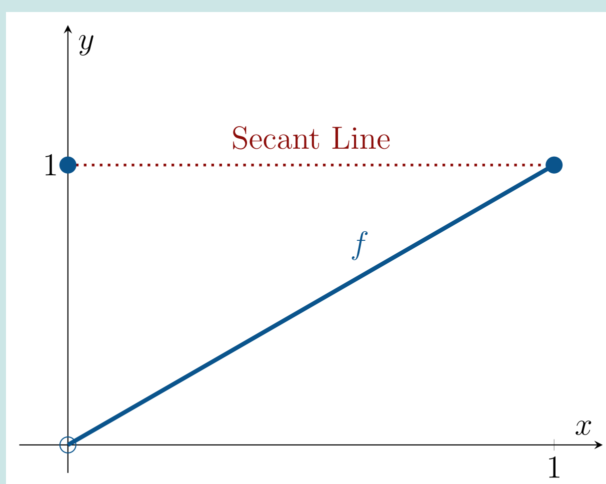
Finally, rearranging gives us that $f'(c) = \dfrac{f(b) - f(a)}{b - a}$. $\square$

REMARKS

- The function $h(x)$ introduced in the proof of MVT gives the height of the function $f(x)$ above the secant line between $(a, f(a))$ and $(b, f(b))$.



- We can think of MVT as a tilted version of Rolle's Theorem. However, Rolle's Theorem is a specific case of the more general MVT.
- Contrary to its name, MVT is a nice theorem. Here, *mean* refers to *average*, as the theorem tells us that for a function meeting its hypothesis, there will be a point where the instantaneous rate of change equals the average rate of change.
- Whenever you wish to use MVT (or Rolle's Theorem), you must always justify that the hypotheses are met. Also be wary of the various kinds of intervals in the hypothesis and conclusion of the theorem.
- MVT is another *existence* theorem. It tells you that c exists, but not where it is or how to find it.
- You may wonder why continuity is required on the stricter closed interval (beyond being a requirement for EVT, which is needed in the proof). Consider for example the function $f(x) = \begin{cases} x & x \neq 0 \\ 1 & x = 0 \end{cases}$ on the interval $[0, 1]$. You will note that the average rate of change over the interval is 0, but there is no point where the instantaneous rate of change is 0.



- You may wonder why differentiability is a requirement here (beyond being required for Fermat's Theorem, which is needed in the proof). Consider for example the function

$f(x) = |x|$ on the interval $[-1, 1]$. You will note the average rate of change over the interval is 0, but there is no point where the instantaneous rate of change is 0.

Example 4.3.4 Can MVT be applied to the following functions on the given intervals? If not, why?

(a) $f(x) = |x|$ on $x \in [-1, 1]$

(b) $g(x) = \dfrac{x+1}{x+3}$ on $x \in [-4, 0]$

(c) $h(x) = \begin{cases} x^3 & x \neq -1 \\ 1 & x = -1 \end{cases}$ on $x \in [-1, 1]$

(d) $j(x) = e^{-x}$ on $x \in [-1, 1]$

Solution:

(a) No, MVT cannot be applied. $f'(0)$ does not exist, so $f(x)$ does not meet the criterion of differentiablility on $(-1, 1)$.

(b) No, MVT cannot be applied. $g(-3)$ is undefined (there is a vertical asymptote there), so $g(x)$ does not meet the criterion of continuity on $[-4, 0]$.

(c) No, MVT cannot be applied. There is a finite jump at $x = -1$, so $h(x)$ does not meet the criterion of continuity on $[-1, 1]$.

However, do note that the average rate of change over $[-1, 1]$ is 0, and $h'(0) = 0$. So, just because MVT is not applicable does not mean there is no point where the instantaneous rate of change equals the average rate of change.

(d) Yes, MVT can be applied. $j(x)$ is continuous on $[-1, 1]$ and differentiable on $(-1, 1)$.

Example 4.3.5 Prove that the equation $\cos(x) = x$ has *at most* one solution on $[0, 1]$.

Solution:

Assume for the sake of contradiction that there exist two solutions to the equation, y and z, in $[0, 1]$, where $y \neq z$. Without loss of generality, let $y < z$.

Now, consider the helper function $f(x) = \cos(x) - x$. This means we have $f(y) = 0$ and $f(z) = 0$. Furthermore, note that $f(x)$ is continuous and differentiable everywhere, and is thus continuous on $[y, z]$ and differentiable on (y, z).

Then, by MVT, there exists a $c \in (y, z)$ such that $f'(c) = \dfrac{f(z) - f(y)}{z - y} = \dfrac{0 - 0}{z - y} = 0$.

But, notice that $f'(x) = -\sin(x) - 1$, which is strictly negative in $[0, 1]$, and thus on (y, z). Thus, we must have that $f'(c) < 0$.

This is a contradiction. Therefore, there is at most one solution to the equation $\cos(x) = x$ on $[0, 1]$.

REMARK

Way back in Example 2.10.5, we showed that the equation $\cos(x) = x$ has *at least* one solution on $(0, 1)$ using IVT. Putting that information together with the preceding example (and the fact that the endpoints of the interval are not themselves solutions), it follows that $\cos(x) = x$ has a *unique* solution on $[0, 1]$.

Section 4.3 Problems

4.3.1. The Netherlands was the first country to introduce average-speed cameras for speed-limit enforcement. The approach has subsequently spread to several European countries and through the Middle East. In the simplest implementation, two cameras are deployed a large distance apart along a highway. As cars pass by the cameras, their licence plates are recorded. By comparing the time-stamps of identifications made by each camera, the average velocity of the vehicle is determined.

(a) Prove that if the average velocity of the vehicle exceeds the posted speed limit, then the vehicle has exceeded the speed limit somewhere along the highway between the cameras.

(b) Prove that the converse is not true: although the average speed is well-below the posted speed limit, a vehicle may have exceeded the speed limit somewhere along the highway.

4.3.2. Prove that

$$|\sin b - \sin a| \le |b - a|.$$

4.3.3. The Mean-Value Theorem can be used to compare the geometric and arithmetic means.

(a) Use the MVT to show that

$$\sqrt{b} - \sqrt{a} < \frac{b-a}{2\sqrt{a}}$$

for $0 < a < b$.

(b) Use part (a) to show that, for $0 < a < b$, the geometric mean $\sqrt{ab}$ is always smaller than the arthicmetic mean $\frac{1}{2}(a + b)$, that is, show that

$$\sqrt{ab} < \frac{a+b}{2}.$$

4.3.4. Show that the function $f(x) = 2x^5 + 2x + 1$ has exactly one root without sketching the graph of the function.

Hint: assume there is more than 1 root and use the MVT to build a contradiction.

4.3.5. Let $f(x)$ be differentiable on (a, b) and $f'(x)$ be continuous on (a, b). Assume there are three points x_1, x_2 and x_3 with each $x_i \in (a, b)$ and with $x_1 < x_2 < x_3$ such that

$$f(x_1) < f(x_2) \quad \text{and} \quad f(x_2) > f(x_3)$$

Use the MVT and the IVT to show that there must be a point $c \in (x_1, x_3)$ such that $f'(c) = 0$.

4.3.6. If $|f(x) - f(y)| \le M|x - y|$ for all $x, y \in \mathbb{R}$, then we say the function f is a Lipschitz function.

(a) Use the Mean Value Theorem to prove that if f is a differentiable function and $|f'(x)| \le M$ for all $x \in \mathbb{R}$ then f is Lipschitz.

(b) Is the converse of part (a) true? Prove it or give a counterexample.

4.4 Antiderivatives

By this point in the course, you might feel rather adept at taking the derivative of nearly any function thrown your way. Perhaps you're even considering adding derivative-taking to your repertoire of party tricks to impress your non-math friends. But, what if we think about going *the other way*?

Definition 4.4.1
Antiderivative

Given a function $f(x)$, an **antiderivative** is a function $F(x)$ such that $F'(x) = f(x)$. If $F'(x) = f(x)$ for all x in an interval I, then we say that $F(x)$ is an antiderivative for $f(x)$ on I.

REMARKS

With antiderivatives, you are being asked to return a function such than when you take its derivative, you get the function that was handed to you.

Unlike the derivative however, antiderivatives are *not unique*. Consider for example being handed the function $f(x) = 5$. You would be correct in saying an antiderivative is $F(x) = 5x$. However, you'd also be correct if you came up with $F(x) = 5x - 1$ or even $F(x) = 5x + e^{123}$. You will note that in each of these three examples, $F'(x) = 5 = f(x)$.

It seems to be the case that there are *infinitely many* possible antiderivatives $F(x)$ for a given function $f(x)$, where each of these antiderivatives differs by only a constant. It turns out that we can arrive at this result via the use of Mean Value Theorem. However, just as when we wanted to prove MVT, we will first need to prove an intermediary theorem.

Theorem 4.4.2

(Constant Function Theorem (CFT))

Assume that $f'(x) = 0$ for all x in some interval I. Then there exists an $\alpha \in \mathbb{R}$ such that $f(x) = \alpha$ for all $x \in I$.

Proof:

Assume that $f'(x) = 0$ for all x in some interval I. Let $x_1, x_2 \in I$, where $x_1 \neq x_2$. Without loss of generality, let $x_1 < x_2$.

Let $f(x_1) = \alpha$.

Since $f(x)$ is differentiable on I, it is also continuous on I.

Then, since $x_1, x_2 \in I$, we can apply MVT to $f(x)$ on the interval $[x_1, x_2]$.

MVT tells us that there exists $c \in (x_1, x_2)$ such that

$$f'(c) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$

But we know that since $c \in (x_1, x_2)$ we must have that $f'(c) = 0$ by our initial assumption.

Then, since $x_1 \neq x_2$, we must have that $0 = f(x_2) - f(x_1)$.

This gives that $f(x_2) = f(x_1)$, but we know that $f(x_1) = \alpha$. So, $f(x_2) = \alpha$ too.

Since x_1 and x_2 were arbitrary, we have that $f(x) = \alpha$ for all $x \in I$. $\square$

With this result in hand, we can move forward to prove our goal.

Theorem 4.4.3 **(Antiderivative Theorem)**

Assume that $F'(x) = G'(x)$ for all x in some interval I. Then, there exists an $\alpha \in \mathbb{R}$ such that $F(x) = G(x) + \alpha$ for all $x \in I$.

Proof:

Assume that $F'(x) = G'(x)$ for all x in some interval I. Define $h(x) = F(x) - G(x)$.

Now, $h(x)$ is continuous and differentiable on I since $F(x)$ and $G(x)$ are.

Note that $h'(x) = F'(x) - G'(x)$. Since $F'(x) = G'(x)$ by our initial assumption, we then have that $h'(x) = 0$ for all $x \in I$.

Then, by CFT we have that $h(x) = \alpha$ for all $x \in I$.

Thus, we have that $F(x) - G(x) = \alpha$, which gives $F(x) = G(x) + \alpha$ for all $x \in I$. $\square$

A LOOK AHEAD

Antiderivatives are the gateway to *integral* calculus, which you will cover in MATH 138. In general, finding antiderivatives is an altogether more daunting task than finding derivatives.

Section 4.4 Problems

4.4.1. Galileo studied the dynamics of falling objects by rolling a brass ball down an inclined plane. He adjusted the spacing of bumps along the plane until the rolling ball clicked at a steady rate. He found that the spacing of the bumps were roughly a sequence of squares so that the total elapsed distance $(x - x_0)$ was related to the time between clicks t as,

$$(x - x_0) \propto t^2, \qquad \text{(Galileo's Law)}.$$

(a) Show that Galileo's Law implies that the ball experiences a constant rate of acceleration. Note that acceleration is the second derivative of position, i.e., $a(t) = x''(t)$.

(b) Conversely, suppose the ball experiences a constant rate of acceleration a,

$$\frac{d^2x}{dt^2} = a,$$

where x is the position of the ball (initially at $x(0) = x_0$ and starting with zero velocity $x'(0) = 0$). Derive Galileo's Law.

4.4.2. The anti-derivatives of the basic trigonometric functions are straightforward,

$$\frac{d}{dx}\Big(-\cos x + C\Big) = \sin x, \quad \frac{d}{dx}\Big(\sin x + C\Big) = \cos x,$$

but the anti-derivative of their quotient $\tan x$ is more complicated. Given that

$$\frac{d}{dx}\Big(-\ln\left[f(x)\right] + C\Big) = \tan x = \frac{\sin x}{\cos x},$$

determine the function $f(x)$, and thereby determine the anti-derivative of $\tan x$.

4.5 The First Derivative and the Shape of a Function

We can use MVT to prove several more theorems, all of which relate to what the first derivative of a function tells us about how the graph of the function looks.

Definition 4.5.1
Increasing and Decreasing Functions

Let I be an interval and let $x_1, x_2 \in I$ such that $x_1 < x_2$. We say that a function $f(x)$ is:

- **increasing** on I if $f(x_1) \leq f(x_2)$ for every such x_1 and x_2.
- **strictly increasing** on I if $f(x_1) < f(x_2)$ for every such x_1 and x_2.
- **decreasing** on I if $f(x_1) \geq f(x_2)$ for every such x_1 and x_2.
- **strictly decreasing** on I if $f(x_1) > f(x_2)$ for every such x_1 and x_2.

Theorem 4.5.2

(Increasing/Decreasing Function Theorem)

Let I be an interval.

1. If $f'(x) \geq 0$ for all $x \in I$, then $f(x)$ is increasing on I.
2. If $f'(x) > 0$ for all $x \in I$, then $f(x)$ is strictly increasing on I.
3. If $f'(x) \leq 0$ for all $x \in I$, then $f(x)$ is decreasing on I.
4. If $f'(x) < 0$ for all $x \in I$, then $f(x)$ is strictly decreasing on I.

Proof:

Here the proof of statement 4 is presented, and the proofs for the rest are nearly identical.

Let I be an interval and let $x_1, x_2 \in I$ such that $x_1 < x_2$. Further, assume that $f'(x) < 0$ for all $x \in I$.

Since $f'(x)$ exists for all $x \in I$, $f(x)$ is differentiable and thus also continuous on I.

By extension, since $x_1, x_2 \in I$, we have that $f(x)$ is continuous on $[x_1, x_2]$ and differentiable on (x_1, x_2).

So, we can apply MVT to $f(x)$ on the interval $[x_1, x_2]$, which tells us that there exists a $c \in (x_1, x_2)$ such that $f'(c) = \dfrac{f(x_2) - f(x_1)}{x_2 - x_1}$.

But since $f'(x) < 0$ for all $x \in I$ and we know $c \in I$, we have that $f'(c) = \dfrac{f(x_2) - f(x_1)}{x_2 - x_1} < 0$.

Now, since $x_2 > x_1$, we have $x_2 - x_1 > 0$.

Therefore, it must be that $f(x_2) - f(x_1) < 0$. Rearranging, we arrive at $f(x_1) > f(x_2)$. $\square$

REMARKS

- If a function is strictly increasing on I, this does *not* imply that $f'(x) > 0$. Consider for example $f(x) = x^3$ on $[-1, 1]$. Here, $f(x)$ is strictly increasing, but $f'(0) = 0$.

- If a function is strictly decreasing on I, this does *not* imply that $f'(x) < 0$. Consider for example $f(x) = -\sqrt[3]{x}$ on $[-1, 1]$. Here, $f(x)$ is strictly decreasing, but $f'(0)$ does not exist.
- Consider the function $f(x) = x^2$. We would say that $f(x)$ is decreasing on $(-\infty, 0]$ and increasing on $[0, \infty)$. This does *not* mean that $f(x)$ is both increasing and decreasing at $x = 0$. In fact, this does not make mathematical sense, as the concept of increase/decrease is defined over an interval, not at a point.
- It is advisable to *not* use union symbols ($\cup$) to combine multiple intervals of increase/decrease, as this can cause issues if you are not careful. Consider for example the function $f(x) = \begin{cases} x & x \leq 1 \\ x - 1 & x > 1 \end{cases}$ which is increasing on $(-\infty, 1]$ and $(1, \infty)$. It would be incorrect to say that $f(x)$ is increasing on $(-\infty, 1] \cup (1, \infty)$ as this is equivalent to $(-\infty, \infty)$, but the function is not increasing everywhere due to the discontinuity at $x = 1$.
- When writing intervals of increase/decrease, you should include an endpoint of an interval if the function is continuous at that point on that interval.

For a visual understanding, we see in Figure 4.5.2 that to the left of the vertex of the parabola, the derivative is negative (the slopes of the tangent lines are negative) and the function is [strictly] decreasing. To the right of the vertex, the derivative is positive (the slopes of the tangent lines are positive) and the function is [strictly] increasing.

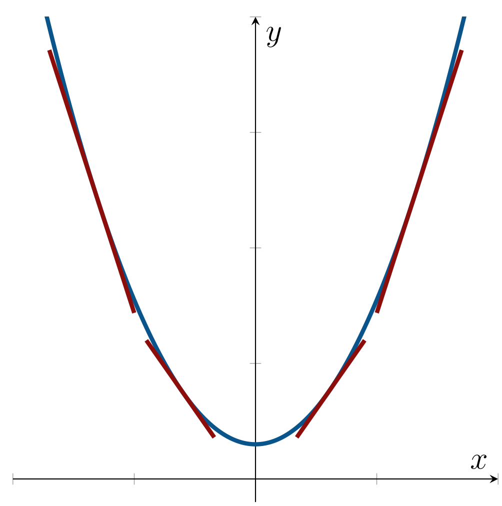


Figure 4.5.2: We see negatively sloped tangent lines to the left of the vertex where the parabola is decreasing, and positively sloped tangent lines to the right of the vertex where the parabola is increasing.

The next theorem allows us to bound our function, given bounds on the first derivative.

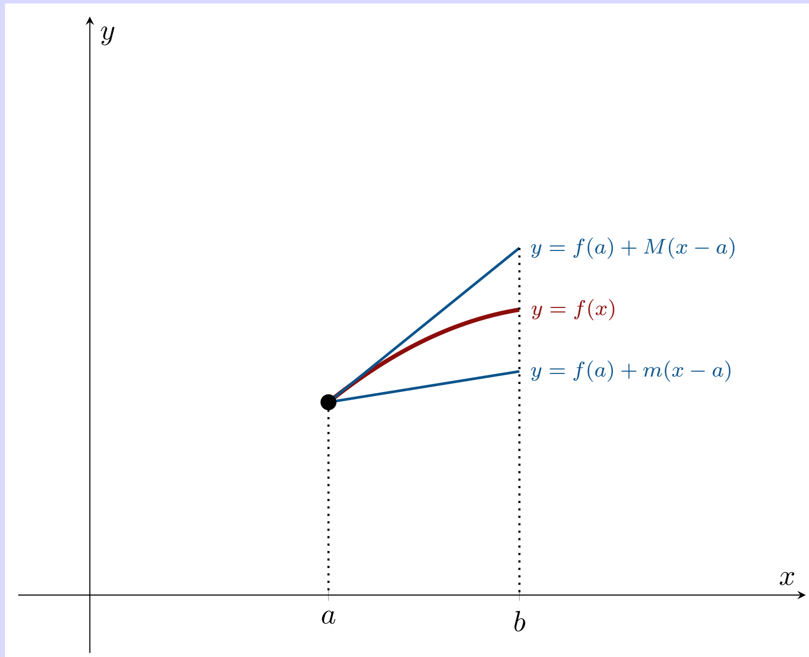
Theorem 4.5.3 **(Bounded Derivative Theorem (BDT))**

Assume that $f(x)$ is continuous on the interval $[a, b]$ and differentiable on the interval (a, b).

Further, assume that $m \leq f'(x) \leq M$ for all $x \in (a, b)$. Then,

$$f(a) + m(x - a) \leq f(x) \leq f(a) + M(x - a)$$

for all $x \in [a, b]$.



Proof:

Assume that $f(x)$ is continuous on the interval $[a, b]$ and differentiable on the interval (a, b). Further, assume that $m \leq f'(x) \leq M$ for all $x \in (a, b)$.

Given these assumptions, we know that we can apply MVT to $f(x)$ on the interval $[a, b]$. Further, MVT would be applicable to $f(x)$ on the interval $[a, x_1]$ for some $x_1 \in (a, b]$.

That is, there exists a $c \in (a, x_1)$ such that $f'(c) = \dfrac{f(x_1) - f(a)}{x_1 - a}$.

But from our assumptions, we know that $m \leq f'(x) \leq M$ for all $x \in (a, b)$, which includes at $x = c$. Thus, we have that $m \leq \dfrac{f(x_1) - f(a)}{x_1 - a} \leq M$.

Now, since $x_1 > a$ and thus $x_1 - a > 0$, we have by multiplying through by $x_1 - a$ that $m(x_1 - a) \leq f(x_1) - f(a) \leq M(x_1 - a)$ which can be further rearranged to give $f(a) + m(x_1 - a) \leq f(x) \leq f(a) + M(x_1 - a)$

Now, x_1 was chosen arbitrarily from $(a, b]$, so we have that

$$f(a) + m(x - a) \leq f(x) \leq f(a) + M(x - a)$$

for all $x \in (a, b]$.

For completeness, we note that at $x = a$ the above inequality reduces to

$$f(a) + 0 \leq f(a) \leq f(a) + 0$$

which is certainly also true. So the inequality holds for all $x \in [a, b]$. □

Example 4.5.4 If $f(12) = 2$ and $1 \leq f'(x) \leq 3$ for all $x \in \mathbb{R}$, find an interval for $f(20)$ using BDT.

Solution:

Since $f(x)$ is differentiable for all $x \in \mathbb{R}$, then it is continuous for all $x \in \mathbb{R}$.

BDT is applicable to $f(x)$ on the interval $[12, 20]$ and says

$$f(12) + 1(x - 12) \leq f(x) \leq f(12) + 3(x - 12)$$

Then, at $x = 20$ we find that $2 + 8 \leq f(20) \leq 2 + 24 \implies 10 \leq f(20) \leq 26$.

That is, BDT tells us that $f(20) \in [10, 26]$.

Example 4.5.5 Prove that $\sqrt{66} \in [8 + \frac{1}{9}, 8 + \frac{1}{8}]$ using BDT.

Solution:

Let $f(x) = \sqrt{x}$, so $f'(x) = \dfrac{1}{2\sqrt{x}}$. We note that $f(x)$ is continuous on $[64, 81]$ and differentiable on $(64, 81)$.

Here, we are using $x = 64$ and $x = 81$ as endpoints as they are the nearest perfect squares below and above 66, and thus are values for which we can evaluate $f'(x)$ easily.

Also, if $x \in [64, 81]$, we note that $f'(x) = \dfrac{1}{2\sqrt{x}} \in \left[\frac{1}{18}, \frac{1}{16}\right]$

BDT then tells us that $\sqrt{64} + \frac{1}{18}(x - 64) \leq \sqrt{x} \leq \sqrt{64} + \frac{1}{16}(x - 64)$.

Then, at $x = 66$ we find that $\sqrt{64} + \frac{1}{18}(2) \leq \sqrt{66} \leq \sqrt{64} + \frac{1}{16}(2) \implies 8 + \frac{1}{9} \leq \sqrt{66} \leq 8 + \frac{1}{8}$.

Our final theorem of this section allows us to compare two functions which meet at a point if we have knowledge of their first derivatives.

Theorem 4.5.6 **(Comparison via First Derivatives Theorem)**

Assume that $f(x)$ and $g(x)$ are continuous at $x = a$ with $f(a) = g(a)$.

1. If both $f(x)$ and $g(x)$ are differentiable for $x > a$ and if $f'(x) \leq g'(x)$ for all $x > a$, then $f(x) \leq g(x)$ for all $x > a$.
2. If both $f(x)$ and $g(x)$ are differentiable for $x < a$ and if $f'(x) \leq g'(x)$ for all $x < a$, then $f(x) \geq g(x)$ for all $x < a$.

Proof:

Here the proof of statement 2 is presented, and the proof for statement 1 is nearly identical.

Assume that $f(x)$ and $g(x)$ are continuous at $x = a$ with $f(a) = g(a)$. Further, assume that $f(x)$ and $g(x)$ are differentiable for $x < a$ and that $f'(x) \leq g'(x)$ for all $x < a$.

Let $h(x) = f(x) - g(x)$.

We note that $h(x)$ is continuous at $x = a$ and differentiable (and thus continuous) for $x < a$ since it is the difference of two such functions. Further, we note that $h(a) = 0$ since $f(a) = g(a)$, and that $h'(x) = f'(x) - g'(x) \leq 0$ for all $x < a$ since $f'(x) \leq g'(x)$ for all $x < a$.

MVT is applicable to $h(x)$ on the interval $[x, a]$, which gives that there exists $c \in (x, a)$ such that $h'(c) = \dfrac{h(a) - h(x)}{a - x}$.

Then, since $c < a$, we have that $h'(c) = \dfrac{h(a) - h(x)}{a - x} \leq 0$.

Now, since $x < a \implies a - x > 0$ and $h(a) = 0$, it must be that $h(x) \geq 0$ for all $x < a$.

That is, $f(x) - g(x) \geq 0 \implies f(x) \geq g(x)$ for all $x < a$. □

REMARK

If instead we have the strict inequality $f'(x) < g'(x)$, statement 1 would give that $f(x) < g(x)$ for all $x > a$ and statement 2 would give that $f(x) > g(x)$ for all $x < a$.

A picture to provide intuition behind this theorem is given below.

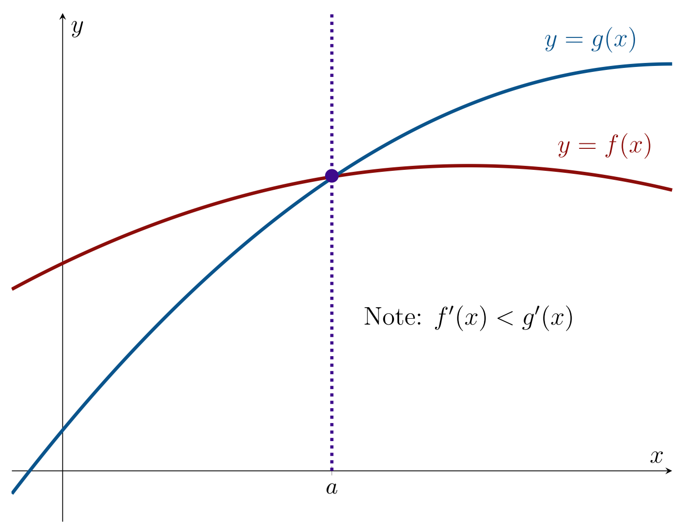


Figure 4.5.3: Here, we see that for $f'(x) < g'(x)$, we have that when $x > a$, $g(x)$ lies above $f(x)$, but when $x < a$, $f(x)$ lies above $g(x)$.

Example 4.5.7

Prove that $x^2 > \ln(1 + x^2)$ for $x < 0$.

Solution:

Let $f(x) = x^2$ and $g(x) = \ln(1 + x^2)$. We know from prior knowledge that these functions are continuous everywhere (as $f(x)$ is a polynomial and $g(x)$ is a composition of continuous functions).

We note that $f(0) = g(0) = 0$.

Taking derivatives, we find that $f'(x) = 2x$ and $g'(x) = \dfrac{2x}{1+x^2}$. We note that these exist everywhere.

Now, for $x < 0$, we have that $1 + x^2 > 1$.

Then, for $x < 0$, we find that $\dfrac{2x}{1+x^2} > \dfrac{2x}{1}$ (if you are confused about the direction of the inequality, note that both sides are negative since $x < 0$).

That is, for $x < 0$, we have that $f'(x) < g'(x)$.

Then, by Comparison via First Derivatives Theorem, we have that $f(x) > g(x)$ or $x^2 > \ln(1 + x^2)$ for $x < 0$.

Section 4.5 Problems

4.5.1. Find the intervals over which the following functions are increasing/decreasing.

(a) $f(x) = x^4 - 8x^2$

(b) $f(x) = \dfrac{1}{x^2 - 1}$

(c) $f(x) = e^x + e^{-x+1}$

(d) $f(x) = x^4 - 4x^3 + 16x - 7$

4.5.2. Show that if f is increasing and differentiable on (a, b) then $f'(x) \geq 0$ for all $x \in (a, b)$.
Hint: You may wish to use the result:

$$\text{If } g(x) > 0 \text{ for all } x \neq a \text{ and } \lim_{x\to a} g(x) = L \text{ then } L \geq 0.$$

4.6 The Second Derivative and the Shape of a Function

Having solidified our understanding of how the first derivative relates to the increase/decrease of a function, we now examine what the second derivative tells about the shape of a function.

First, some terminology.

Examining the curves in Figure 4.6.4 below, we note that for the curve on the left, the secant line joining any two points lies above the curve. For the curve on the right, the secant line joining any two points lies below the curve.

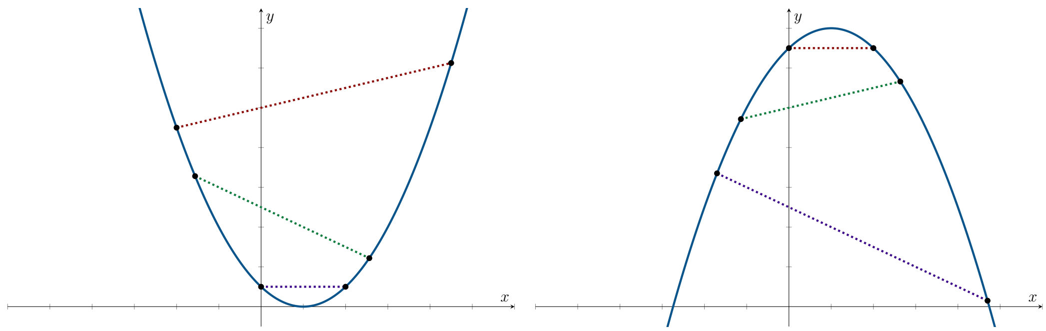


Figure 4.6.4: The function on the left is concave up, as the secant line joining any two points on the curve always lies above the curve. The function on the right is concave down, as the secant line joining any two points on the curve always lies below the curve.

Definition 4.6.1

Concave Up & Concave Down

A function $f(x)$ is **concave up** on an interval I if for all $a, b \in I$ the secant line joining $(a, f(a))$ to $(b, f(b))$ sits above the graph of $f(x)$.

A function $f(x)$ is **concave down** on an interval I if for all $a, b \in I$ the secant line joining $(a, f(a))$ to $(b, f(b))$ sits below the graph of $f(x)$.

REMARKS

Be aware that there are different ways to define concavity. By our definition, a line is *neither* concave up *nor* concave down, as any secant line would lie directly on top of the line. Further, consider that our definition would tell us that $f(x) = |x|$ is neither concave up nor concave down on $[-1, 1]$. This is since any secant line in the interval with $x = 0$ as an endpoint would lie directly on the graph of the function.

Now, bringing ourselves to the second derivative, we remind ourselves what it actually *is*: it is the rate of change of the rate of change. That is, the second derivative tells us how the slopes of the tangent lines to a function are changing.

Examining the curves in Figure 4.6.5 below, we note that for the curve on the left that as we move our eyes from left to right, the tangent line slopes move from steeply negative, to shallow and negative, to zero, to shallow and positive, to steeply positive. That is, the value of the first derivative is increasing. Looking at the curve on the right, we see that the value of the first derivative is decreasing.

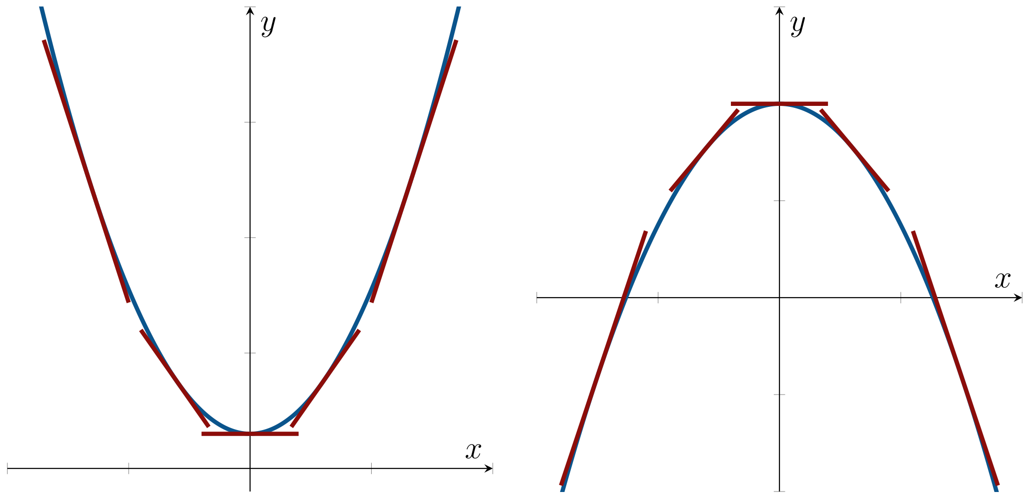


Figure 4.6.5: For the concave up function on the left, we see that the value of the first derivative (the slopes of the tangent lines) is increasing. For the concave down function on the right, we see that the value of the first derivative is decreasing.

This idea leads to the following theorem:

Theorem 4.6.2 **(The Second Derivative and Concavity)**

1. If $f''(x) > 0$ for all x in an interval I then $f(x)$ is concave up on I.
2. If $f''(x) < 0$ for all x in an interval I then $f(x)$ is concave down on I.

Now, it will be of particular interest to us to note where concavity *changes*. We have a name for such points.

Definition 4.6.3
Points of Inflection (POIs)

A point $(c, f(c))$ is called a **point of inflection** (or **POI** for short) of $f(x)$ if $f(x)$ is continuous at $x = c$ and the concavity of $f(x)$ changes at $x = c$.

Now, points of inflection will usually occur when $f''(x)$ changes sign at $x = c$. If we happen to know that $f''(x)$ is continuous at $x = c$, our old friend IVT would require that $f''(c) = 0$.

Theorem 4.6.4 **(Test for Points of Inflection)**

If $f''(x)$ is continuous at $x = c$ and $(c, f(c))$ is a point of inflection of $f(x)$, then $f''(c) = 0$.

REMARKS

As with many other theorems we have seen, you must proceed with caution.

- The converse is false. That is, $f''(c) = 0$ does *not* imply that there is a point of inflection at $(c, f(c))$. Consider for example $f(x) = x^4$, for which $f''(0) = 0$ but is concave up everywhere.

- Points of inflection do not *only* occur where $f''(c) = 0$. Consider for example $f(x) = \sqrt[3]{x}$ which has a point of inflection when $x = 0$, but for which $f''(0)$ does not exist. So, we will be interested in checking where the second derivative doesn't exist as well as where it is zero.
- Just because $f''(c) = 0$ or doesn't exist does *not* mean there is a point of inflection at $(c, f(c))$. These are just *candidates* for points of inflection. To confirm whether or not they are POIs, you must test if concavity *changes* at $x = c$.

Example 4.6.5 Find the intervals of concavity and any points of inflection of $f(x) = x^4 - 6x^2$.

Solution:

An excellent place to start would be to find the second derivative.

We have that $f'(x) = 4x^3 - 12x$ and thus $f''(x) = 12x^2 - 12$.

We note that $f''(x)$ is a polynomial so it exists everywhere. Thus, we are only concerned about where $f''(x) = 0$.

Solving $12x^2 - 12 = 0$ gives $x = \pm 1$. Again, since $f''(x)$ is a polynomial, it is continuous at these values.

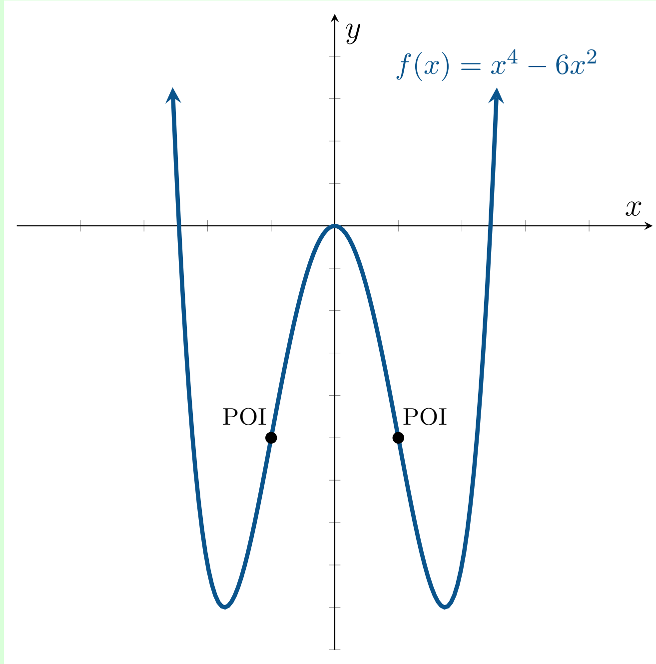
To organize our information, we draw a number line delineated by these candidate values, and then choose test values in each of the three sections (for example -2, 0, and 2) to plug into $f''(x)$. The only thing we care about is whether $f''(x)$ at our test values is positive or negative. We then denote the concavity of $f(x)$ based on this information in each of the three sections.

x		-1		1	
$f''(x)$	$+$	0	$-$	0	$+$
$f(x)$	concave up		concave down		concave up

We see that $f(x)$ is concave up on $(-\infty, -1]$ and $[1, \infty)$ and that $f(x)$ is concave down on $[-1, 1]$.

Finally, we note that $f(x)$ does in fact change concavity at $x = \pm 1$ while being continuous there, so there are points of inflection at $(-1, -5)$ and $(1, -5)$.

The function is pictured below, with our points of inflection indicated.



REMARKS

In the previous example, you will notice that when we listed the intervals of concavity, these intervals were *inclusive* of $x = \pm 1$. You should include an endpoint when listing intervals of concavity if your function is continuous there.

You will also notice that each of $x = \pm 1$ was included in both a concave up interval *and* a concave down interval. It is important to understand that this is *not* saying that $f(x)$ is both concave up and concave down at $x = \pm 1$. In fact, that statement does not make mathematical sense, as concavity is defined over an interval, not at a point.

Similar to intervals of increase and decrease, it is advisable to *not* use union symbols ($\cup$) to combine multiple intervals of concavity, as this can cause issues if you are not careful. Consider for example the function $f(x) = x^{2/3}$ which is concave down on $(-\infty, 0]$ and $[0, \infty)$. It would be incorrect to say that $f(x)$ is concave down on $(-\infty, 0] \cup [0, \infty)$ as this is equivalent to $(-\infty, \infty)$, but the function is not concave down across the cusp at $x = 0$.

Example 4.6.6 Find the intervals of concavity and any points of inflection of $f(x) = -\dfrac{1}{x}$.

Solution:

Again, we begin by finding the second derivative.

We find that $f'(x) = \dfrac{1}{x^2}$ and thus $f''(x) = -\dfrac{2}{x^3}$.

Now, we note that $f''(x) \neq 0$ as the numerator is non-zero. However, we notice that $f''(x)$ does not exist for $x = 0$, so we should investigate concavity on either side.

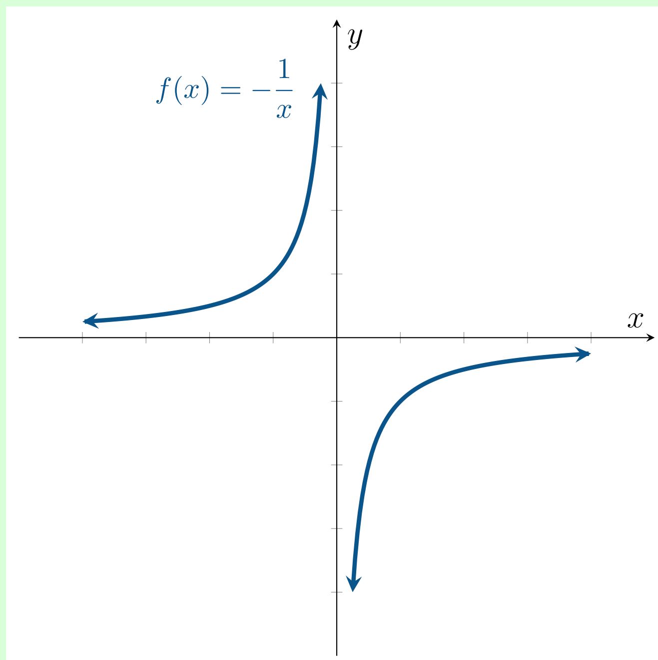
We again organize our information using a number line delineated into two sections by our sole point of interest. We test values in each section (for example -1 and 1), determine the sign of $f''(x)$ at those test values, and translate that into information about the concavity of $f(x)$ in each section.

x		0	
$f''(x)$	$+$	DNE	$-$
$f(x)$	concave up		concave down

We see that $f(x)$ is concave up on $(-\infty, 0)$ and that $f(x)$ is concave down on $(0, \infty)$.

Now, although concavity does change at $x = 0$, we notice that $f(x)$ is not continuous there, so there is *not* a point of inflection there. Overall, this function has no points of inflection.

The function is pictured below.



Section 4.6 Problems

4.6.1. For each of the following functions find the intervals where the function is concave up or concave down and state any points of inflection.

(a) $f(x) = x^4 + 6x^3 - 60x^2 + 6x - 60$

(b) $f(x) = 1 + \frac{1}{x} - \frac{1}{x^3}$

(c) $f(x) = \frac{x^2}{x + x^3}$

(d) $f(x) = \dfrac{\sin(x)}{1+\cos(x)}$

(e) $f(x) = x^2 - 9x^{\frac{1}{3}}$

(f) $f(x) = x\ln(x^2 - 2x)$

4.6.2. Consider the function

$$f(x) = \begin{cases} \dfrac{x^3}{|x|} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

(a) Compute $f'(x)$ for all possible x. Your answer can be piecewise if necessary.

(b) Compute $f''(x)$ for all possible x. Your answer can be piecewise if necessary.

(c) Does f have an inflection point at $x = 0$?

4.6.3. Let f and g be twice differentiable functions (that is, both $f''(x)$ and $g''(x)$ exist for all $x \in \mathbb{R}$). Prove or disprove (that is, find a counterexample to) the following statements:

(a) If the graph of f is concave up on $\mathbb{R}$, the graph of g is concave down on $\mathbb{R}$, and $f(x) > g(x)$ for all $x \in \mathbb{R}$, then the graph of $f + g$ is concave up on $\mathbb{R}$.

(b) If the graph of f is concave up on $\mathbb{R}$, and the graph of g is concave down on $\mathbb{R}$, then the graph of $f + g$ is neither concave up nor concave down on all of $\mathbb{R}$.

4.6.4. Prove that if f is a twice differentiable function such that $f''(x) \neq 0$ for all $x \in \mathbb{R}$, f is positive on $\mathbb{R}$, and the graph of f is concave up on $\mathbb{R}$, then the graph of $g = (f)^2$ is concave up on $\mathbb{R}$.

4.7 Classifying Critical Points

We learned in Section 4.2 that critical points are candidates for local extrema. However, at that point, we did not have the tools to determine whether each such critical point was a local maximum, local minimum, or neither.

Now that we understand what the first and second derivatives tell us about the shape of a function, we have not one, but two tools at our disposal for critical point classification. Each tool has its pros and cons. In many situations, you will be able to choose whichever of the two you prefer. Sometimes, you will be forced into using one over the other.

Let us first think about what the information the first derivative can provide us. Examining the curves in Figure 4.7.6 below, we note that for the curve on the left depicting a local maximum, the first derivative is non-negative to the left of the critical point, and non-positive to the right. That is, the function switches from increasing to decreasing through the critical point. For the curve on the right depicting a local minimum, the first derivative is non-positive to the left of the critical point, and non-negative to the right. That is, the function switches from decreasing to increasing through the critical point.

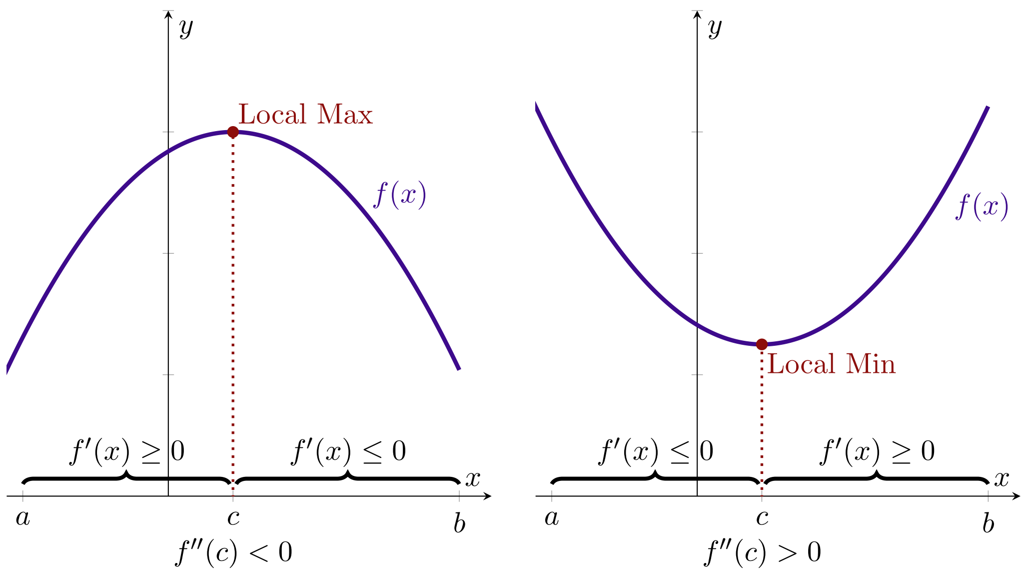


Figure 4.7.6: For the local maximum in the figure on the left, we see that the first derivative is non-negative to the left of the critical point and non-positive to the right. Also, notice that the second derivative is negative at the critical point. For the local minimum in the figure on the right, we see that the first derivative is non-positive to the left of the critical point and non-negative to the right. Also, notice that the second derivative is positive at the critical point.

Theorem 4.7.1 **(First Derivative Test (FDT))**

Assume that there is a critical point for $f(x)$ at $x = c$, and that $f(x)$ is continuous at $x = c$. If there is an open interval (a, b) containing c such that

1. $f'(x) \geq 0$ for all $x \in (a, c)$ and $f'(x) \leq 0$ for all $x \in (c, b)$ then there is a local maximum for $f(x)$ at $x = c$.
2. $f'(x) \leq 0$ for all $x \in (a, c)$ and $f'(x) \geq 0$ for all $x \in (c, b)$ then there is a local minimum for $f(x)$ at $x = c$.

Otherwise, the critical point at $x = c$ is neither a local maximum nor a local minimum.

Now, let us consider what the information the second derivative can provide us. Examining again the curves in Figure 4.7.6, we note that for the curve on the left depicting a local maximum, the second derivative is negative at the critical point. That is, the function is concave down around the critical point. For the curve on the right depicting a local minimum, the second derivative positive at the critical point. That is, the function is concave up around the critical point.

Theorem 4.7.2 **(Second Derivative Test (SDT))**

Assume that $f'(c) = 0$, and that $f''(x)$ is continuous at $x = c$.

1. If $f''(c) < 0$ then there is a local maximum for $f(x)$ at $x = c$.
2. If $f''(c) > 0$ then there is a local minimum for $f(x)$ at $x = c$.
3. If $f''(c) = 0$ then we have no information. That is, there might be a local maximum, local minimum, or neither at $x = c$.

REMARKS

It is worth highlighting now the pros and cons of FDT and SDT so you can make an informed decision when you are classifying critical points.

- FDT can be used to conclusively classify any critical points. SDT cannot be used to classify critical points that occur where $f'(x)$ does not exist, nor those where $f'(x) = 0$ but $f''(x)$ is discontinuous.
- Even for critical points where SDT is applicable, sometimes SDT gives us no information. In these scenarios, we would have to default back to FDT.
- However, with FDT, we must test to either side of each critical point, whereas with SDT we must test *at* each critical point. So, SDT requires less testing than FDT.

A LOOK AHEAD

Another advantage of SDT is that it generalizes more easily to the multivariable setting than FDT does. You will have this to look forward to when you take MATH 237!

Example 4.7.3

Find the x-values of, and classify, the local extrema of $f(x) = \dfrac{1}{3}x^3 - \dfrac{3}{2}x^2 + 2x + 1$ using FDT and SDT.

Solution:

Given the names of our tests, a good place to start would be to find the first and second derivatives.

We have that $f'(x) = x^2 - 3x + 2$ and $f''(x) = 2x - 3$.

Now, we seek to find our critical points. We note that $f'(x)$ is a polynomial, so it exists everywhere. Thus, we seek only to find where $x^2 - 3x + 2 = 0$. This gives $(x-2)(x-1) = 0$, or $x = 1, 2$.

We will first employ FDT. We draw a number line delineated into three sections by our critical points and choose one test value for $f'(x)$ in each section (for example 0, $\frac{1}{2}$ and 3). The only thing we care about is the sign of the first derivative at each test value. We then relate this information back to the shape of $f(x)$.

x		1		2	
$f'(x)$	$+$	0	$-$	0	$+$
$f(x)$	increasing		decreasing		increasing

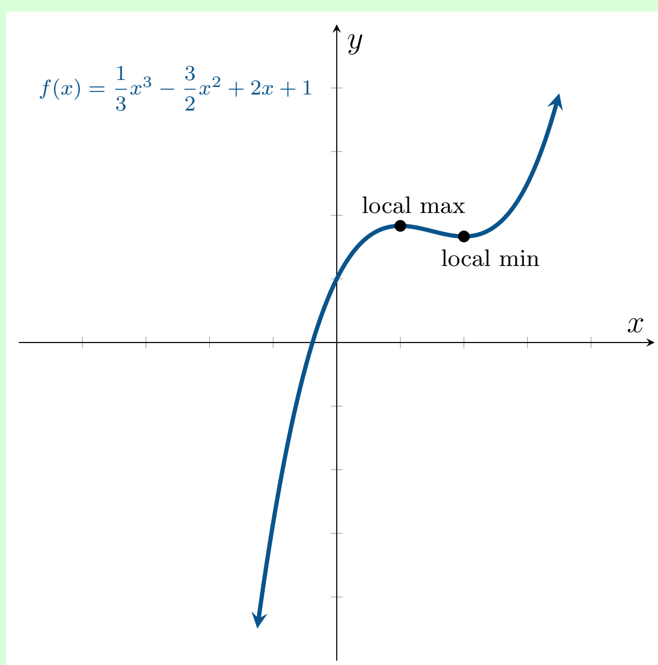
We see then by FDT that there is a local maximum at $x = 1$ and a local minimum at $x = 2$.

Next, we use SDT. We draw a number line delineated by our critical points and use the critical points themselves as our test values for $f''(x)$. The only thing we care about is the sign of the second derivative at each test value. We then relate this information back to the shape of $f(x)$.

x	1	2
$f''(x)$	$-$	$+$
$f(x)$	concave down	concave up

Our results from SDT mirror those from FDT: there is a local maximum at $x = 1$ and a local minimum at $x = 2$.

The function is pictured below, with these extrema indicated.



Example 4.7.4 Find the co-ordinates of, and classify, all extrema of $f(x) = x\sqrt[3]{4-x}$ on the interval $[0,5]$ given that $f'(x) = \dfrac{4(3-x)}{3(4-x)^{2/3}}$ and $f''(x) = \dfrac{4(x-6)}{9(4-x)^{5/3}}$.

Solution:

We notice some immediate differences in this example from the previous one. Here, we are being asked for *all* extrema, so we will want to find both global and local extrema. We also note that we have been given an interval, so we should remember that endpoints are candidates for global (but not local) extrema.

Let us start then by calculating that $f(0) = 0$ and $f(5) = -5$.

Next, we should identify our critical points. We note that $x = 4$ is a point in the domain of $f(x)$ at which $f'(x)$ does not exist (as then the denominator is zero), so this is the location of a critical point. Further, we note that $f'(x) = 0$ when $x = 3$ (as then the numerator is zero), so this is also the location of a critical point.

Now, we should work to classify these critical points. We notice however, that SDT is not applicable to the critical point at $x = 4$, as this occurs where $f'(x)$ does not exist, rather than where $f'(x) = 0$. So, we will proceed with FDT.

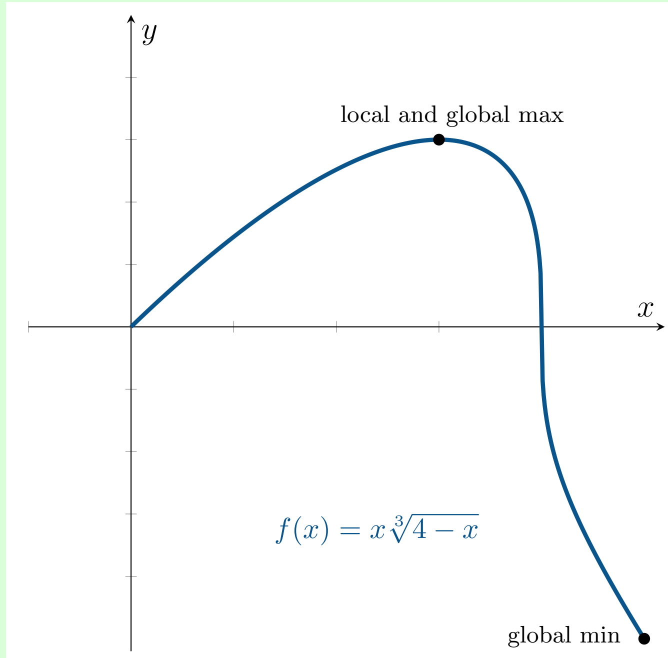
x	0		3		4		5
$f'(x)$		$+$	0	$-$	DNE	$-$	
$f(x)$		increasing		decreasing		decreasing	

We see then that FDT indicates that there is a local maximum at $x = 3$, and that the critical point at $x = 4$ is neither a local maximum nor a local minimum. For sanity, we also note that $f''(3) < 0$, so SDT returns the same result for the critical point at $x = 3$ as $f(x)$ is concave down around this point.

In order to determine our global extrema, we will need to calculate that $f(3) = 3$.

We compare this to the fact that $f(0) = 0$ and $f(5) = -5$ to realize that the global maximum is at $(3,3)$ and the global minimum is at $(5,-5)$. For completeness, we reiterate that $(3,3)$ is also a local maximum. There are no local minima.

The function is pictured below, with these extrema indicated.



Section 4.7 Problems

4.7.1. Locate and classify the critical points of each of the following functions:

(a) $f(x) = 3x^5 - 20x^3 + 15$

(b) $f(x) = \dfrac{x+1}{x-1} - x$

(c) $f(x) = \dfrac{x+1}{x-1} + x$

(d) $f(x) = 12x^{\frac{1}{3}} + |x|$

(e) $f(x) = 4e^x - e^{2x} - x$

(f) $f(x) = \ln(x+1) + x^2 - 6x$

(g) $f(x) = |x| + \sin^2(x)$

4.7.2. Let $f(x) = ax^3 + bx^2 + cx + d$ be a cubic equation.

(a) Find $a, b, c, d \in \mathbb{R}$ such that $f(x)$ has a local maximum at $(-2, 3)$ and local minimum at $(1, 0)$.

(b) Use (a) to find the intervals of concavity and the inflection points of $f(x)$.

4.8 L'Hôpital's Rule

You may remember when you were a young(er) MATH 137 student, bright-eyed and excited. Back then, you spent a great deal of time (two whole chapters worth) learning about limits. You worked hard to understand how we describe the concept of being infinitely close, mathematically, through the use of epsilon. Then, you were finally grown up, and ready to move on to learn about derivatives! It is only natural, then, to drag you back into the world of limits.

You likely find yourself asking 'why now?' and thinking to yourself 'not again!' There is a good reason we are here again: we actually needed to first understand derivatives in order to work with the limit technique of this section, and its proof (beyond the scope of this course) relies upon MVT. In fact, there were many limits we purposely avoided dealing with in the earlier part of the course. Additionally, there were even some limits we previously went to great lengths to solve, which will shortly fall apart in just a few quick lines.

Before we get into the mathematics, a word of caution and a historical note.

The rule you are about to learn holds great power. In fact, so much power that you (as many a calculus student before you) may find it irresistible to try and use it all the time. However, you should be wary of when you can, and more importantly when you can not, use it.

Now, given the name of this section, you might naturally assume that L'Hôpital's Rule was theorized by French mathematician Guillaume De l'Hôpital. Unfortunately, this is not the case. The rule was actually discovered by Swiss mathematician Johann Bernoulli. The story goes that L'Hôpital had an agreement where he paid Bernoulli in exchange for credit for any findings that the latter made. It seems certain that this contract pushed the *limits* of reasonability.

When dealing with a limit problem, our first tactic is typically to 'plug in' the value we are approaching. Once we do this, there are certain forms that require special attention.

Definition 4.8.1

Indeterminate Forms

The following seven types of limit are a complete list of **indeterminate forms**:

$$\frac{0}{0}, \quad \pm\frac{\infty}{\infty}, \quad 0 \times \pm\infty, \quad \infty - \infty, \quad 1^{\infty}, \quad \infty^{0}, \quad 0^{0}$$

REMARKS

- Indeterminate forms are so-called because we cannot determine them just by looking at them, and require specialized manipulation to evaluate.
- The list above are the *only* indeterminate forms. You may run into forms that look a lot like they may be indeterminate, but are in fact *determinate*. Some such examples are:
  - $\infty \times \infty$ evaluates to ∞
  - $\dfrac{\infty}{0}$ does not exist (it evaluates to ∞ if we have 0^{+} and to $-\infty$ if we have 0^{-})
  - 0^{∞} evaluates to 0

To handle these indeterminate forms, here is our central tool:

Theorem 4.8.2 **(L'Hôpital's Rule (LHR))**

Assume that $f'(x)$ and $g'(x)$ exist near $x = a$, except possibly at $x = a$, and that $g'(x) \neq 0$ near $x = a$, except possibly at $x = a$. Further, assume that $\lim_{x \to a} \frac{f(x)}{g(x)}$ is of indeterminate form $\frac{0}{0}$ or $\pm\frac{\infty}{\infty}$. Then,

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)}$$

if the latter limit exists or is $\pm\infty$.

REMARKS

- Notice that L'Hôpital's Rule is only applicable directly to $\frac{0}{0}$ or $\pm\frac{\infty}{\infty}$ indeterminate forms.
- Be aware that you are not using quotient rule, but are rather taking the derivative of the top and bottom functions independently.
- L'Hôpital's Rule applies for $a \in \mathbb{R}$, when a is $\pm\infty$, and for one-sided limits.
- You may have to repeatedly apply L'Hôpital's Rule.
- You must use the notation $\stackrel{\text{LHR}}{=}$ to indicate where exactly L'Hôpital's Rule has been applied in a limit calculation.

Now, let's examine how to apply LHR to the different indeterminate forms described above.

4.8.1 Type $\frac{0}{0}$

With this type, we can apply LHR directly.

Example 4.8.3 Evaluate $\lim_{x \to 2} \dfrac{2-x}{\sqrt{2}-\sqrt{x}}$.

Solution:

$$\begin{aligned}
&\lim_{x \to 2} \frac{2-x}{\sqrt{2}-\sqrt{x}} && \text{(type } \tfrac{0}{0}\text{)} \\
&\stackrel{\text{LHR}}{=} \lim_{x \to 2} \frac{-1}{\left(\frac{-1}{2\sqrt{x}}\right)} \\
&= \lim_{x \to 2} 2\sqrt{x} \\
&= 2\sqrt{2}
\end{aligned}$$

Example 4.8.4 Evaluate $\lim_{x\to 0}\frac{\sin(x)}{x}$.

Solution:

We recognize that this is the Fundamental Trig Limit, which we used geometry and Squeeze Theorem to solve in Section 2.4. Now, let's use LHR instead.

$$\begin{aligned}
&\lim_{x\to 0}\frac{\sin(x)}{x} && \left(\text{type } \tfrac{0}{0}\right)\\
&\overset{\text{LHR}}{=}\lim_{x\to 0}\frac{\cos(x)}{1}\\
&=1
\end{aligned}$$

This was much quicker! However, it is worth noting that there is some circular logic at work here. To use LHR, we had to take the derivative of $\sin(x)$, but you may remember we derived that derivative from first principles by making use of the Fundamental Trig Limit.

Example 4.8.5 Evaluate $\lim_{x\to 0}\frac{e^x-x-1}{5\cos(x)+x^2-5}$.

Solution:

$$\begin{aligned}
&\lim_{x\to 0}\frac{e^x-x-1}{5\cos(x)+x^2-5} && \left(\text{type } \tfrac{0}{0}\right)\\
&\overset{\text{LHR}}{=}\lim_{x\to 0}\frac{e^x-1}{-5\sin(x)+2x} && \left(\text{type } \tfrac{0}{0}\right)\\
&\overset{\text{LHR}}{=}\lim_{x\to 0}\frac{e^x}{-5\cos(x)+2}\\
&=-\frac{1}{3}
\end{aligned}$$

4.8.2 Type $\pm\frac{\infty}{\infty}$

Again, we can apply LHR directly to this type.

Example 4.8.6 Evaluate $\lim_{x\to\infty}\frac{x^3-2x+7}{2x^3+x^2+x+1}$.

Solution:

$$\begin{aligned}
&\lim_{x\to\infty}\frac{x^3-2x+7}{2x^3+x^2+x+1} && \left(\text{type } \tfrac{\infty}{\infty}\right)\\
&\overset{\text{LHR}}{=}\lim_{x\to\infty}\frac{3x^2-2}{6x^2+2x+1} && \left(\text{type } \tfrac{\infty}{\infty}\right)\\
&\overset{\text{LHR}}{=}\lim_{x\to\infty}\frac{6x}{12x+2} && \left(\text{type } \tfrac{\infty}{\infty}\right)
\end{aligned}$$

$$\begin{aligned} &\overset{\text{LHR}}{=} \lim_{x\to\infty} \frac{6}{12} \\ &= \frac{1}{2} \end{aligned}$$

While LHR worked here, it is arguably faster to use the dominating power technique we learned earlier in the course for rational function limits.

Example 4.8.7 Evaluate $\displaystyle\lim_{x\to\infty} \frac{\ln(x)}{x}$.

Solution:

We recognize that this is the Fundamental Log Limit, which we used Squeeze Theorem to solve in Section 2.6. Now, let's use LHR instead.

$$\begin{aligned} &\lim_{x\to\infty} \frac{\ln(x)}{x} && \left(\text{type } \tfrac{\infty}{\infty}\right) \\ &\overset{\text{LHR}}{=} \lim_{x\to\infty} \frac{\frac{1}{x}}{1} \\ &= 0 \end{aligned}$$

We recover our expected result much quicker with LHR.

4.8.3 Type $0 \times \pm\infty$

Here, we cannot apply LHR directly. Our strategy for this type of indeterminate form will be to rewrite the multiplication as the division by the reciprocal of one of the functions. By doing this, we will recover a $\dfrac{0}{0}$ or $\pm\dfrac{\infty}{\infty}$ indeterminate form to which we can apply LHR.

Example 4.8.8 Evaluate $\displaystyle\lim_{x\to 0^+} x\ln(x)$.

Solution:

$$\begin{aligned} &\lim_{x\to 0^+} x\ln(x) && (\text{type } 0 \times -\infty) \\ &= \lim_{x\to 0^+} \frac{\ln(x)}{\frac{1}{x}} && \left(\text{type } \tfrac{-\infty}{\infty}\right) \\ &\overset{\text{LHR}}{=} \lim_{x\to 0^+} \frac{\frac{1}{x}}{-\frac{1}{x^2}} \\ &= \lim_{x\to 0^+} -x \\ &= 0 \end{aligned}$$

REMARK

We should generally try to simplify our limit after applying LHR, before applying it again. In the previous example, we *could* have used LHR again on $\lim\limits_{x\to 0^+} \frac{\frac{1}{x}}{-\frac{1}{x^2}}$ as this is a $\frac{\infty}{-\infty}$ type, but this would just lead to more indeterminate limits.

Example 4.8.9

Evaluate $\lim\limits_{x\to 0^+} xe^{1/x}$.

Solution:

$$\begin{aligned}
&\lim_{x\to 0^+} xe^{1/x} && (\text{type } 0\times\infty)\\
&= \lim_{x\to 0^+} \frac{e^{1/x}}{\frac{1}{x}} && (\text{type } \tfrac{\infty}{\infty})\\
&\overset{\text{LHR}}{=} \lim_{x\to 0^+} \frac{e^{1/x}\left(-\frac{1}{x^2}\right)}{-\frac{1}{x^2}}\\
&= \lim_{x\to 0^+} e^{1/x}\\
&= +\infty
\end{aligned}$$

REMARK

It can be tricky to determine which function you want to move to the denominator. If you find that the limit gets worse when using LHR with one of the two functions as a reciprocal, you can always try the problem again with the other function as the reciprocal. Generally, you will want to choose the function whose reciprocal is simpler to differentiate.

Example 4.8.10

Evaluate $\lim\limits_{x\to\infty} xe^{-x}$.

Solution:

$$\begin{aligned}
&\lim_{x\to\infty} xe^{-x} && (\text{type } \infty\times 0)\\
&= \lim_{x\to\infty} \frac{x}{e^x} && (\text{type } \tfrac{\infty}{\infty})\\
&\overset{\text{LHR}}{=} \lim_{x\to\infty} \frac{1}{e^x}\\
&= 0
\end{aligned}$$

4.8.4 Type $\infty - \infty$

Again, we cannot apply LHR directly. Here, our strategy will be to combine our expression in to a single term in some way. This might involve rationalizing, factoring, simplifying, or similar. This will transform our limit into one of the three types previously discussed.

Example 4.8.11 Evaluate $\lim_{x\to\frac{\pi}{2}^-}(\sec(x)-\tan(x))$.

Solution:

$$\begin{aligned}
&\lim_{x\to\frac{\pi}{2}^-}(\sec(x)-\tan(x)) && (\text{type } \infty-\infty)\\
&=\lim_{x\to\frac{\pi}{2}^-}\left(\frac{1}{\cos(x)}-\frac{\sin(x)}{\cos(x)}\right)\\
&=\lim_{x\to\frac{\pi}{2}^-}\frac{1-\sin(x)}{\cos(x)} && (\text{type } \tfrac{0}{0})\\
&\overset{\text{LHR}}{=}\lim_{x\to\frac{\pi}{2}^-}\frac{-\cos(x)}{-\sin(x)}\\
&=0
\end{aligned}$$

Example 4.8.12 Evaluate $\lim_{x\to\infty}\left(\ln(3x)+\ln\left(\frac{17}{x+7}\right)\right)$.

Solution:

$$\begin{aligned}
&\lim_{x\to\infty}\left(\ln(3x)+\ln\left(\frac{17}{x+7}\right)\right) && (\text{type } \infty-\infty)\\
&=\lim_{x\to\infty}\ln\left(3x\times\frac{17}{x+7}\right)\\
&=\lim_{x\to\infty}\ln\left(\frac{51x}{x+7}\right)\\
&=\ln\left(\lim_{x\to\infty}\frac{51x}{x+7}\right) && (\text{since } \ln(x) \text{ is continuous} \quad \big(\text{type } \tfrac{\infty}{\infty}\big))\\
&\overset{\text{LHR}}{=}\ln\left(\lim_{x\to\infty}\frac{51}{1}\right)\\
&=\ln(51)
\end{aligned}$$

REMARK

We learned that we could distribute a limit inside a continuous function in Lemma 2.8.16. This will be a common strategy in the remainder of this section.

4.8.5 Types 1^∞, ∞^0, and 0^0

Once again, we cannot apply LHR directly to these types. Here, our strategy will be to rewrite our limit in the form $e^{\ln(\text{indeterminate form})}$. Thanks to log laws, the exponent will transform into the form $0\times\pm\infty$. We can then use continuity of e^x to distribute the limit into the exponent and then implement what we learned for the previous indeterminate forms.

Example 4.8.13

Evaluate $\lim\limits_{x\to 0^+} x^x$.

Solution:

$$\begin{aligned}
&\lim_{x\to 0^+} x^x && \text{(type } 0^0\text{)}\\
&= \lim_{x\to 0^+} e^{\ln(x^x)}\\
&= \lim_{x\to 0^+} e^{x\ln(x)}\\
&= e^{\lim\limits_{x\to 0^+} x\ln(x)} && \text{(since } e^x \text{ is continuous} \quad \text{(type } 0\times -\infty\text{))}\\
&\overset{\text{LHR}}{=} e^0 && \text{(we solved the limit in the exponent with LHR in Example 4.8.8)}\\
&= 1
\end{aligned}$$

Example 4.8.14

Evaluate $\lim\limits_{x\to\infty}\left(1+\dfrac{1}{x}\right)^x$.

Solution:

$$\begin{aligned}
&\lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x && \text{(type } 1^\infty\text{)}\\
&= \lim_{x\to\infty} e^{\ln\left(\left(1+\frac{1}{x}\right)^x\right)}\\
&= \lim_{x\to\infty} e^{x\ln\left(1+\frac{1}{x}\right)}\\
&= e^{\lim\limits_{x\to\infty} x\ln\left(1+\frac{1}{x}\right)} && \text{(since } e^x \text{ is continuous} \quad \text{(type } \infty\times 0\text{))}\\
&= e^{\lim\limits_{x\to\infty} \frac{\ln\left(1+\frac{1}{x}\right)}{\frac{1}{x}}} && \text{(type } \tfrac{0}{0}\text{)}\\
&\overset{\text{LHR}}{=} e^{\lim\limits_{x\to\infty} \frac{\frac{1}{1+\frac{1}{x}}\left(-\frac{1}{x^2}\right)}{-\frac{1}{x^2}}}\\
&= e^{\lim\limits_{x\to\infty} \frac{1}{1+\frac{1}{x}}}\\
&= e^1 = e
\end{aligned}$$

Example 4.8.15

Evaluate $\lim\limits_{x\to\frac{\pi}{2}^-} (\sec(x))^{\cos(x)}$.

Solution:

$$\begin{aligned}
&\lim_{x\to\frac{\pi}{2}^-} (\sec(x))^{\cos(x)} && \text{(type } \infty^0\text{)}\\
&= \lim_{x\to\frac{\pi}{2}^-} e^{\cos(x)\ln(\sec(x))}
\end{aligned}$$

$$= e^{\lim_{x\to\frac{\pi}{2}^-} \cos(x)\ln(\sec(x))} \qquad \text{(since } e^x \text{ is continuous} \quad \text{(type } 0\times\infty))$$

$$= e^{\lim_{x\to\frac{\pi}{2}^-} \frac{\ln(\sec(x))}{\sec(x)}} \qquad \text{(type } \tfrac{\infty}{\infty})$$

$$\overset{\text{LHR}}{=} e^{\lim_{x\to\frac{\pi}{2}^-} \frac{\frac{1}{\sec(x)}\sec(x)\tan(x)}{\sec(x)\tan(x)}}$$

$$= e^{\lim_{x\to\frac{\pi}{2}^-} \frac{1}{\sec(x)}}$$

$$= e^0 = 1$$

REMARK

It is acceptable and equivalent to calculate the limit of these three indeterminate forms of type $f(x)^{g(x)}$ by:

1. naming the limit L;
2. calculating $\ln(L) = \ln\left(\lim_{x\to a} f(x)^{g(x)}\right)$ by bringing the limit outside the logarithm via continuity, using log laws to bring down the power, and then applying previous indeterminate form strategies; and
3. exponentiating to get L from $\ln(L)$.

Section 4.8 Problems

4.8.1. Evaluate the following indeterminate forms using l'Hôpital's rule.

$$\text{a)}\ \lim_{x\to\pi}\frac{x\sin x}{x-\pi}, \qquad \text{b)}\ \lim_{x\to\frac{\pi}{2}}\frac{\ln(2-\sin x)}{\ln(1+\cos x)}, \qquad \text{c)}\ \lim_{x\to\infty}\frac{\ln x}{\sqrt{x}},$$

$$\text{d)}\ \lim_{x\to 0} x^{\varepsilon}\ln x \ \ (\varepsilon>0), \qquad \text{e)}\ \lim_{x\to\infty} x^r e^{-x} \ \ (r\in\mathbb{R}, r>0).$$

4.8.2. Evaluate

$$\lim_{h\to 0}\frac{f(x+h)-2f(x)+f(x-h)}{h^2},$$

where f is a twice-differentiable function.

4.8.3. What is wrong with the following use of l'Hôpital's rule:

$$\lim_{x\to 1} = \frac{x^3+x-2}{x^2-3x+2} = \lim_{x\to 1}\frac{3x^2+1}{2x-3} = \lim_{x\to 1}\frac{6x}{2} = 3$$

(The limit is actually -4).

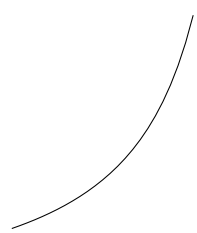
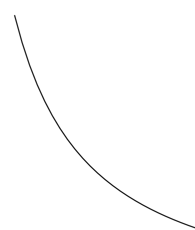
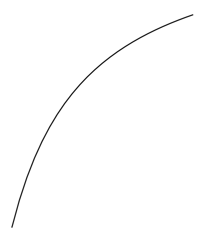
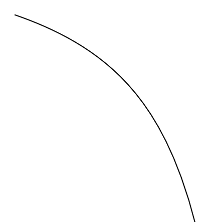
4.9 Curve Sketching

It's time now to tie everything together we've learned so far, via curve sketching. Curve sketching is akin to putting a puzzle together. Although time-consuming, you will see all of the pre-calculus and calculus knowledge that we have developed so far fit together. In the process outlined below, redundancies are built in so that you have extra information to ensure that you are on the right track.

Method

Curve Sketching Process

1. Identify function domain and possibly function values at endpoints.
2. Identify x- and y-intercepts.
3. Examine $\lim\limits_{x\to\pm\infty} f(x)$ and identify any horizontal asymptotes.
4. Identify any holes and vertical asymptotes by checking $\lim\limits_{x\to a^{\pm}} f(x)$ at domain issues.
5. Find all critical points ($f'(x) = 0$ or DNE).
6. Find all candidate points of inflection ($f''(x) = 0$ or DNE).
7. Investigate intervals of increase/decrease and concavity, divided by points in Steps 5 and 6 and discontinuities.
8. Identify local extrema and points of inflection from info in Step 7.
9. Plot all x-intercepts, y-intercepts, asymptotes, extrema, and points of inflection and label them.
10. Connect the dots in Step 9 using info from Step 7:

	increasing	decreasing
concave up		
concave down		

Example 4.9.1

Sketch the curve $f(x) = \dfrac{x^5 - 16x^3}{x}$.

Solution:

1. The domain of the function is $x \in (-\infty, 0) \cup (0, \infty)$ due to the issue in the denominator.

2. We note that $f(x) = 0$, when the numerator is zero. That is, when

$$x^5 - 16x^3 = x^3(x^2 - 16) = x^3(x-4)(x+4) = 0.$$

With the domain restriction in mind, we note then that there are x-intercepts at $(4, 0)$ and $(-4, 0)$.

We further note that due to the domain restriction, there are no y-intercepts.

3. We have that $\lim\limits_{x\to-\infty} \dfrac{x^5 - 16x^3}{x} = \lim\limits_{x\to-\infty} x^4 - 16x^2 = \lim\limits_{x\to-\infty} x^2(x^2 - 16) = \infty$.

And, $\lim\limits_{x\to\infty} \dfrac{x^5 - 16x^3}{x} = \lim\limits_{x\to\infty} x^4 - 16x^2 = \lim\limits_{x\to\infty} x^2(x^2 - 16) = \infty$.

Thus, there are no horizontal asymptotes.

4. We seek to investigate the domain issue at $x = 0$.

We find that $\lim\limits_{x\to 0^-} \dfrac{x^5 - 16x^3}{x} = \lim\limits_{x\to 0^-} x^4 - 16x^2 = 0$.

We also see that $\lim\limits_{x\to 0^+} \dfrac{x^5 - 16x^3}{x} = \lim\limits_{x\to 0^+} x^4 - 16x^2 = 0$.

Thus, given that the one-sided limits are equal, we have that $\lim\limits_{x\to 0^+} f(x) = 0$ while $f(0)$ does not exist. This is characteristic of a hole.

So, we have a hole at $(0, 0)$.

5. We note that $f(x) = x^4 - 16x^2$, $x \neq 0$. Thus, $f'(x) = 4x^3 - 32x$, $x \neq 0$.

While $f'(0)$ is not defined, $x = 0$ would not be a critical point, since $f(x)$ is not defined there either. However, we will still investigate increase/decrease on either side of $x = 0$.

So, we are only concerned with when $f'(x) = 0$. That is, when

$$4x^3 - 32x = 4x(x^2 - 8) = 4x(x - \sqrt{8})(x + \sqrt{8}) = 0$$

With the domain restriction in mind, we note then that there are critical points at $(\sqrt{8}, -64)$ and $(-\sqrt{8}, -64)$.

6. Given that $f'(x) = 4x^3 - 32x$, $x \neq 0$, we find that $f''(x) = 12x^2 - 32$, $x \neq 0$.

While $f''(0)$ is not defined, $x = 0$ would not be a candidate point of inflection, since $f(x)$ is not defined there either. However, we will still investigate concavity on either side of $x = 0$.

So, we are only concerned with when $f''(x) = 0$. That is, when

$$12x^2 - 32 = 4(3x^2 - 8) = 4\left(x - \sqrt{\frac{8}{3}}\right)\left(x + \sqrt{\frac{8}{3}}\right) = 0.$$

We note then that there are candidate points of inflection at $\left(\sqrt{\frac{8}{3}}, -\frac{320}{9}\right)$ and $\left(-\sqrt{\frac{8}{3}}, -\frac{320}{9}\right)$.

7./8. We now conduct testing:

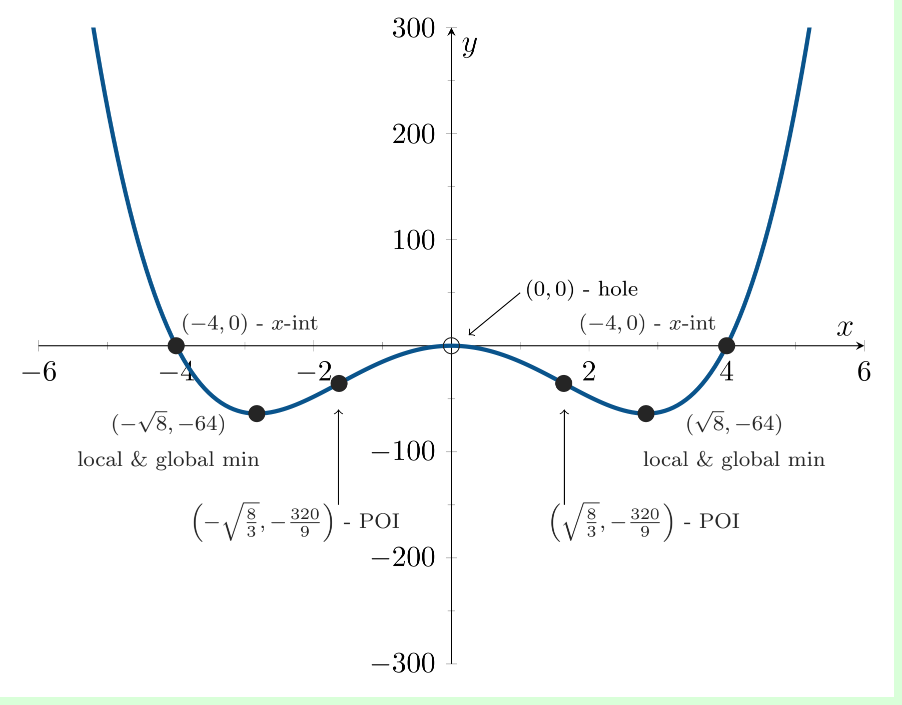
x		$-\sqrt{8}$		0		$\sqrt{8}$	
$f'(x)$	$-$	0	$+$	DNE	$-$	0	$+$

We see that $f(x)$ is increasing on $[-\sqrt{8},0)$ and $[\sqrt{8},\infty)$ and decreasing on $(-\infty,-\sqrt{8}]$ and $(0,\sqrt{8}]$.
By FDT, there are local minimums at $(\sqrt{8},-64)$ and $(-\sqrt{8},-64)$.

x		$-\sqrt{\frac{8}{3}}$		0		$\sqrt{\frac{8}{3}}$	
$f''(x)$	$+$	0	$-$	DNE	$-$	0	$+$

We see that $f(x)$ is concave down on $\left[-\sqrt{\frac{8}{3}},0\right)$ and $\left(0,\sqrt{\frac{8}{3}}\right]$ and concave up on $\left(-\infty,-\sqrt{\frac{8}{3}}\right]$ and $\left[\sqrt{\frac{8}{3}},\infty\right)$.
So there are confirmed points of inflection at $\left(\sqrt{\frac{8}{3}},-\frac{320}{9}\right)$ and $\left(-\sqrt{\frac{8}{3}},-\frac{320}{9}\right)$.

9./10. Putting all of the information together, we arrive at the sketch:



Example 4.9.2 Sketch the curve $f(x)=xe^{-2x}$ where $f'(x)=(1-2x)e^{-2x}$ and $f''(x)=(4x-4)e^{-2x}$.

Solution:

1. The domain of the function is $x\in(-\infty,\infty)$.
2. We note that $f(0)=0$, so there is an x-intercept at $(0,0)$. We further note that due to the exponential, $f(x)=0$ only when $x=0$, so the only y-intercept is the one found previously, $(0,0)$.
3. We have that $\lim\limits_{x\to-\infty} xe^{-2x}$ is of the form $-\infty\times\infty$, so $\lim\limits_{x\to-\infty} xe^{-2x}=-\infty$. And, $\lim\limits_{x\to\infty} xe^{-2x}$ is of the form $\infty\times 0$, so we rewrite:

$$\lim_{x\to\infty} xe^{-2x}=\lim_{x\to\infty}\frac{x}{e^{2x}}\overset{\text{LHR}}{=}\lim_{x\to\infty}\frac{1}{2e^{2x}}=0$$

so there is a horizontal asymptote at $y = 0$ for $x \to \infty$.

4. We note that the domain of $f(x)$ is $x \in (-\infty, \infty)$, so there are no holes nor vertical asymptotes to check for.
5. We see that $f'(x)$ is defined everywhere, and that $f'(x) = 0$ only when $x = \frac{1}{2}$. So, there is a critical point at $\left(\frac{1}{2}, \frac{1}{2}e^{-1}\right)$.
6. We note that $f''(x)$ is defined everywhere, and that $f''(x) = 0$ only when $x = 1$. So, there is a candidate point of inflection at $(1, e^{-2})$.

7./8. We now conduct testing:

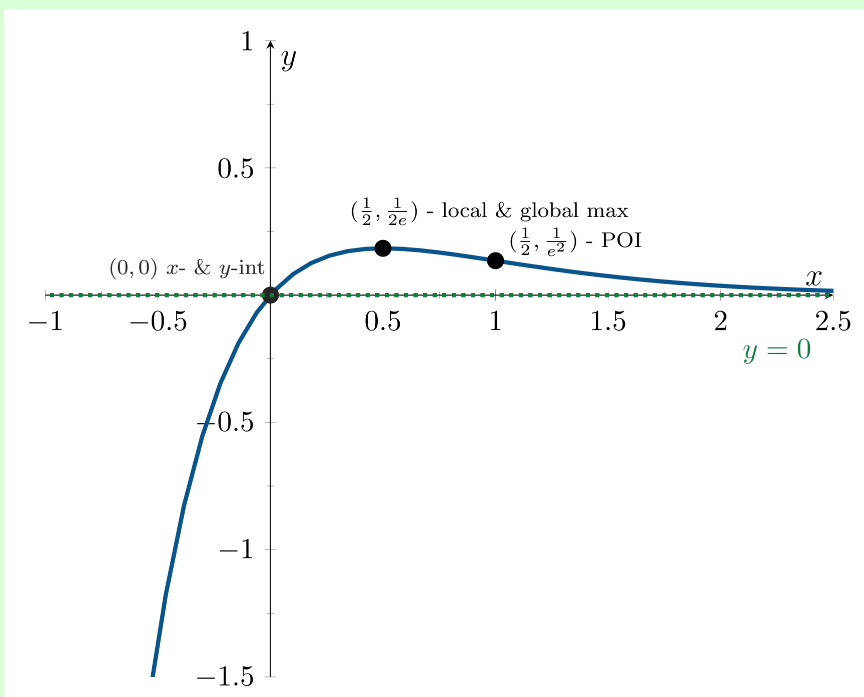
x		$\frac{1}{2}$	
$f'(x)$	$+$	0	$-$

We see that $f(x)$ is increasing on $(-\infty, \frac{1}{2}]$ and decreasing on $[\frac{1}{2}, \infty)$.
So by FDT, there is a local max at $\left(\frac{1}{2}, \frac{1}{2}e^{-1}\right)$.

x		1	
$f''(x)$	$-$	0	$+$

We see that $f(x)$ is concave down on $(-\infty, 1]$ and concave up on $[1, \infty)$.
So there is a point of inflection at $(1, e^{-2})$.

9./10. Putting all of the information together, we arrive at the sketch:



Example 4.9.3 Sketch the curve $f(x) = \dfrac{x^2}{x^2 - 4}$ where $f'(x) = \dfrac{-8x}{(x^2 - 4)^2}$ and $f''(x) = \dfrac{8\left(3x^2 + 4\right)}{(x^2 - 4)^3}$.

Solution:

1. The domain of the function is $x \in (-\infty, -2) \cup (-2, 2) \cup (2, \infty)$ due to the issue in the denominator.
2. We note that $f(0) = 0$, when the numerator is zero. So, we have our only x- and y-intercept at $(0, 0)$.
3. We have that $\lim\limits_{x \to -\infty} \dfrac{x^2}{x^2 - 4} = \lim\limits_{x \to -\infty} \dfrac{x^2(1)}{x^2(1 - \frac{4}{x^2})} = \lim\limits_{x \to -\infty} \dfrac{1}{1 - \frac{4}{x^2}} = 1.$

   Similarly, $\lim\limits_{x \to \infty} \dfrac{x^2}{x^2 - 4} = 1.$

   Thus, there is a horizontal asymptote at $y = 1$ for $x \to \pm\infty$.
4. We seek to investigate the domain issues at $x = \pm 2$.

   We find that $\lim\limits_{x \to 2^-} \dfrac{x^2}{x^2 - 4}$ is of the form $\frac{4}{0^-}$ so $\lim\limits_{x \to 2^-} \dfrac{x^2}{x^2 - 4} = -\infty$.

   This is sufficient information to conclude there is a vertical asymptote at $x = 2$, but let's gather a bit more information at $x = 2$ that will help us sketch.

   We find that $\lim\limits_{x \to 2^+} \dfrac{x^2}{x^2 - 4}$ is of the form $\frac{4}{0^+}$ so $\lim\limits_{x \to 2^+} \dfrac{x^2}{x^2 - 4} = +\infty$.

   Moving on, we find that $\lim\limits_{x \to -2^-} \dfrac{x^2}{x^2 - 4}$ is of the form $\frac{4}{0^+}$ so $\lim\limits_{x \to -2^-} \dfrac{x^2}{x^2 - 4} = +\infty$.

   This is sufficient information to conclude there is a vertical asymptote at $x = -2$, but let's gather a bit more information at $x = -2$ that will help us sketch.

   We find that $\lim\limits_{x \to -2^+} \dfrac{x^2}{x^2 - 4}$ is of the form $\frac{4}{0^-}$ so $\lim\limits_{x \to -2^+} \dfrac{x^2}{x^2 - 4} = -\infty$.
5. While $f'(\pm 2)$ are not defined, $x = \pm 2$ would not be critical points, since $f(x)$ is not defined there either. We will still investigate whether the function is increasing/decreasing on the intervals around these points.

   So, we are only concerned with when $f'(x) = 0$. This occurs when the numerator, $-8x$ is zero. This occurs at $x = 0$, so there is a critical point at $(0, 0)$.
6. While $f''(\pm 2)$ are not defined, $x = \pm 2$ would not be candidate points of inflection, since $f(x)$ is not defined there either.

   So, we are only concerned with when $f''(x) = 0$. However, the numerator of $f''(x)$ is strictly positive, so this never occurs. Thus, there are no candidate points of inflection, and by extension, no points of inflection.

7./8. We now conduct testing:

x		-2		0		2	
$f'(x)$	$+$	DNE	$+$	0	$-$	DNE	$-$

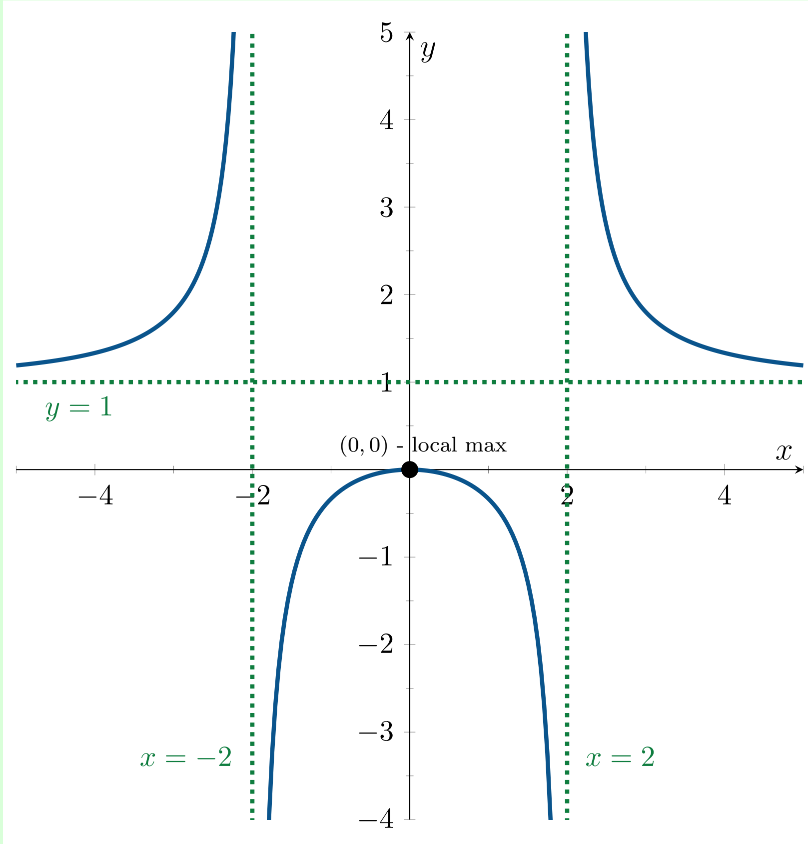
We see that $f(x)$ is increasing on $(-\infty, -2)$ and $(-2, 0]$ and decreasing on $[0, 2)$ and $(2, \infty)$.

By FDT, there is a local maximum at $(0, 0)$.

x		-2		2	
$f''(x)$	$+$	DNE	$-$	DNE	$+$

We see that $f(x)$ is concave down on $(-2,2)$ and concave up on $(-\infty,-2)$ and $(2,\infty)$.

9./10. Putting all of the information together, we arrive at the sketch:



Section 4.9 Problems

4.9.1. For each of the following functions perform all the steps we outlined to sketch the curve with calculus, and then sketch the curve. This includes listing the domain, any asymptotes, intervals of increase/decrease, intervals of concavity, and along with the sketch plotting any critical points, points of inflection, and intercepts.

(a) $f(x) = x^3 - 6x^2 + 9x$.

(b) $f(x) = \dfrac{x^2+1}{x^2-9}$.

(c) $f(x) = \ln(18 - 2x^2)$.

4.9.2. Given the function $f(x) = \frac{1}{x^3-x}$ which of the following properties are true:

(a) There is a point of inflection at $x = -1$.

(b) f is concave up from $[2, 5]$.

(c) $\displaystyle\sum_{c\in D} f(c) = 0$ where D is the set of local extrema.

(d) f is decreasing on $[\frac{1}{\sqrt{3}}, \infty)$.

(e) f is concave down from $[0, 1]$.

(f) There are more vertical asymptotes than horizontal asymptotes.

(g) f has no x or y intercepts.

(h) $f(-x) = f(x)$ for all x in the domain of f.

(i) Let $b > 1$ and $a < -1$, there is a $c \in (-1, 1)$ such that $f'(c) = \frac{f(b)-f(a)}{b-a}$.

4.10 Optimization

You're almost there: you've very nearly *maximized* your differential calculus knowledge. Soon, your time spent reading these course notes will be *minimized*. In this final section, we look at the practical application of optimization.

As mentioned at the end of Section 4.2, many real-world optimization problems are not closed interval method friendly, due to the occasional lack of closed intervals. In these situations, we will be able to apply our newly minted knowledge of the First and/or Second Derivative Tests to justify the extrema we find. We will also see how optimization problems generally come with some kind of *constraint*.

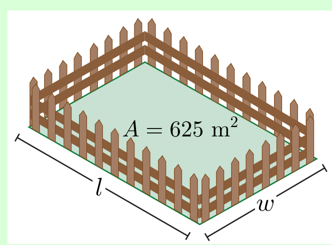
We begin with a classic fencing (the less sporty kind) example, before discussing a general strategy to optimization problems.

Example 4.10.1 Graeme the Groundskeeper is looking to build a new enclosure for Andrew the Aardvark. To make sure that Andrew is comfortable, Graeme wants to ensure the enclosure has an area of 625 m^2. If fencing costs \$15 per metre, what dimensions of rectangular enclosure will minimize cost? What is this minimum cost?

Solution:

In this problem, we identify that the key optimization word is *minimize*. In particular, we are seeking to find the global minimum of the cost of the enclosure.

We'll start by drawing a quick diagram of an enclosure as described in the problem.



[Fence] - Ket4up/iStock via Getty Images

Now, we focus on the constraints we have been given in the problem. We are told that the area of the rectangular enclosure must be 625 m^2, which gives us $wl = 625$. Given that the enclosure must physically exist, we know that the width and length of the enclosure must be positive. So, this gives us that $w \in (0, \infty)$ and $l \in (0, \infty)$.

Next, we build the function which we seek to minimize, the cost function. The cost is based on the perimeter of the enclosure, where each metre of the perimeter costs \$15. Thus, we have that $C = 15(2w + 2l) = 30w + 30l$.

The concerning feature of the cost function is that it is of two variables! Thankfully, we can make use of the constraint that $wl = 625$ to rewrite the cost function in one variable.

We can rewrite the constraint as $w = \dfrac{625}{l}$ (with no division by zero concern given the constraint on l). Note that you could also have isolated for l and proceeded.

Substituting this into the cost function gives that $C = 30\left(\dfrac{625}{l}\right) + 30l = \dfrac{18750}{l} + 30l$.

So, to reframe our problem, we want to find the global minimum of $C(l) = \dfrac{18750}{l} + 30l$ for $l \in (0, \infty)$. We see here explicitly that the closed interval method would not be applicable.

Having set up our problem, it is now time to whack it with calculus. We should first find the critical points, which occur where $C'(l) = 0$ or does not exist.

Taking the derivative, we find that $C'(l) = -\dfrac{18750}{l^2} + 30$.

Although $C'(0)$ does not exist, this is not a critical point, as it is not in the domain of the function (it is also outside the constraint on l).

So we now proceed to determine where $C'(l) = 0$:

$$\begin{aligned} 0 &= -\frac{18750}{l^2} + 30 \\ 18750 &= 30l^2 \\ l &= \pm\sqrt{\frac{18750}{30}} \\ l &= \pm 25 \end{aligned}$$

However, given our constraint, we discard $l = -25$. Thus, our only critical point occurs when $l = 25$.

While you might be excited to find that there is only one critical point and be ready to declare it our global minimum, you must *justify* that you have actually found a minimum!

Here, you have your choice of FDT or SDT. In this example, we will proceed to use SDT, beginning with finding that the second derivative is $C''(l) = \dfrac{37500}{l^3}$.

For the SDT, we test *at* the critical point, finding that $C''(25) > 0$, meaning that $C(l)$ is concave up around $l = 25$. SDT thus tells us that we have indeed found a local minimum at $l = 25$.

We note that there are no endpoints to consider as candidates for global extrema. Further, we note that $C''(l) > 0$ for all $l > 0$, so the graph is *always* concave upwards on our domain.

Thus, there is in fact a global minimum where $l = 25$.

Alas, we are not yet finished the problem! We were asked to give both the dimensions of minimum cost, and the minimum cost itself.

Returning to the fact that $w = \dfrac{625}{l}$, we find $w = \dfrac{625}{25} = 25$. And, inputting these into our cost function, we find that $C = 30(25) + 30(25) = 1500$.

Therefore, the dimensions for the minimum cost enclosure are 25 m by 25 m, and the minimum cost is \$1500.

Here now is a general strategy for tackling applied optimization problems.

Method

Approach for Applied Optimization Problems

1. Determine what needs to be optimized and in which fashion (max or min).

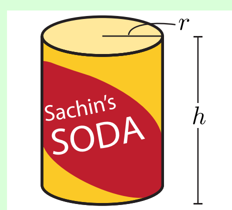
2. Draw a picture of the situation (if applicable) and introduce variables.
3. Determine all constraints.
4. Create the function to be optimized, the objective function.
5. Use the constraints to write the objective function in one variable.
6. Use calculus to optimize the objective function of one variable.
   (a) Ensure you test the critical points with FDT/SDT to classify them.
   (b) Keep the constraints in mind.
   (c) Check endpoints if applicable.
7. State your conclusions, ensuring appropriate real-world units.

Example 4.10.2 Sachin's Soda has teamed up with Burcu's Beverage Bottling to redesign the can for their most popular pop. The new cans are to be cylindrical, with a fixed volume of V cm^3. Metal for the top, bottom, and sides of the can is at a fixed cost of c cents per cm^2. Determine the radius, r, of the top and bottom of the can, and the height, h, of the can that will minimize the cost.

Solution:

Who said math problems needed to involve any numbers?!

In this problem, we are seeking to minimize the cost of the new can. We draw a quick sketch of the new can with relevant labels.



We now focus on identifying our constraints. We are told that the volume is fixed, giving us that $\pi r^2 h = V$. We also need for the can to physically exist, so we have that $r \in (0, \infty)$ and $h \in (0, \infty)$.

Now, the cost of the can is based on multiplying a fixed cost, c, per cm^2 by the surface area of the can. Because of this, we can make our lives very slightly easier by realizing that minimizing the cost is thus equivalent to minimizing the surface area of the can. Then, our objective function is $S = 2\pi r^2 + 2\pi r h$.

To make this a function of one variable, we use one of our constraints to note that $h = \dfrac{V}{\pi r^2}$.

Thus, our objective function can be written as $S(r) = 2\pi r^2 + 2\pi r \dfrac{V}{\pi r^2} = 2\pi r^2 + \dfrac{2V}{r}$, for which we seek to find a global minimum for $r \in (0, \infty)$.

We begin the process of finding critical points by finding that $S'(r) = 4\pi r - \dfrac{2V}{r^2}$. Now, although $S'(0)$ does not exist, this is not a critical point due to our constraint. So, we focus on finding where $S'(r) = 0$.

$$\begin{aligned} 0 &= 4\pi r - \frac{2V}{r^2} \\ \frac{2V}{r^2} &= 4\pi r \\ r &= \sqrt[3]{\frac{2V}{4\pi}} = \sqrt[3]{\frac{V}{2\pi}} \end{aligned}$$

To classify this critical point, we will appeal to SDT, first noting that $S''(r) = 4\pi + \dfrac{4V}{r^3}$. Now, since $\sqrt[3]{\dfrac{V}{2\pi}}$ is positive, we note then that $S''\left(\sqrt[3]{\dfrac{V}{2\pi}}\right) > 0$, which tells us by SDT that there is a local minimum there.

There are no endpoints to consider here, and we note that $S''(r) > 0$ f‘or all $r > 0$, meaning the graph is always concave upwards on our domain, and therefore we have found the global minimum.

We now must find the related height, for which we will appeal to the fact that $h = \dfrac{V}{\pi r^2}$. This gives us that $h = \dfrac{V}{\pi\left(\sqrt[3]{\frac{V}{2\pi}}\right)^2} = 2\sqrt[3]{\dfrac{V}{2\pi}}$.

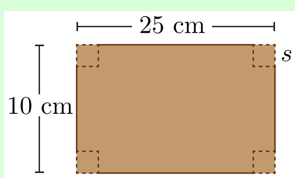
Therefore, the radius that minimizes the cost of the can is $r = \sqrt[3]{\dfrac{V}{2\pi}}$ and the height that minimizes the cost of the can is $h = 2\sqrt[3]{\dfrac{V}{2\pi}}$.

Of particular interest, note that in this optimal solution, $h = 2r$.

Example 4.10.3 Peter’s Packaging is working on designing a new, cutting-edge box. Peter takes a piece of cardboard, which has dimensions 25 cm by 10 cm, and cuts out squares from each corner. He then folds the net he has made to create an open-top box. What side length of square should Peter cut from each corner to maximize the volume of the resulting box?

Solution:

In this problem, we want to maximize the volume of the box. We draw a picture below of what the cardboard net looks like.



We see in our diagram that the height of the resulting box will be the side length of square we cut out, s, and that the dimensions of the base of the box will be $25 - 2s$ and $10 - 2s$.

As we need the box to physically exist, we must have that $s \in (0, \infty)$, $25 - 2s \in (0, \infty)$, and $10 - 2s \in (0, \infty)$.

Cleaning up the second and third of these constraints, we simultaneously require $s \in (0, \infty)$, $s \in (-\infty, \frac{25}{2})$, and $s \in (-\infty, 5)$.

The intersection of these three constraints is that we must have $s \in (0, 5)$.

Our objective function is the volume of the box, which is given by

$$V(s) = s(25 - 2s)(10 - 2s) = 4s^3 - 70s^2 + 250s$$

for which we aim to find a global maximum subject to combined constraint $s \in (0, 5)$.

We note that $V'(s) = 12s^2 - 140s + 250$ is a polynomial, so exists everywhere. Thus, we seek only to find the zeros of this quadratic, which via the quadratic formula yields roots at $s = \dfrac{5}{6}(7 \pm \sqrt{19})$.

A bit of calculation work would demonstrate that $s = \dfrac{5}{6}(7 + \sqrt{19})$ violates the constraint while the other root, $s = \dfrac{5}{6}(7 - \sqrt{19})$ satisfies the constraint. Thus, this is our only critical point.

For the sake of utilizing FDT, we will note that $s = \dfrac{5}{6}(7 - \sqrt{19}) \approx 2.201$.

We now conduct testing, perhaps using the test values $s = 1$ and $s = 3$:

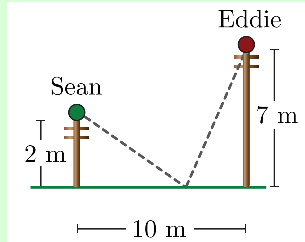
s		$\frac{5}{6}(7 - \sqrt{19})$	
$V'(s)$	$+$	0	$-$

We see that FDT confirms that this is the location of a local maximum.

There are no endpoints to consider as candidates for global maximum. We note that $V''(s) = 24s - 140 < 0$ for $s \in (0, 5)$, so the graph is concave downwards on this interval. Thus we have found the global maximum.

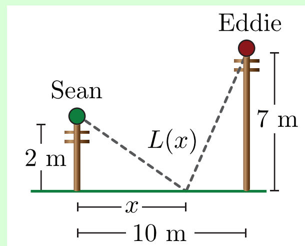
Therefore, Peter should cut $\dfrac{5}{6}(7 - \sqrt{19}) \approx 2.201$ cm sided squares from each corner to maximize the volume of the box.

Example 4.10.4 Sean the Sparrow and Eddie the Eagle are resting in their nests on the top of two telephone poles 10 metres apart, after a long day of flapping around. Sean's pole is 2 metres tall, and Eddie's pole is 7 metres tall. A wire has been attached to the top of both poles and to the ground somewhere between the poles, including possibly at the base of either pole, as shown in the figure below. Determine how far away from the base of Sean's pole the wire should be attached to the ground in order to minimize the length of wire required.



Solution:

Our goal in this problem is to minimize the length of the wire. We add some labels to the situation.



In the diagram, we have labelled the distance from Sean's pole to the place where the wire is attached to the ground as x. This means the distance from the ground point of the wire to Eddie's pole is $10 - x$. Our constraints are thus $x \in [0, \infty)$ and $10 - x \in [0, \infty)$. The intersection of these constraints leads to $x \in [0, 10]$.

We see that the length of the wire is made up of two segments. We can find these lengths via the Pythagorean Theorem. Doing so would give us that

$$L(x) = \sqrt{2^2 + x^2} + \sqrt{7^2 + (10 - x)^2} = \sqrt{4 + x^2} + \sqrt{x^2 - 20x + 149}$$

for which we seek a global minimum for $x \in [0, 10]$. While we could use the closed interval method here, we will still use our derivative tests for practice.

We find that $L'(x) = \dfrac{x}{\sqrt{4 + x^2}} + \dfrac{x - 10}{\sqrt{x^2 - 20x + 149}}$. The quadratics in both denominators have a negative discriminant, so they are both never zero. Thus, $L'(x)$ exists everywhere.

We now aim to find where $L'(x) = 0$.

$$\begin{aligned}
0 &= \frac{x}{\sqrt{4 + x^2}} + \frac{x - 10}{\sqrt{x^2 - 20x + 149}} \\
\frac{10 - x}{\sqrt{x^2 - 20x + 149}} &= \frac{x}{\sqrt{4 + x^2}} \\
(10 - x)\sqrt{4 + x^2} &= x\sqrt{x^2 - 20x + 149} \\
(10 - x)^2(4 + x^2) &= x^2(x^2 - 20x + 149) \\
x^4 - 20x^3 + 104x^2 - 80x + 400 &= x^4 - 20x^3 + 149x^2 \\
45x^2 + 80x - 400 &= 0
\end{aligned}$$

Then, via quadratic formula, this yields $x = -4$, which we discard due to our constraint, and $x = \dfrac{20}{9}$.

For the sake of utilizing FDT, we will note that $x = \dfrac{20}{9} \approx 2.222$.

We now conduct testing, perhaps using the test values $x = 1$ and $x = 3$:

x		$\frac{20}{9}$	
$L'(x)$	$-$	0	$+$

We see that the FDT confirms that this is the location of a local minimum.

Since we have endpoints, we find the value of the function at the endpoints and compare it with the value of the function at the local minimum. Doing so, we'd find

$$\begin{aligned} L(0) &\approx 14.207 \\ L\left(\frac{20}{9}\right) &\approx 13.454 \\ L(10) &\approx 17.198 \end{aligned}$$

Thus, we have confirmed that there is a global minimum at $x = \dfrac{20}{9}$.

Therefore, the wire should be attached to the ground $\dfrac{20}{9}$ metres away from the base of Sean's pole to minimize the length of the wire.

REMARK

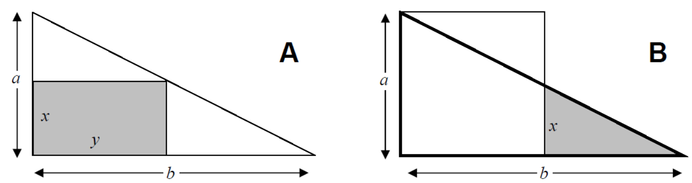
If the preceding example had asked us to maximize the length of wire required, we would have found this occured when the wire is attached to the ground 10 metres away from Sean's pole (ie. at the base of Eddie's pole). So, you should remember that it *is* still possible for the global extremum of interest to occur at an endpoint!

You've reached the end of MATH 137. You are probably feeling a little sad. This is okay. Sadness is a normal human emotion. We encourage you to watch the movie Inside Out 2 - one of the best movies of all time. Bye for now!

Section 4.10 Problems

4.10.1. A swimmer is 200 m out in the ocean and would like to arrive at a town 3 km down the coast. The swimmer will swim to shore, then jog along the beach to the town. She can swim at 2.5 km/h and jog at 10 km/h. To what point along the shore should she swim so that the total time it takes to get to town is a minimum? What about if the time it takes is a maximum?

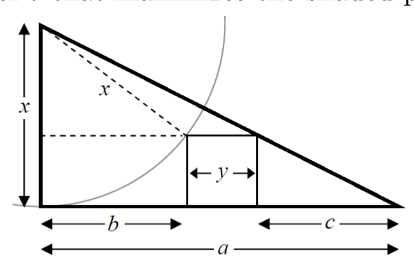
4.10.2. The following geometry problems appear on small wooden tablets hung in Japanese temples (Sangaku).



A. (For Ex. 2a) Given side lengths a and b, find x and y to maximize the area of the inscribed rectangle. **B.** (For Ex. 2b) Maximize the area of the shaded triangle as a function of a, assuming the base b is fixed.

(a) A rectangle shares two edges with the perpendicular sides of a right angled triangle; the fourth vertex lies on the hypotenuse. Find the side lengths of the inscribed rectangle that result in the maximum area

(b) A right-angled triangle, with perpendicular side-lengths a and b, shares an edge with a square of side-length a. Assume that the edge-length a is variable, and find the value of a in terms of b that maximizes the shaded portion of the triangle.

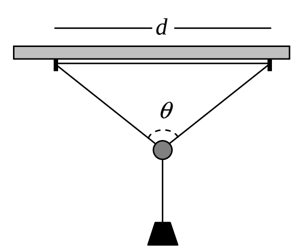


(For Ex. 2c) Maximize y as a function of x assuming that the base a is fixed.

(c) In a given right triangle, draw a circle of radius x corresponding to one of the perpendicular side-lengths, and centered at the vertex. Consider a square, one of whose sides lies on the other perpendicular side and touches both the circle and the hypotenuse. If y is the length of the side of the square, and the common side-length a is held constant, find the maximum value of y as a function of x.

4.10.3. Consider a rope of length ℓ with a mass W attached at one end, and a pulley attached at the other. The rope is suspended from two hooks (a distance d apart), and the weight is passed through the pulley (see figure below). After the weight is released and comes to equilibrium, determine the angle between ropes at the pulley, θ, assuming that the distance between the weight and the ceiling is a maximum. From the symmetry of the system, you can assume that the weight comes to rest half-way between the hooks.

Hint: First show that the distance of the weight from the ceiling is given by $L(\theta) = \ell - d - \dfrac{d}{\sin(\theta/2)} + \dfrac{d}{2\tan(\theta/2)}$.



(For Ex. 3) Assume the distance between the ceiling and the weight is a maximum at equilibrium.

4.10.4. Consider the function $f(x) = x^x$, $x \geq 0$.

(a) The point $x = 0$ is undefined in the original function definition; determine the value of $f(0)$ that ensures $f(x) = x^x$ is continuous at $x = 0$.

(b) Find the point $x^\star$ where the function achieves its minimum. Hint: $a = e^{\ln a}$ when $a > 0$.

4.10.5. What is the point on the parabola $y = x^2$ that is closest to point $(x_0, y_0) = (2, 1)$? Express your answer correct to two decimals. Hint: Minimizing the distance is equivalent to minimizing the distance-squared, simplifying the algebra.